Persistent Homology for the Analysis of Stratified Spaces

by

1079615

A thesis presented for the degree of

MSc (Mathematics and Foundations of Computer Science)

2024

Mathematical Institute

University of Oxford

Acknowledgements

First and foremost, I would like to thank my dissertation supervisor, Vidit Nanda, for granting me the freedom to explore a topic of my interest and for offering invaluable guidance throughout the process. Your brilliance as a mathematician has been a constant source of inspiration. Additionally, I would like to thank those who too often go unacknowledged. I'm deeply grateful to all the scouts, kitchen staff, porters, and members of the Magdalen College staff whose hard work afforded me the precious time I had to focus on my studies.

Abstract

This dissertation investigates the application of persistent homology to the analysis of stratified spaces, focusing on word embeddings as a case study. Stratified spaces present unique challenges for traditional topological data analysis techniques, particularly in identifying and analyzing singularities. To address these challenges, we discuss persistent intersection homology and extend the concepts of kernel, image, and cokernel persistence. We also present a novel approach to analyzing singularities using bifiltrations and sliding window (co)kernel diagrams. In addition, we discuss multiparameter persistence. These tools allow for a more nuanced analysis, enabling the differentiation of various types of singularities, such as those associated with polysemous words in word embeddings.

We demonstrate the practical utility of these theoretical advancements through computational pipelines and experiments on word embeddings, specifically analyzing how dimensionality reduction techniques influence the topology of local neighborhoods. The results indicate that our methods can reveal significant structural differences in word embeddings, offering new insights into their topological properties.

This work advances the theoretical understanding of persistent homology in stratified spaces and opens new avenues for applying these techniques to other complex, high-dimensional datasets. We conclude with a discussion of potential future research directions, including extending these methods to multifiltrations, exploring different invariants, and broader applications beyond word embeddings.

Contents

1 Introduction

Many real-world, high-dimensional datasets are commonly assumed to reside on low-dimensional manifolds embedded within the high-dimensional space. Persistent homology can be successfully employed in such cases to extract features, remove noise, and compare data sets [23]. However, challenges arise when the data represents several manifolds of possibly varying dimensions.

For example, when considering word embeddings, it has been argued that word vectors lie on a pinched manifold, that is the quotient of a manifold obtained by identifying some of its points [38]. These singular points correspond to polysemous words, i.e., words with multiple meanings, which suggests that monosemous and polysemous words can be differentiated based on the topology of their neighborhoods.

Persistent homology, however, is often unable to detect the topology of such singular neighborhoods. For instance, when computing the Vietoris-Rips complex of a point cloud, it becomes impossible to identify singularities using only homological methods [63]. Furthermore, certain desirable properties, such as Poincaré duality, generally fail for spaces that are not a manifold.

Some of the challenges mentioned earlier can be addressed by loosening the assumption that data resides on a single manifold and instead considering that data may exist in a stratified space. A stratified space is, in simple terms, a collection of manifolds with potentially different dimensions that are "glued together" in a well-behaved manner. However, fewer methods are available for analyzing data under this assumption.

This dissertation aims to bridge the gap between topological data analysis and the study of stratified spaces. It is organized into three main sections. First, we provide the theoretical background necessary for our work, offering an accessible and intuitive introduction to stratified spaces and intersection homology. We also discuss kernel, image, and cokernel persistence and demonstrate how they can be used to analyze singularities. Additionally, we provide a brief overview of multiparameter persistence. Next, we present computational methods for analyzing singularities in datasets using the theoretical framework outlined in the previous section. Finally, we apply these techniques to the study of word embeddings.

1.1 Contributions

We present several key contributions that advance theoretical understanding and computational techniques in the study of stratified spaces using persistent homology.

Theoretical Contributions

1. We introduce a general partial order on perversities, which induces an inclusion of intersection chain complexes and subsequently leads to a homomorphism on intersection homology.

2. We propose the concepts of kernel, image, and cokernel persistence as novel tools for constructing a bifiltration to analyze singularities.

3. We develop the idea of sliding window (co)kernel diagrams to quantify the interaction of singularities with the persistent homology of a space, thereby enabling the differentiation of various types of singularities.

4. We present a new method for testing possible stratifications of data using kernel persistence

Computational Contributions

We introduce three computational approaches for analyzing singularities within datasets:

1. We use kernel, image, and cokernel persistence to identify which singularities significantly impact the persistent homology of the data.
2. We apply multiparameter persistence to construct candidate decompositions and compute the invariants of the resulting bifiltration.
3. We leverage multiparameter persistence-based clustering and persistent intersection homology to derive more refined topological invariants of stratified data.

Finally, we demonstrate the practical application of our methods through a case study on word embeddings. We investigate and compare the local neighborhoods of two words, examine the effect of dimensionality reduction on these neighborhoods, and differentiate between their potential number of senses.

2 Theory

We begin by covering relevant theoretical background, starting with an introduction to stratified spaces, intersection homology, and persistent intersection homology. We then discuss kernel, image, and cokernel persistence. Finally, we briefly introduce multiparameter persistence.

2.1 Stratified Spaces

Recall that a smooth manifold M is second countable Hausdorff topological space such that for every point q in M, there exists a neighborhood U of q with a homeomorphism ϕ mapping U to an open subset of $\mathbb{R}^n$.

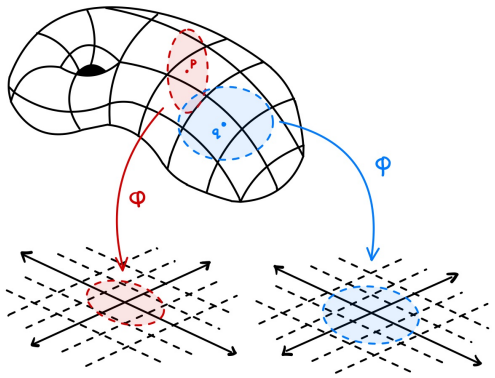


Figure 1: For every point q in M, there exists a neighborhood U of q with a homeomorphism ϕ mapping U to an open subset of $\mathbb{R}^n$.

In other words, a manifold is a topological space that locally resembles Euclidean space near each point.

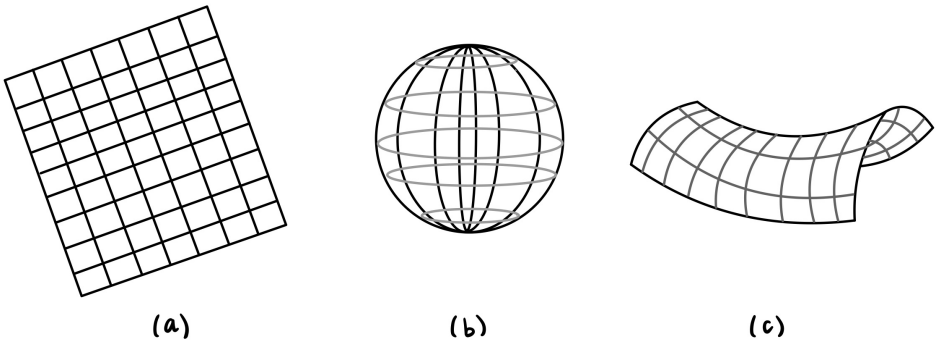


Figure 2: Some manifolds

If a topological space fails to be a manifold, it is because of the existence of singular points that do not have Euclidean neighborhoods. For instance, consider the figure-8. It is not a 1-manifold because the singular crossing point does not have a neighborhood homeomorphic to $\mathbb{R}$. However, every other point in the figure-8 does have such a neighborhood. The figure-8 minus the crossing point is a 1-manifold, while the crossing point itself is a 0-manifold.

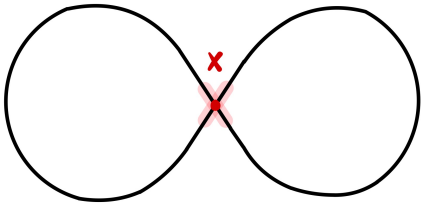


Figure 3: The figure-8

Stratified spaces generalize the concept of a manifold to address singular spaces. A stratified space is a topological space that can be decomposed into manifold pieces of different dimensions (called *strata*), which "fit together nicely." In general, a **stratification** of a topological space $\mathbb{X}$ is a filtration by closed subsets,

$$\mathbb{X} = \mathbb{X}_m \supseteq \mathbb{X}_{m-1} \supseteq \ldots \supseteq \mathbb{X}_0 \supseteq \mathbb{X}_{-1} = \emptyset$$

such that $\mathbb{X}_i - \mathbb{X}_{i-1}$ is a (possibly empty) manifold for each i. The connected components of the set $\mathbb{X}_i - \mathbb{X}_{i-1}$ are called **strata**. In this context, the dimension of points can vary, which offers more flexibility than working with manifolds. However, the restrictions we apply ensure that we focus on a relatively "nice" class of spaces, which are well-behaved enough to exclude things like fractals. Given their ability to describe a broader range of spaces, it has been argued that stratified spaces are ideal for analyzing real-world data [72].

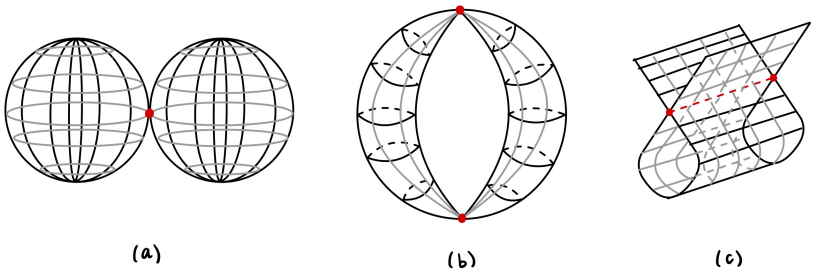


Figure 4: Some stratified spaces

2.1.1 Historical Remarks

Partitioning a singular space into manifold-like pieces is not a new idea. However, the study of how these strata fit together and the geometry of their neighborhoods is a more recent development [37]. Before 1960, few mathematicians paid much attention to singular spaces. During this period, there was a remarkable surge in results concerning manifolds, like Lefschetz theory, de Rham cohomology, Hodge theory, characteristic numbers, cobordism, the Hirzebruch-Riemann-Roch theorem, surgery, handlebodies, and the Atiyah-Singer index theorem. The standard belief at the time was that if you had a singularity, you should resolve it and get a manifold. [48].

However, spaces with singularities are both important and not always pathological. For example, any manifold with a boundary is a stratified space consisting of two strata: the boundary and the interior. Similarly, any polyhedron can be stratified in a reasonable way [73]. The natural idea of dividing a singular space into manifolds was at least partially realized in the study of simplicial complexes and regular cell complexes, even before the notion of a manifold was well defined [33]. In fact, there were several early attempts to triangulate algebraic sets, including Poincaré [61] and Lefschetz [44].

Though there are many different definitions of these kinds of spaces (for a survey see [37]), we will use the one used by Goresky and MacPherson in [31] which works most naturally with intersection homology, namely *piecewise linear (PL) stratified pseudomanifolds*. To make the technical aspects of the definition more intuitive, we first discuss some preliminary definitions and results.

2.1.2 Filtered Spaces

We introduce a series of definitions, starting with the most general: the filtered space.

Filtered Space [29]

A **filtered space** is a Hausdorff topological space X together with a sequence of closed subspaces

$$\emptyset = X_{-1} \subseteq X_0 \subseteq X_1 \subseteq X_2 \subseteq \cdots \subseteq X_{n-1} \subseteq X_n = X$$

for some integer $n \geq -1$.

The space X_i is called the **i-skeleton**. The index i is called the **formal dimension** of the skeleton. Note that it is possible to have $X_i = X_{i-1}$, the smallest index is always -1, and X_{-1} is always empty. The connected components of $X_i - X_{i-1}$ are called the **strata**

of X. If $S \subseteq X_i - X_{i-1}$ is a stratum of X, we say that S is a stratum of **dimension** i and **codimension** $n - i$.

Example Recall that the i**-skeleton** of a simplicial complex is the subspace X_i formed by taking the union of all the simplices in X that have dimensions up to i. Any finite-dimensional simplicial complex is a filtered space, filtered by the i-dimensional skeleton. In this setting, the strata are the open simplices (i.e. the interiors of the simplices) of X. For example, consider the simplicial complex below.

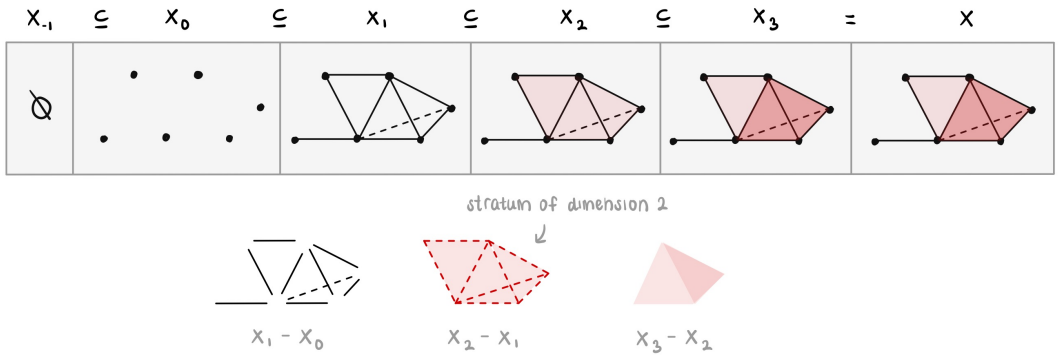


Figure 5: Filtered space

Regular and Singular Strata

In the study of filtered spaces, the strata of the highest possible dimension are particularly important.

Strata [29]

If X is a filtered space of dimension n, the components of $X_n - X_{n-1}$ are called the **regular strata** of X, and all other strata are called **singular strata**. The union of the singular strata, denoted by Σ_X, is known as the **singular locus** of X.

While $\Sigma_X = X_{n-1}$, the notation Σ_X provides a convenient way to refer to the singular locus without explicitly mentioning the formal dimension of X.

Example Let X be an n-dimensional simplicial complex filtered by the i-dimensional skeleton as above. Then, the interiors of the n-simplices are the regular strata, and the interiors of all other faces are the singular strata.

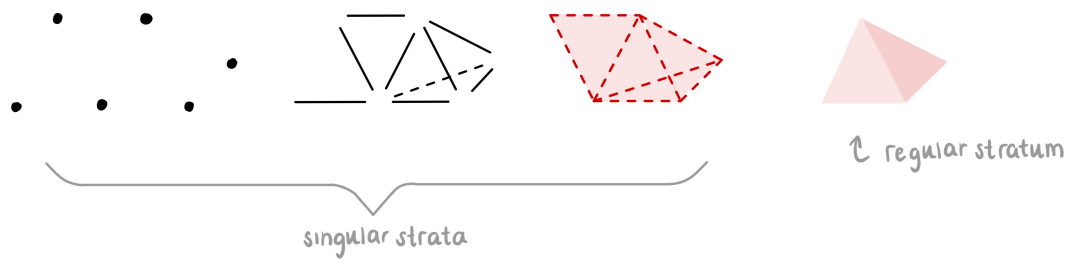


Figure 6: Regular and singular strata

The singular locus is the union of the singular strata $\Sigma_X = \{X_0 - X_{-1}\} \cup \{X_1 - X_0\} \cup \{X_2 - X_1\}$.

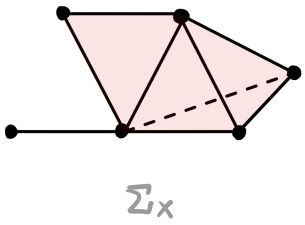


Figure 7: Singular locus

Cones

Cones are a common tool when constructing certain types of filtered spaces, so we introduce them here.

Open Cone [50]

If X is a compact Hausdorff space, then the **open cone** on X is defined as

$$\mathring{c}X = X \times [0,1)/X \times \{0\}$$

The open cone $\mathring{c}X$ on X is the result of identifying the subset $X \times \{0\}$ of $X \times [0,1)$ to a single point (called the vertex of the cone). Intuitively, one can imagine stretching X along the real line and then adding a point at one end to create a cone-like structure. We will let $*$ denote the vertex of the cone and adopt the convention that the cone on the empty set is a point.

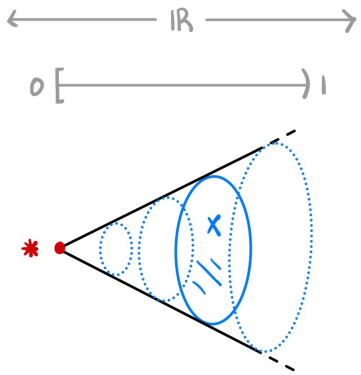


Figure 8: $\mathring{c}X$

One way of creating new filtered spaces from old ones is by taking cones. If X is a compact filtered space of dimension $n-1$, we may construct a filtration of dimension n on the open cone $\mathring{c}X$ as follows. We define the filtration on $\mathring{c}X$ so that $(\mathring{c}X)_i = \mathring{c}\,(X_{i-1})$ for all i with $0 \leq i \leq n$. With this definition, $\mathring{c}X$ has dimension n, and the strata of $\mathring{c}X$ are the cone point $*$ and the products of the strata of X with the open interval. Notice that we always have $(\mathring{c}X)_0 = \{*\}$, with the possibility that $(\mathring{c}X)_i = \emptyset$ for some $i > 0$ if $X_{i-1} = \emptyset$. By definition, the skeleton $(\mathring{c}X)_{-1}$ is empty. The suspension of a compact filtered space can be filtered in a similar manner.

Example Consider the filtration of the 2-simplex $X = \Delta^2$.

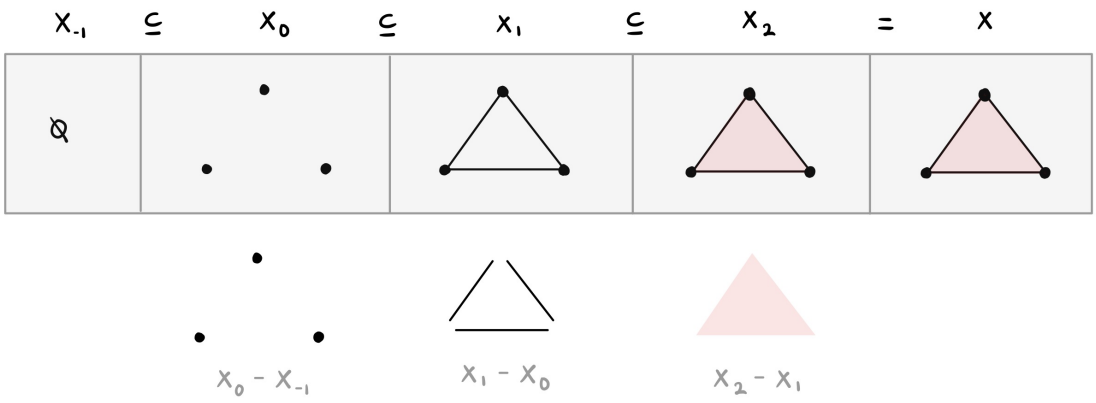


Figure 9: Filtration of 2-simplex

The corresponding filtration on $\mathring{c}(\Delta^2)$ is illustrated below.

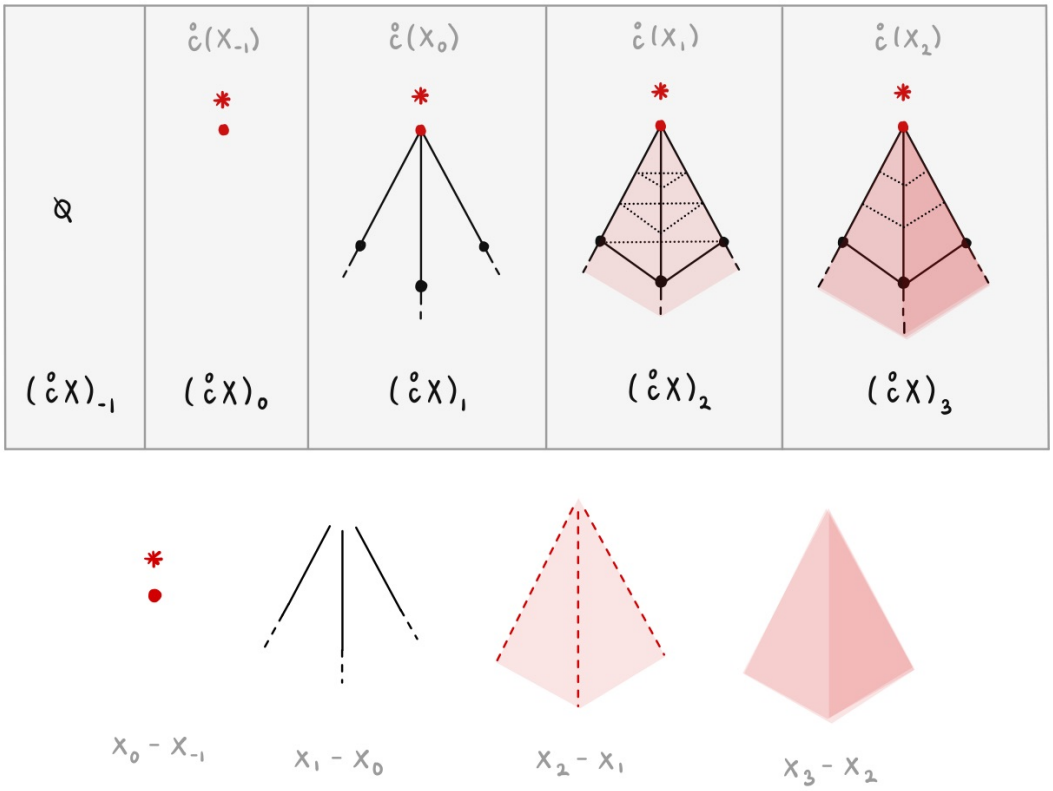


Figure 10: Filtration of $\mathring{c}(\Delta^2)$

Note that the filtration has dimension 3. Furthermore, we have $(\mathring{c}X)_0 = \{*\}$ and $(\mathring{c}X)_{-1}$ is empty.

2.1.3 Stratified Spaces

We can now start defining stratified spaces. Recall that a stratified space is a topological space that can be decomposed into manifold pieces that "fit together nicely." The Frontier Condition formalizes this idea.

The Frontier Condition [29]

A filtered space X satisfies the **Frontier Condition** if for any two strata S, T of X such that

$$S \cap \bar{T} \neq \emptyset$$

then

$$S \subseteq \bar{T}$$

where $\bar{T}$ denotes the closure of T.

The Frontier Condition establishes that the set of strata exhibits a well-defined structure. In particular, we obtain a partial order $<$ defined by $S \prec T$ if $S \subseteq \bar{T}$.

Example Let X be an n-dimensional simplicial complex filtered by the i-dimensional skeleton. Suppose S, T are strata of X. If $S \cap \bar{T} \neq \emptyset$ then S is a face of T. This implies $S \subset \bar{T}$. In other words, each open simplex (i.e., the interior of a simplex) intersects only the simplices of which it is a face (which implies it is contained in the closure of those simplices). This means X satisfies the Frontier Condition. In this context, the partial order $<$ corresponds to the face relations among the simplices.

For example, consider the 2-simplex below, letting S be the 1-stratum and T be the 2-stratum.

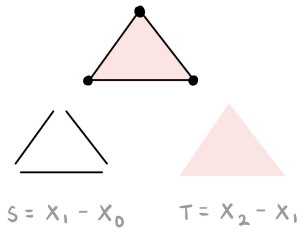


Figure 11: Frontier condition

On the other hand, the following filtered space does not satisfy the Frontier Condition.

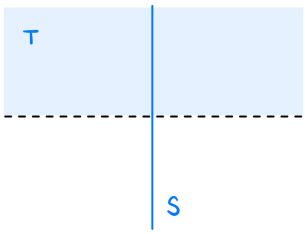


Figure 12: Frontier condition (Non-example)

Consider $S \subset T$ where S is the y-axis in the plane, and T is the union of S with the

open upper half-plane. Here, S intersects the closure of $T - S$ within T, but S is not contained in that closure.

Stratified Space [29]

A filtered space X is a **stratified space** if it satisfies the Frontier Condition.

The partial order induced by the Frontier Condition gives us a structured way of decomposing spaces into strata that fit together in a well-defined manner. More specifically, the closure of any stratum T comprises T and all lower-dimensional strata that intersect T, ensuring that the stratified space is coherently built up from its lower-dimensional pieces.

Proposition [29]

If X is a stratified space, the relation $S \prec T$ defines a partial order. The closure of any stratum T consists of the union of T and all strata of lower dimension that intersect T, given by

$$\bar{T} = \bigcup_{S \prec T} S.$$

Proof. **Reflexivity**: It suffices to show that $S \prec S$ for any stratum S in X. By definition, $S \prec T$ means $S \subseteq \bar{T}$. Since the closure of a stratum S always contains S itself, it follows that $S \subseteq \bar{S}$. Therefore, $S \prec S$.

Transitivity: Assume $S_1 \prec S_2$ and $S_2 \prec S_3$ for strata S_1, S_2, and S_3 in X. This means $S_1 \subseteq \bar{S_2}$ and $S_2 \subseteq \bar{S_3}$. It suffices to show that $S_1 \prec S_3$, i.e., $S_1 \subseteq \bar{S_3}$. Since $S_1 \subseteq \bar{S_2}$ and $S_2 \subseteq \bar{S_3}$, it follows that $\bar{S_2} \subseteq \bar{\bar{S_3}} = \bar{S_3}$ because taking closures is idempotent (i.e., $\bar{\bar{S_3}} = \bar{S_3}$). Therefore, $S_1 \subseteq \bar{S_2} \subseteq \bar{S_3}$, which implies $S_1 \prec S_3$.

Antisymmetry: Suppose $S \prec T$ and $T \prec S$, i.e $S \subseteq \bar{T}$ and $T \subseteq \bar{S}$. It suffices to show $S = T$. Assume $S \subset X_i = X^i - X^{i-1}$ and $T \subset X_j = X^j - X^{j-1}$. Since X^i is closed in X, it follows that $\bar{S} \subseteq X^i$, and similarly $\bar{T} \subseteq X^j$. Given $T \subseteq \bar{S} \subseteq X^i$, we have $j \leq i$. Similarly, since $S \subseteq \bar{T} \subseteq X^j$, we have $i \leq j$. Therefore, $i = j$.

Since $i = j$, both S and T are contained within the same $X_i = X_j$. Moreover, since $S \subseteq \bar{T}$ and $T \subseteq \bar{S}$, and both S and T are connected components of X_i, each must be closed in X_i. This implies there exists a closed set C in X such that $S = X_i \cap C$ and $T = X_i \cap C$, which implies that $S = T$.

Finally, we verify that $\bar{T}$ is precisely the union of T and all strata S such that $S \prec T$. By definition, we have that $\bigcup_{S \prec T} S \subseteq \bar{T}$. Now, suppose $x \in \bar{T}$. Since the strata partition X, x must belong to some stratum S. By the Frontier Condition, if $S \cap \bar{T} \neq \emptyset$, then $S \subseteq \bar{T}$, implying $S \prec T$. Therefore, $x \in S$ for some $S \prec T$, meaning $x \in \bigcup_{S \prec T} S$. This shows $\bar{T} \subseteq \bigcup_{S \prec T} S$.

Thus, we conclude that $\bar{T} = \bigcup_{S \prec T} S$.

□

Example Let X be an n-dimensional simplicial complex filtered by its i-dimensional skeleton. Recall that in this setting, the partial order induced by the filtration corresponds to the face relations among simplices. Specifically, for two simplices S and T in X, we say $S < T$ if and only if S is an open face of the closed simplex $\bar{T}$. Furthermore, the closure of an open simplex T is the disjoint union of T with all of its open faces. This hierarchical structure ensures that each simplex "fits together nicely" with its faces.

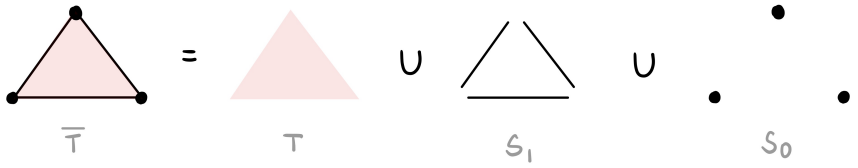


Figure 13: The closure of an open simplex T is the disjoint union of T with all of its open faces.

A particularly nice class of stratified spaces is one in which every stratum is a manifold.

Manifold Stratified Space [29]

A **manifold stratified space** is a stratified space in which every i-dimensional stratum is an i-dimensional manifold.

Example A finite-dimensional simplicial complex is a manifold stratified space. In this setting, the strata are the open faces of its simplices, and each of these faces is homeomorphic to some Euclidean space.

Figure 14: Manifold stratified space

Example Suppose X is a manifold stratified space. Then, the open cone $\mathring{c}X$ is also a manifold stratified space. We define the filtration on $\mathring{c}X$ such that $(\mathring{c}X)_0 = \{*\}$ is the cone point, and for $i > 0$ we have

$$(\mathring{c}X)_i = (0,1) \times X_{i-1}$$

This means the strata of $\mathring{c}X$ are the cone point and the products of the strata of X with the interval $(0,1)$, which are themselves manifolds since X is manifold stratified and we take products. Furthermore, the Frontier Condition holds on $\mathring{c}X$ because it holds for X. Additionally, the cone point is in the closure of every non-empty stratum.

2.1.4 Topologically Stratified Spaces

Manifold stratified spaces can be decomposed into a partially ordered set of strata, where each stratum is a manifold. However, this framework is often still too broad to yield useful results. Manifold theory usually applies conditions like local flatness to prevent problematic cases. Likewise, we can define a local topological structure at each point for stratified spaces to achieve a more manageable and well-behaved space. A tubular neighborhood around a singular stratum is usually too much to expect because singular strata often have irregular structures that prevent them from being approximated by such

neighborhoods. However, by imposing some additional conditions, we can achieve locally uniform "normal behavior" along singular strata.

This need for uniformity along singular strata leads to the local normal triviality condition. This condition states that any point x in a singular stratum $X_j - X_{j-1}$ should have a neighborhood N_x that looks like $N_x \cong \mathbb{R}^j \times \mathring{c}L$, where L is a compact filtered space called the "link." The homeomorphism maps $\mathbb{R}^j \times \{v\}$ (with v being the cone point) to a neighborhood of x in $X_j - X_{j-1}$. This means that for each point x in a singular stratum $X_j - X_{j-1}$, we have a bundle over some neighborhood of x, with the fiber being a cone on a lower-dimensional stratified space.

In fact, the link L can be considered as the intersection of a "normal slice" at x with a sphere. One can show that the homeomorphism type of this normal slice is independent of both the choice of a sufficiently small neighborhood of x and the choice of x within a particular connected component of $X_j - X_{j-1}$ [3]. Consequently, these manifold pieces fit uniformly into the larger space.

Historically, Whitney was the first to highlight that a good stratification should meet specific regularity conditions along strata [74] [75]. These are famously known as "Whitney conditions (a) and (b)." Topologically stratified spaces offer a purely topological framework for studying singularities, similar to Whitney's more differential-geometric theory. These spaces were introduced by Thom, who demonstrated that every Whitney stratified space is also a topologically stratified space with the same strata [68].

Topologically Stratified Space

We define a **topologically stratified space** by induction on dimension as follows:

(i) A **0-dimensional topologically stratified space** is a countable set with the discrete topology.

(ii) For $m > 0$, an m**-dimensional topologically stratified space** is a filtered space X

$$X = X_m \supseteq X_{m-1} \supseteq \cdots \supseteq X_1 \supseteq X_0$$

so that the following **local normal triviality** condition is satisfied:

Local Normal Triviality

If $x \in X_j - X_{j-1}$ there exists

1. a neighbourhood N_x of x in X,
2. a compact $(m-j-1)$-dimensional topologically stratified space L with filtration
$$L = L_{m-j-1} \supseteq \cdots \supseteq L_1 \supseteq L_0$$
and
3. a homeomorphism
$$\phi : N_x \to \mathbb{R}^j \times \mathring{c}L$$
where $\mathring{c}L$ is the open cone on L,

such that ϕ takes $N_x \cap X_{j+i+1}$ homeomorphically onto

$$\mathbb{R}^j \times \mathring{c}\,(L_i)$$

for $m - j - 1 \geq i \geq 0$, and ϕ takes $N_x \cap X_j$ homeomorphically onto

$$\mathbb{R}^j \times \{ \text{ vertex of } \mathring{c}L \}$$

This guarantees that the subset $X_j - X_{j-1}$ is a topological manifold of dimension j, which we denote as S_j and call the j**th stratum** of X. The connected components of the strata are called **pieces**, and X is locally normal trivial along these pieces. The union of lower strata X_{d-1} is Σ_X, the **singular locus** of X.

L is called the **link** of the stratum containing x. Up to stratified diffeomorphism, the link L depends only on the connected component of the stratum $X_j - X_{j-1}$ (the piece) containing x. Hence, the condition on the neighborhoods ensures that each point within a piece has the same local structure.

Example [4] Consider the pinched torus with a disc stretched across the center. We'll call the boundary of this disc C. If we remove C, we obtain a disconnected 2-manifold, with the two pieces forming the stratum $X_2 - X_1$. Note that C itself is a one-manifold.

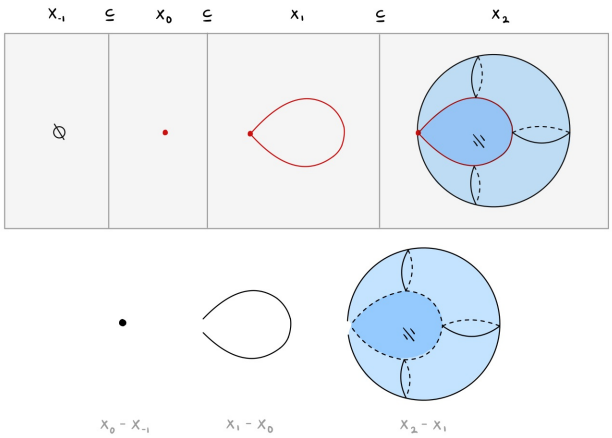


Figure 15: Stratification of pinched torus

However, not all points on C are singularities of the same kind. Let x be the pinch point, and let $y \in C$ where $y \neq x$.

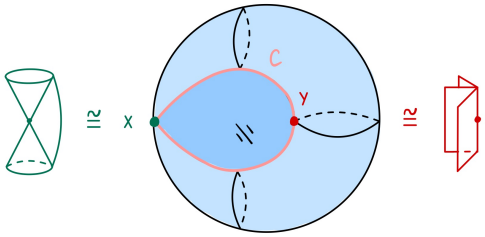


Figure 16: Pinched torus

Then y has a neighborhood that is the product of an interval in C and a cone on three points: one from the disc and two from the torus (this resembles three sheets glued

together along a line). However, x has no such neighborhood. Instead, its neighborhood consists of a cone on "two circles joined by a line".

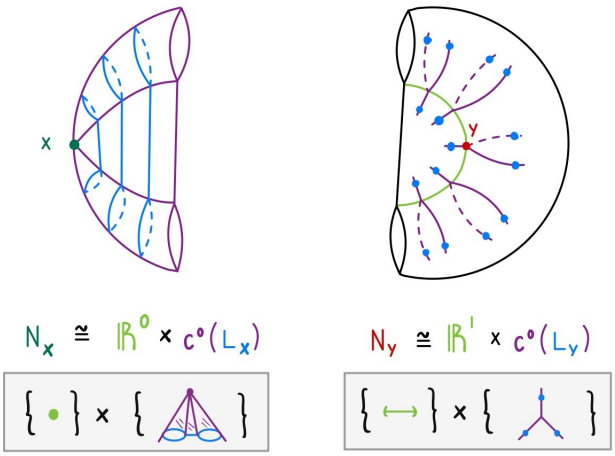


Figure 17: The neighborhood of x is the product of the point $\mathbb{R}^0$ and a cone on "two circles joined by a line." In contrast, the neighborhood of y is the product of the interval $\mathbb{R}^1$ and a cone on three points.

Thus, the "local normal triviality" condition requires we place x in a separate stratum.

Other notions of stratified spaces have evolved with various approaches to intersection homology. The following kinds of spaces were the ones originally considered by Goresky and MacPherson in [31].

Topological Pseudomanifold

A **topological pseudomanifold** of dimension m is a para-compact Hausdorff topological space X which possesses a topological stratification such that

$$X_{m-1} = X_{m-2}$$

and $X - X_{m-2}$ is dense in X

Overall, the definition tells us the following. An m-dimensional topological pseudomanifold X is mostly the m-manifold $X - X_{m-2}$, which is dense in X. The rest of X is made up of manifolds of various dimensions, and these must fit together nicely, in the sense that each point in each stratum should have a neighborhood that is a trivial fiber bundle, whose fibers are cones on lower-dimensional stratified spaces.

Examples of such spaces are copious. Any complex analytic or algebraic variety can be given such a structure, as can certain quotient spaces of manifolds by group actions. Other simple examples arise by taking open cones on manifolds, suspending manifolds (or by repeated suspensions), gluing manifolds and pseudomanifolds together in allowable ways, etc. [28].

2.2 Intersection Homology

Having introduced stratified spaces, we now turn to discuss intersection homology, which is defined by modifying the definition of homology groups to place "allowability" conditions

on the dimensions in which chains can meet singularities. These geometric conditions are controlled by a perversity parameter $\bar{p}$, which assigns an integer to each i-skeleton (or stratum) of the space.

2.2.1 Homology with $\mathbb{Z}/2\mathbb{Z}$ Coefficients

Homology provides an algebraic framework for understanding and quantifying the "holes" in a given shape. We briefly recall some background here following [20] and refer the reader to [35] or [22] for a more extensive treatment of the subject.

Let K be a simplicial complex. An **i-chain** is a subset of i-simplices in K. These i-chains can be added together, where the sum of two i-chains is defined as the symmetric difference of their respective sets. Thus, any i-simplex that is an element in both chains cancels out in modulo-2 arithmetic. The **boundary** of an i-simplex is the set of its $(i-1)$-dimensional faces, forming an $(i-1)$-chain.

The set of all i-chains, together with the sum operation, forms a group denoted by $\mathrm{C}_i(K)$. Furthermore, there exists a boundary operator $\partial_i : \mathrm{C}_i(K) \to \mathrm{C}_{i-1}(K)$ that maps an i-chain to the sum of the boundaries of its simplices.

An **i-cycle** is an i-chain with an empty boundary, and an **i-boundary** is an i-cycle that is itself the boundary of some $(i+1)$-chain. These i-cycles and i-boundaries form their own groups, denoted $Z_i(K)$ and $B_i(K)$, respectively. Since every i-boundary is also an i-cycle, and every i-cycle is an i-chain, these groups are nested: $B_i(K) \subseteq Z_i(K) \subseteq \mathrm{C}_i(K)$.

Now, two i-cycles are **homologous** if their sum is an i-boundary. Equivalently, two i-cycles are homologous if one can be transformed into the other by adding an i-boundary. Homology defines an equivalence relation, and the equivalence classes under this relation form the i-th homology group, denoted $H_i(K) = Z_i(K)/B_i(K)$.

All of these groups—$\mathrm{C}_i(K)$, $Z_i(K)$, $B_i(K)$, and $H_i(K)$—are vector spaces, so their ranks correspond to their dimensions. The rank of the i-th homology group is particularly important and is known as the i-th **Betti number** of K, given by

$$\operatorname{rank} H_i(K) = \operatorname{rank} Z_i(K) - \operatorname{rank} B_i(K).$$

Consider a subcomplex L of K. The concept of **relative homology** describes the connectivity of the pair (K, L), which can be thought of geometrically as treating L as a single point within K. The **relative chain group** $\mathrm{C}_i(K, L)$ is the quotient $\mathrm{C}_i(K)/\mathrm{C}_i(L)$. Relative cycles and boundaries are defined similarly to their absolute counterparts, with the key distinction that a relative cycle's boundary must lie within L. The relative homology groups of the pair (K, L) are denoted by $Z_i(K, L)$. Importantly, for the types of spaces we consider, the relative homology of the pair (K, L) is isomorphic to the homology of the quotient space K/L. It is well known that the homology groups of K, L, and (K, L) are related by a **long exact sequence**:

$$\ldots \to H_i(L) \to H_i(K) \to H_i(K, L) \to H_{p-1}(L) \to \ldots$$

Assuming only finitely many groups have non-zero ranks, a fundamental property of long exact sequences is that the alternating sum of the dimensions of the vector spaces in the sequence vanishes.

Vanishing Alternating Sum [20]

Let $L \subseteq K$ be simplicial complexes. Then

$$\sum_{i \in \mathbb{Z}} (-1)^i \left[\operatorname{rank} H_i(L) - \operatorname{rank} H_i(K) + \operatorname{rank} H_i(K, L)\right] = 0.$$

Proof Sketch. By the definition of exactness, each homology group's rank can be expressed as the sum of two non-negative integers, one shared with the preceding group and the other with the succeeding group in the sequence. Since only finitely many groups have non-zero ranks, the alternating sum of these ranks must vanish.

2.2.2 Motivation

Before discussing intersection homology, we begin with some motivation and general remarks. One of the most significant results in the topology of manifolds is Poincaré Duality [60] [61]. To state the Poincaré duality isomorphism, we follow [50] and recall the cap product on an n-manifold M. The cap product is a map

$$C^i(M) \times C_n(M) \xrightarrow{\frown} C_{n-i}(M),$$

where $C_i(M)$ and $C^i(M)$ are the i-chains and i-cochains on M. For $a \in C^{n-i}(M)$, $b \in C^i(M)$, and $\sigma \in C_n(M)$, the cap product satisfies:

$$a(b \frown \sigma) = (a \smile b)(\sigma).$$

This operation is compatible with boundary maps, allowing it to extend to cohomology and homology:

$$H^i(M; \mathbb{Z}) \times H_n(M; \mathbb{Z}) \xrightarrow{\frown} H_{n-i}(M; \mathbb{Z}).$$

For a closed, oriented, connected topological n-manifold M, it's known that $H_n(M; \mathbb{Z}) = \mathbb{Z}$, where the generator of this group referred to as the fundamental class of M, denoted by $[M]$. This leads us to Poincaré Duality:

Poincaré Duality

Let M be a closed, connected, oriented topological n-manifold with fundamental class $[M]$. Then capping with $[M]$ induces an isomorphism:

$$H^i(M; \mathbb{Z}) \xrightarrow{\cong} H_{n-i}(M; \mathbb{Z}),$$

for all integers i.

For a modern proof, see [35]. A consequence of Poincaré Duality is the existence of a non-degenerate pairing:

$$H_i(M, \mathbb{C}) \otimes H_{n-i}(M; \mathbb{C}) \longrightarrow \mathbb{C}.$$

This pairing implies that the Betti numbers of M in complementary degrees are equal:

$$\dim_{\mathbb{C}} H_i(M; \mathbb{C}) = \dim_{\mathbb{C}} H_{n-i}(M; \mathbb{C}).$$

However, Poincaré duality generally fails for spaces that are not manifolds. Even a single point that is not locally Euclidean can cause this failure. For example, in the one-point union of two n-dimensional spheres $S^n \vee S^n$, where $n > 0$, we have $H_0(S^n \vee S^n) \cong \mathbb{Z}$ but $H^n(S^n \vee S^n) \cong \mathbb{Z} \oplus \mathbb{Z}$.

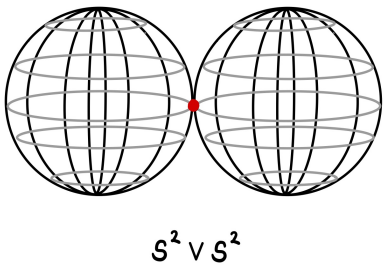


Figure 18: We have $H_0(S^2 \vee S^2) \cong \mathbb{Z}$ but $H^2(S^2 \vee S^2) \cong \mathbb{Z} \oplus \mathbb{Z}$

A more involved example is the suspension of the 2-torus, ST^2, which has two isolated singularities at the cone points a and b.

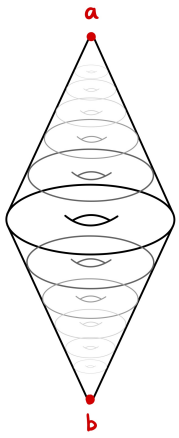


Figure 19: The suspension of the 2-torus cT^2

Each of a and b have a neighborhood homeomorphic to the cone on the torus, cT^2, however the cone point of cT^2 does not have a neighborhood homeomorphic to $\mathbb{R}^3$. This can be demonstrated as follows [29]. Let v be the cone point of cT^2. Since cones are contractible, the long exact sequence of the pair and the homotopy invariance of homology gives us

$$H_2(cT^2, cT^2 - \{v\}) \cong H_1(cT^2 - \{v\}) \cong H_1(T^2) \cong \mathbb{Z} \oplus \mathbb{Z}$$

However, if v had a neighborhood homeomorphic to $\mathbb{R}^3$, then by excision we would have

$$H_2(cT^2, cT^2 - \{v\}) \cong H_2(\mathbb{R}^3, \mathbb{R}^3 - \{v\}) \cong H_1(\mathbb{R}^3 - \{v\}) \cong H_1(S^2) \cong \mathbb{Z}$$

Therefore, cT^2 is not a manifold. Routine computations show that the homology groups are:

$$\begin{aligned} H_3(cT^2) &= \mathbb{Z}, \\ H_2(cT^2) &= \mathbb{Z} \oplus \mathbb{Z}, \\ H_1(cT^2) &= 0, \\ H_0(cT^2) &= \mathbb{Z}, \end{aligned}$$

and the cohomology groups are:

$$H^3(cT^2) = \mathbb{Z},$$
$$H^2(cT^2) = \mathbb{Z} \oplus \mathbb{Z},$$
$$H^1(cT^2) = 0,$$
$$H^0(cT^2) = \mathbb{Z}.$$

Thus, for example, $H_2(cT^2) \ncong H^1(cT^2)$, indicating that Poincaré duality fails.

The failure of Poincaré duality in spaces with singularities arises from issues with transversality. Transversality describes how two objects intersect and can be thought of as the "opposite" of tangency.

Transversality

Two submanifolds M and N of a manifold Y are **transverse** if, at every point $x \in M \cap N$,

$$\text{span}\{T_xM, T_xN\} = T_xY$$

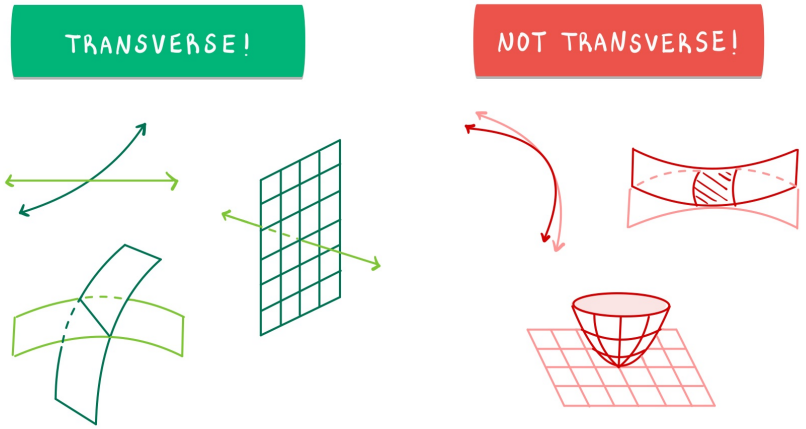


Figure 20: Transversality

In smooth manifold theory, submanifolds can be adjusted via small isotopies to achieve transversality. Similar techniques exist for moving polyhedra in general position, and chains can be pushed into general position with respect to submanifolds or other chains without changing their homology class.

For an n-dimensional PL stratified pseudomanifold X, when computing the homology $H_*(X)$ from the chain complex $C_\bullet(X)$, we typically use arbitrary PL chains. A PL i-chain ξ is transverse to the stratification of X if

$$\dim(|\xi| \cap X_{n-k}) = i + (n-k) - n = i - k,$$

for all $k \geq 2$. According to a theorem by McCrory [51], if we define a chain complex to include only transverse PL chains and compute its i-th homology, we obtain the cohomology $H^{n-i}(X)$ of X. Thus, if every chain could be made transverse to the stratification, Poincaré duality would hold. The failure then of Poincaré duality in stratified spaces is related to the difficulty of making chains transverse to the stratification [73]. A homology theory for such spaces must address this issue.

2.2.3 Historical Remarks

In the late 1970s and early 1980s, Mark Goresky and Robert MacPherson first defined (in [31] for PL pseudomanifolds and in [32] for topological pseudomanifolds) a collection of groups $IH_*^{\bar{p}}(X)$, called intersection homology groups. Their goal was to extend some of the most significant tools of manifold theory, such as Poincaré duality and signatures, to spaces with singularities. Goresky and MacPherson introduced a multi-index "perversity" parameter to specify allowable "deviation" from full transversality and to associate a group to each value of the parameter, resulting in a spectrum of groups that interpolate between cohomology and homology [3].

Remarkably, the groups in complementary dimensions with opposite perversities exhibit a type of Poincaré duality that holds for all pseudomanifolds. Goresky and MacPherson demonstrated that if X is a compact n-dimensional stratified pseudomanifold and $\bar{p}$ and $\bar{q}$ are complementary perversities such that $\bar{p}(k) + \bar{q}(k) = k - 2$, then there exists an intersection pairing $I^{\bar{p}}H_i(X) \otimes I^{\bar{q}}H_{n-i}(X) \to \mathbb{Z}$ which is nondegenerate when all groups are tensored with the rationals $\mathbb{Q}$. Although intersection homology was initially introduced using piecewise-linear (PL) chain complexes, Goresky and MacPherson extended the definition of these groups using sheaf theory [32]. This reformulation proved to be highly successful, particularly in its applications to algebraic geometry and representation theory.

King later introduced a singular version of intersection homology and used it to provide another proof of invariance [40]; this version was eventually shown to be isomorphic to the sheaf formulation. The result of topological invariance was further extended to the broader category of locally conelike topological spaces [34]. For a more detailed exposition on the development of intersection homology, see the survey by Kleiman [42].

2.2.4 Perversity

Intersection homology groups are defined by considering the simplices and chains that we ordinarily use to define homology groups, but placing some limitations on how these chains are allowed to interact with simplices in X. In practice, this is done by restricting the dimensions of the intersections of chains with skeleta or strata [29]. There can be many different ways to impose these constraints: we could completely forbid a chain from intersecting an i-skeleton, impose no limitations at all, or something in between. These choices are encoded in a parameter called the "perversity". Because there are numerous options for these parameters, many different kinds of intersection homology groups exist.

The most general definition of perversity we will use is as follows.

General Perversities [48]

Let X be a filtered space. A **general perversity** on X is a function mapping the skeleta of X to the integers

$$\bar{p} : \{X_n, X_{n-1}, \ldots, X_0\} \to \mathbb{Z}$$

In other words, a perversity is a sequence of integers

$$\bar{p} = (p_0, p_1, \ldots p_n)$$

For defining intersection homology, a perversity function takes the codimension of a skeleton as its input. The idea is that the index k is the codimension of the skeleton X_{n-k} in an n-dimensional filtered space. The output determines how much chains in our homology

computations can interact with these skeleta. If $\bar{p}$ and $\bar{q}$ are perversities on X and $\bar{p}_k \leq \bar{q}_k$ for all k, then we write $\bar{p} \leq \bar{q}$.

When Goresky and MacPherson first introduced intersection homology, they required some specific conditions for the perversity parameters. These rules were necessary for two main reasons: to ensure the intersection homology groups met a generalized version of Poincare duality and to guarantee these groups were topological invariants. In other words, the groups needed to be independent of the choice of stratification of X. The original definition, which we refer to as "GM Perversity," is as follows.

Goresky-MacPherson Perversities

A **GM perversity** $\bar{p}$ is a function

$$\begin{aligned} \bar{p} : \mathbb{N} &\to \mathbb{N} \\ i &\mapsto \bar{p}_i \end{aligned}$$

satisfying the following properties

1. $\bar{p}_0 = \bar{p}_1 = \bar{p}_2 = 0$
2. $\bar{p}_k \leq \bar{p}_{k+1} \leq \bar{p}_k + 1$

These conditions define a perversity as a step function. It starts at 0, and with each increase of one in the input, the output either remains the same or increases by one.

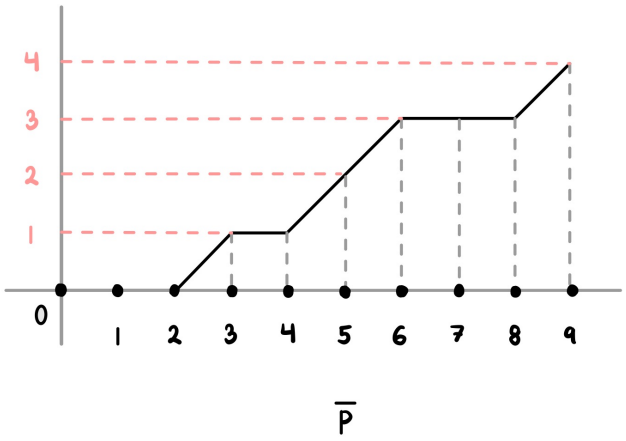


Figure 21: The perversity $\bar{p} = (0, 0, 0, 1, 1, 2, 3, 3, 3, 4)$

Since GM perversities always evaluate to 0 on codimension-zero strata and we typically assume there are no codimension-one strata when using GM perversities, it is standard to represent GM perversities as functions with domain $\{2, 3, 4, \dots\}$. One convenient way to describe GM perversities is to think of them as sequences

$$(\bar{p}_2, \bar{p}_3, \bar{p}_4, \dots)$$

Over time, it has become clear that the conditions imposed by Goresky and McPherson can be relaxed [28]. To quote Friedman, "In relaxing these restrictions, one usually loses the topological invariance of intersection homology. However, this should be seen not as a loss but as an opportunity to study stratification data by using intersection homology theory

to assess these stratifications and perhaps measure the difference between them. In this context, we most certainly do not desire stratification independence" [28]. Nonetheless, duality results usually persist, provided the appropriate generalizations of intersection homology are chosen.

Example Several GM perversities play a significant role in intersection homology, as intersection homology groups dualize not only with respect to dimension but also with respect to perversities.

The minimal **zero perversity** $\bar{0}$ is the function

$$\bar{0}_k = (0, 0, 0, 0, \dots)$$

The maximal **top perversity** $\bar{t}$ is the function $\bar{t}$ is

$$\bar{t}_k = (k-2) = (0, 1, 2, 3, \dots)$$

The **lower middle perversity**

$$\bar{m}_k = \left\lfloor \frac{k-2}{2} \right\rfloor = (0, 0, 1, 1, 2, 2, \dots)$$

and the **upper middle perversity**

$$\bar{n}_k = \left\lfloor \frac{k-1}{2} \right\rfloor = (0, 1, 1, 2, 2, 3, \dots)$$

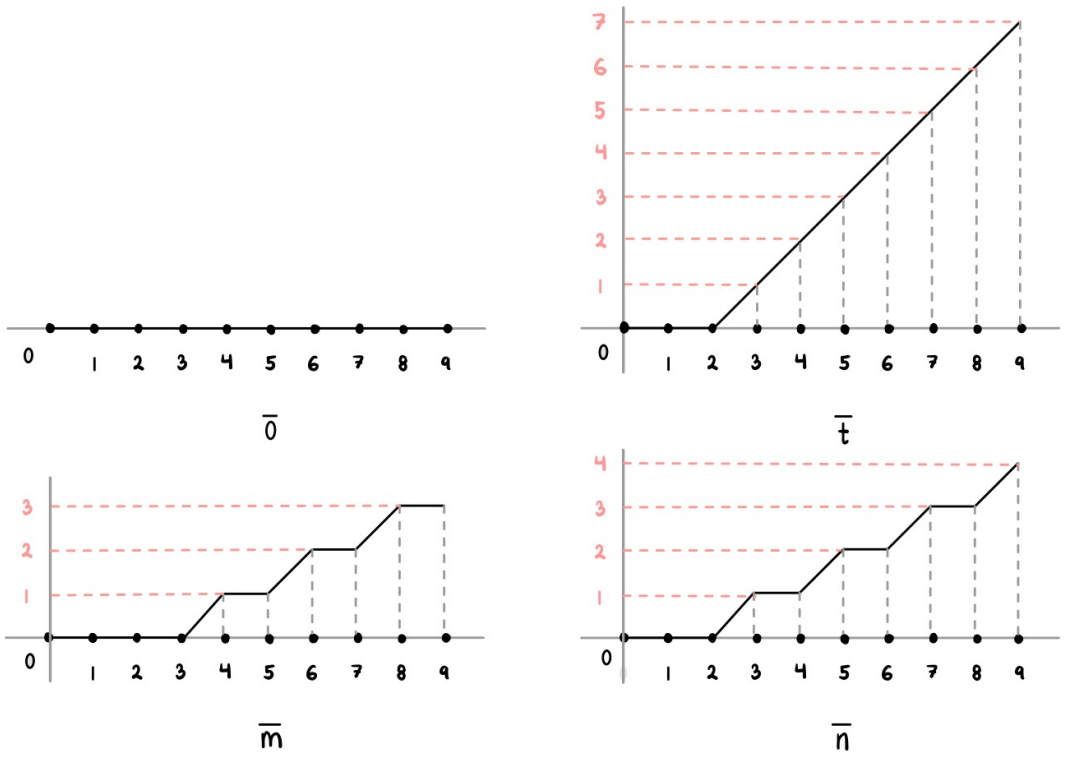


Figure 22: The zero perversity $\bar{0}$, top perversity $\bar{t}$, lower middle perversity $\bar{m}_k$, and upper middle perversity $\bar{n}_k$

We call two perversities $\bar{p}$ and $\bar{q}$ **complementary** or **dual** if $\bar{p} + \bar{q} = \bar{t}$. For example, if $d = 2$, then $(-1, 0)$ and $(0, 0)$ are dual perversities. Observe that the lower middle and upper middle perversities $\bar{m} + \bar{n} = \bar{t}$ are dual.

We will use perversities to determine the permissible level of intersection between simplices and lower-dimensional skeleta.

2.2.5 Allowable Simplices

We begin with a simplicial version of intersection homology. A simplicial filtered space is a filtered space with a fixed triangulation, where each skeleton in the filtration is a simplicial subcomplex.

Allowable Simplices

Let X be a simplicial filtered space with general perversity $\bar{p}$. An i-simplex σ of X is **$\bar{p}$-allowable** (or **$\bar{p}$-proper**) if its closure $\bar{\sigma}$ satisfies

$$\dim(\bar{\sigma} \cap X_{n-k}) \leq i - k + \bar{p}_k$$

for each codimension-k skeleton X_{n-k}

Since we are working in a simplicial setting, this intersection will be a union of faces of σ, and $\dim(\bar{\sigma} \cap X_{n-k})$ will be the highest dimension of such a face. If $\sigma \cap X_{n-k} = \emptyset$, we adopt the convention that $\dim(\emptyset) = -\infty$.

The intuition behind this inequality is as follows. Consider a m-manifold M^m with submanifolds N^n and P^p. We say that N and P are in **general position** if

$$\dim(N \cap P) \leq n + p - m$$

In particular, this condition is satisfied when N and P intersect transversely. Given that the codimension of P in M is $m - p$, we can rewrite this as

$$\dim(N \cap P) \leq n - \operatorname{codim}(P)$$

Applying this to the simplicial setting, if an i-dimensional simplex σ intersects a codimension-k skeleton X_{n-k} transversely, the dimension of the intersection will satisfy

$$\dim(\sigma \cap X_{n-k}) \leq i - \operatorname{codim}(X_{n-k}) = i - k$$

Non-transverse intersections would result in a higher dimension. Thus, when $\bar{p}_k = 0$, we require that σ intersects X_{n-k} transversely for it to be allowable. Allowing $\bar{p}_k > 0$ relaxes this requirement by a degree controlled by $\bar{p}$. On the other hand, setting $\bar{p}_k < 0$, which is less common, strengthens the requirement for general position.

Since $\bar{p}$ is a function defined on the skeleta, the perversity provides a way to control, on a skeleton-by-skeleton basis, how much deviation from general position is acceptable when defining intersection chains. This is the origin of the term "perversity" — in a sense, it is perverse not to always require general position!

Example Consider the filtered simplicial space X and the 1-simplex σ, which is made up of two manifolds joined at a singular point.

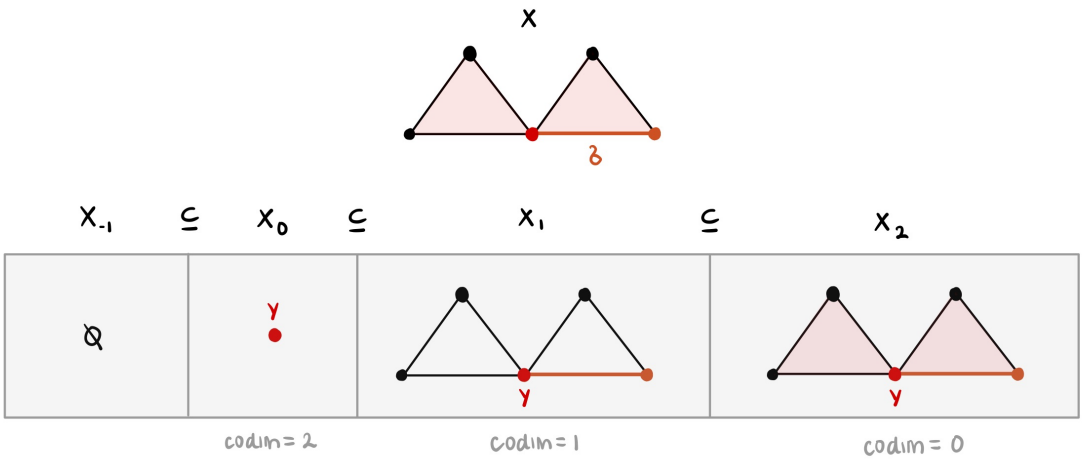


Figure 23: The filtered simplicial space X and the 1-simplex σ

Consider the intersection of σ with the singular point in X_0. We have

$$\dim(\sigma \cap X_0) = 0$$

If the 1-dimensional simplex σ were to intersect the codimension-2 skeleton X_0 transversely, we would have

$$\dim(\sigma \cap X_0) \leq -1$$

This implies that no simplex can intersect the point transversely. Thus, for $p_2 = 0$, we have that no simplex containing the singular point in X_0 is proper. This reflects the fact that X is made up of two pieces. However, we can set $p_2 = 1$ to relax this condition. Then, every simplex that intersects the 0-dimensional point is proper, and the singular point now leads to a proper connected component.

2.2.6 Allowable Chains

Now that we have defined allowable simplices, we can extend this definition to allowable chains. Recall that every element ξ in $\mathbf{C}_i(X)$, known as an i-chain of X, can be uniquely expressed as a linear combination:

$$\xi = \sum_{\sigma} \xi_\sigma \cdot \sigma,$$

where σ ranges over the i-simplices of X, and the coefficients ξ_σ are chosen from our coefficient field $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$.

We refer to σ as a "simplex of ξ" if $\xi_\sigma \neq 0$. Each i-simplex σ in X serves as a basis vector in $\mathbf{C}_i(X)$. This basis vector corresponds to the chain ξ where all coefficients are zero except for ξ_σ, which is equal to the multiplicative identity $1 \in \mathbb{F}$.

Allowable Chains

Let X be a simplicial filtered space with general perversity $\bar{p}$. A chain $\xi \in C_\bullet(X)$ is **$\bar{p}$-allowable** (or **$\bar{p}$-proper**) if all of the simplices of ξ and all of the simplices of $\partial\xi$ (with non-zero coefficient) are $\bar{p}$-allowable.

The $\mathbb{Z}/2\mathbb{Z}$-vector space with basis the allowable i-chains for perversity $\bar{p}$ is denoted

$$IC_i^{\bar{p}}(X)$$

The complex of intersection chains is a subcomplex of the simplicial chains $C_\bullet(X)$. Let us take a moment to discuss the boundary condition in the above definition. To define homology, we need a sequence of groups and boundary maps ∂_i between them that satisfy $\partial_i \circ \partial_{i+1} = 0$. In the simplicial setting, the group $C_i(K)$ is typically defined as a vector space with the i-simplices of K as its basis. The boundary map then maps a sum of i-simplices to a sum of $(i-1)$-simplices, ensuring ∂_i is a well-defined homomorphism $\partial_i : C_i(K) \to C_{i-1}(K)$.

In the context of intersection homology with a given perversity $\bar{p}$, it might seem natural to define a "proper" i-chain as a sum of proper i-simplices. However, this approach doesn't work because the boundary map would not be well-defined: there's no guarantee that the boundary of a proper i-simplex will be a sum of proper $(i-1)$-simplices. To see why, consider the following example [4].

Example Suppose that X_0 is a singular point of codimension two and that $\bar{p} = (0, 0, 0)$. We have

$$\dim(A \cap X_0) = 0 \leq 2 - \operatorname{codim}(X_0) + \bar{p}_0 = 0$$

Thus, the triangle A is itself an allowable 2-simplex with respect to X_0. On the other hand, ∂A is not a sum of allowable 1-simplices: for example, the edge δ is not an allowable 1-simplex since

$$\dim(\delta \cap X_0) = 0 \not\leq 1 - \operatorname{codim}(X_0) + \bar{p}_0 = -1$$

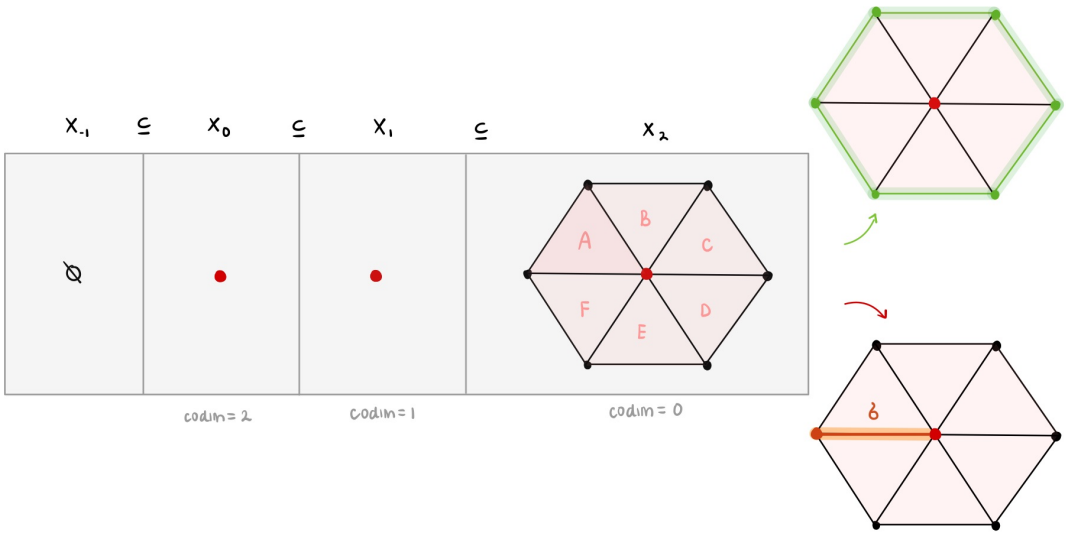


Figure 24: A cannot be part of the 2-dimensional chain group. On the other hand, the sum of all the 2-simplices could be [4]

Thus, to have well-defined boundary maps when defining allowable chains, we require all of the simplices of ξ and all of the simplices of $\partial\xi$ to be $\bar{p}$-allowable.

Now suppose ξ is an allowable i-chain. Note that $\partial(\partial\xi) = 0$. Specifically, $\partial(\partial\xi)$ can be written as a sum of proper $(i-2)$-simplices, meaning that $\partial\xi$ itself is an allowable

$(i-1)$-chain. Consequently, the boundary maps ∂_i provide a sequence of well-defined homomorphism

$$\partial_i : IC_i^{\bar{p}}(X) \to IC_{i-1}^{\bar{p}}(X)$$

with $\partial_i \circ \partial_{i+1} = 0$, forming a chain complex.

2.2.7 Intersection Homology

The intersection homology groups are defined for the intersection chain complex above.

Intersection Homology Groups [3]

The homology groups of the intersection chain complex,

$$IH_i^{\bar{p}}(X) = H_i\left(IC_\bullet^{\bar{p}}(X)\right)$$

are called the **perversity $\bar{p}$ intersection homology groups** of the stratified space X.

So, we define $IH_i^{\bar{p}}(X)$, the i-th intersection homology group with perversity $\bar{p}$ of X, as the kernel of the map $\partial_i : IC_i^{\bar{p}}(X) \to IC_{i-1}^{\bar{p}}(X)$ modulo the image of the map $\partial_{i+1} : IC_{i+1}^{\bar{p}}(X) \to IC_i^{\bar{p}}(X)$.

Thus, the i-th intersection homology group is a $\mathbb{Z}/2\mathbb{Z}$-vector space with a basis consisting of those allowable i-cycles that are not the boundary of an allowable $(i+1)$-chain.

We remark that the definition of intersection homology groups does not use the locally cone-like structure of topologically stratified spaces. In fact, intersection homology groups can be defined for any filtered space. The pseudomanifold structure is typically used to demonstrate that intersection homology groups are topological invariants.

In [31], Goresky and MacPherson impose the condition that their stratified spaces have no strata of codimension 1. This requirement was not problematic since their primary application was to complex algebraic varieties. However, we aim to apply this theory to more general types of stratified spaces and, therefore, do not wish to impose any assumptions on the stratum codimension.

Dropping this assumption has two consequences. Firstly, Poincaré Duality no longer holds, but this can be resolved by using the relative chain groups $C_*(X, \Sigma_X)$, where Σ_X is the singular set of X. As proven in [4], this restores Poincaré Duality. Secondly, intersection homology groups become dependent on the choice of stratification. However, we see this as a useful feature. Persistent intersection homology, being an invariant of a space equipped with both a filtration and a stratification, will change if either is altered. This makes it a valuable tool for identifying where changes in stratification cause significant structural changes in our data.

Example [50] Consider $X = \Sigma(S^1 \sqcup S^1)$, the suspension of a disjoint union of two circles. We denote these circles by A and B, with points $a \in A$ and $b \in B$.

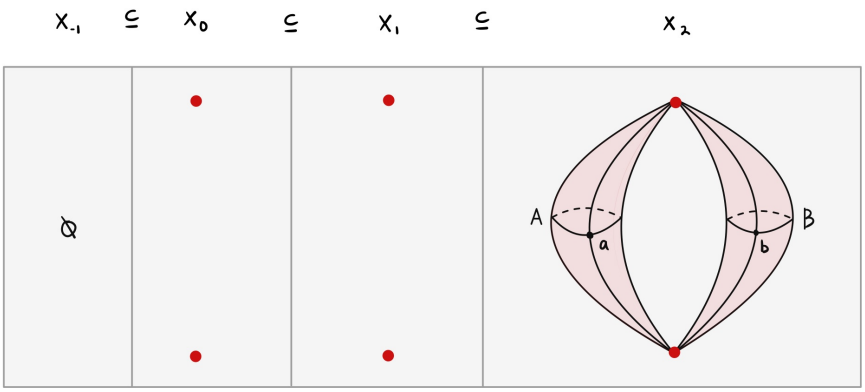


Figure 25: $X = \Sigma(S^1 \sqcup S^1)$

Let $p_a = \text{cone}(a)$ and $p_b = \text{cone}(b)$ be the paths joining a and b to the top suspension point.

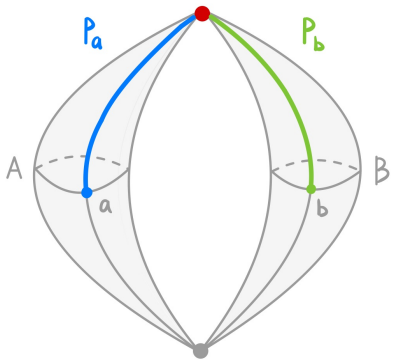


Figure 26: Paths joining a and b to the top suspension point

Let $\Sigma(a)$ and $\Sigma(b)$ be the geodesic paths joining the two suspension points, which pass through a and b.

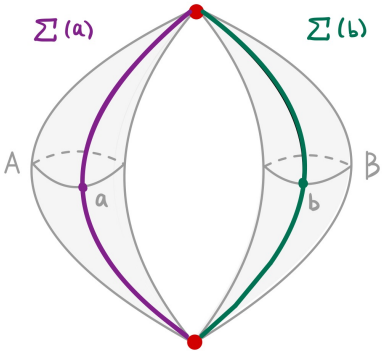


Figure 27: Geodesic paths joining the two suspension points

Finally, let $\Sigma(A)$ and $\Sigma(B)$ be the two 2-spheres obtained by suspending the circles A and B.

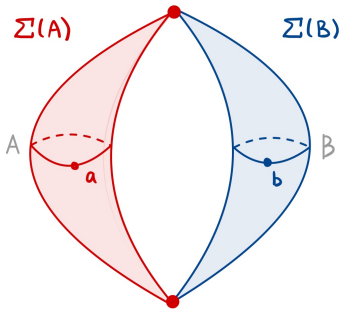


Figure 28: The two 2-spheres obtained by suspending the circles A and B

The homology groups of X are

- $H_0(X) = \mathbb{Z} = \langle [a] \rangle = \langle [b] \rangle$
- $H_1(X) = \mathbb{Z} = \langle [\Sigma(a) - \Sigma(b)] \rangle$
- $H_2(X) = \mathbb{Z} \oplus \mathbb{Z} = \langle [\Sigma(A)], [\Sigma(B)] \rangle$

To calculate intersection homology with perversity $\bar{0} = (0, 0, 0)$, consider the allowable 0-chains with respect to the 0-skeleton:

$$IC_0^{\bar{0}}(X) = \left\{ \xi^0 \in C_0(X) : \dim\left(\xi^0 \cap X_0\right) \leq 0 - 2 + \bar{0}(2) = \boxed{\mathbf{-2}} \right\}$$

Hence, no 0-chains (points) can intersect X_0. Similarly, no 1-chains can intersect X_0:

$$IC_1^{\bar{0}}(X) = \left\{ \xi^1 \in C_1(X) : \dim\left(\xi^1 \cap X_0\right) \leq 1 - 2 + \bar{0}(2) = \boxed{\mathbf{-1}} \right\}$$

Next, consider the allowable 2-chains with respect to the 0-skeleton:

$$IC_2^{\bar{0}}(X) = \left\{ \xi^2 \in C_2(X) : \dim\left(\xi^2 \cap X_0\right) \leq 2 - 2 + \bar{0}(2) = \boxed{\mathbf{0}} \right\}$$

Hence, all 2-chains can intersect X_0, but their boundaries (1-chains) cannot. Therefore, the $\bar{0}$-intersection homology groups of X are as follows.

The 1-chain $p_a - p_b$ (which passes through X_0) is not allowed. Thus, $IH_0^{\bar{0}}(X)$ is generated by points in different components.

- $IH_0^{\bar{0}}(X; \mathbb{Z}) = \mathbb{Z} \oplus \mathbb{Z} = \langle [a], [b] \rangle$

The 1-chain $\Sigma(a) - \Sigma(b)$ is not allowable because it passes through X_0. Furthermore, $\text{cone}(A), \text{cone}(B)$ are allowable 2-chains whose boundaries A and B do not intersect X_0. Thus, A and B are boundaries.

- $IH_1^{\bar{0}}(X; \mathbb{Z}) = 0$

The 2-chains $\Sigma(A)$ and $\Sigma(B)$ are allowable because they intersect X_0 only at points. Thus, these remain generators in intersection homology.

- $IH_2^{\bar{0}}(X; \mathbb{Z}) = \mathbb{Z} \oplus \mathbb{Z} = \langle [\Sigma(A)], [\Sigma(B)] \rangle$

2.2.8 Local Homology

Recall that the **local homology groups** of a space X at a point $x \in X$ are defined as the groups $H_n(X, X - \{x\})$. For any open neighborhood U_x of x, excision gives isomorphisms

$$H_n(X, X - \{x\}) \approx H_n(U_x, U_x - \{x\})$$

assuming points are closed in X. This means the groups $H_n(X, X - \{x\})$ depend only on the local topology of X at x [35]. A homeomorphism $f : X \to Y$ must induce isomorphisms

$$H_n(X, X - \{x\}) \approx H_n(Y, Y - \{f(x)\})$$

for all x and n. Therefore, local homology groups can be used to determine when spaces are not locally homeomorphic at certain points.

If X is a m-manifold, or if x is a point in the top-dimensional stratum of an m-dimensional stratified space, the local homology groups are rank one in dimension m and trivial in all other dimensions. However, the local homology groups for points in lower strata are more enlightening.

Thus, local homology is a valuable tool for studying stratified spaces. If x and y are sufficiently close points in the same stratum, there is a natural isomorphism between their local homology groups:

$$H(X, X - \{x\}) \cong H(X, X - \{y\})$$

This idea has been used to define filtrations on point cloud data [6]. Although the filtration of a stratified space is not unique, there is a natural coarsest filtration defined by considering points x and y equivalent if there exist neighborhoods of x and y with a homeomorphism between them that maps x to y. If two points have such neighborhoods, their local homology groups are the same. The contrapositive of this statement is used to find the best stratification of point cloud data. If two points do not have equivalent local homology groups, then they must belong to different strata.

2.3 Invariance of Intersection Homology

Following [4], we briefly address the independence of intersection homology groups on the choice of stratification and triangulation of a stratified space, as well as the types of homotopy under which intersection homology is invariant. We then proceed to discuss the effects of perversity.

2.3.1 Stratification

A natural question is whether the intersection homology groups of a stratified space X depend on its stratification. In our more general context, the intersection homology groups will indeed rely on the stratification. However, under certain assumptions on the space and perversity, such as having no codimension-1 strata and $p_i \leq p_{i+1} \leq p_i + 1$, independence can be guaranteed [31].

2.3.2 Triangulation

Intersection homology groups are not independent of the triangulation of a space.

Fortunately, this dependence is not very strong. We can achieve independence of triangulation by assuming some minor conditions. Specifically, the intersection homology groups stabilize with repeated barycentric subdivision [49].

2.3.3 Flaglike Triangulations

A triangulation of a space X with stratification $\{X_k\}$ is called **flaglike** if, for every simplex σ and every k, the intersection $\bar{\sigma} \cap X_k$ is a single face of $\bar{\sigma}$. The first barycentric subdivision of any triangulation is always flaglike [49].

Flaglike triangulations are ideal for computing intersection homology. If K is a flaglike triangulation, any further subdivision will yield isomorphic intersection homology groups [41].

2.3.4 Homotopy Type

Recall that a homotopy equivalence between topological spaces induces homology isomorphisms. This is not necessarily true for intersection homology; specifically, a problem arises if an allowable chain is mapped to a non-allowable one.

Intersection homology is, however, preserved by a specific type of homotopy equivalence [4]. A map $f : X \to Y$ between two stratified spaces X and Y is called **stratum-preserving** if the image of each stratum of X under f is contained within the stratum of Y of the same codimension. When stratifying the space $X \times I$ by setting $(X \times I)_k = X_k \times I$, a stratum-preserving map $F : X \times I \to Y$ is referred to as a **stratum-preserving homotopy** from X to Y.

A map $f : X \to Y$ is called a **stratum-preserving homotopy equivalence** if there exists a map $g : Y \to X$ such that $f \circ g$ and $g \circ f$ are both homotopic to the identity via stratum-preserving homotopies. Such a stratum-preserving homotopy equivalence induces intersection homology isomorphisms in all dimensions [27].

2.3.5 Perversity

Recall that given two perversities $\bar{p}$ and $\bar{q}$, we say $\bar{p} \leq \bar{q}$ if $\bar{p}_k \leq \bar{q}_k$ for all k. This gives us a natural inclusion:

$$IC_\bullet^{\bar{p}}(X) \hookrightarrow IC_\bullet^{\bar{q}}(X)$$

The following is a more general partial order on perversities, which also induces an inclusion of intersection chain complexes.

Partial Order on Perversities

For perversities $\bar{p}$ and $\bar{q}$, we say $\bar{p} \preceq \bar{q}$ if one of the following holds for each k:

1. $\bar{p}_k \leq \bar{q}_k$
2. $\bar{q}_k < \bar{p}_k$ and $\bar{p}_k < i - k$
3. $\bar{q}_k < \bar{p}_k$ and for $\gamma \in IC_i^{\bar{p}}(X)$ such that $\dim(\gamma \cap X_{n-k})$ is maximal, we have $\bar{q}_k \geq \dim(\gamma \cap X_{n-k}) - i + k$

for all i and k

Proof. Let $\sigma \in IC_i^{\bar{p}}(X)$ be an arbitrary basis simplex.

Case 1 Suppose $\bar{p}_k \leq \bar{q}_k$. Then

$$\dim(\sigma \cap X_{n-k}) \leq i - k + \bar{p}_k \leq i - k + \bar{q}_k$$

implies σ is $\bar{q}_k$-allowable.

Case 2

Suppose $\bar{q}_k < \bar{p}_k$ and $\bar{p}_k < i - k$. Then no i-simplex σ is $\bar{p}_k$-allowable, since

$$\begin{aligned}\dim(\sigma \cap X_{n-k}) &\leq i - k + \bar{p}_k \\ &< i - k - (i - k) \\ &< 0\end{aligned}$$

Case 3

Suppose $\bar{q}_k > \bar{p}_k$ Let $\gamma \in IC_i^{\bar{p}}(X)$ be a $\bar{p}_k$-allowable simplex such that $\dim(\gamma \cap X_{n-k})$ is maximal. This means

$$\dim(\sigma \cap X_{n-k}) \leq \dim(\gamma \cap X_{n-k})$$

By assumption, we have $\dim(\gamma \cap X_{n-k}) - i + k \leq \bar{q}_k$. Then

$$\dim(\sigma \cap X_{n-k}) \leq \dim(\gamma \cap X_{n-k}) \leq i - k + \bar{q}_k$$

implies that σ is also $\bar{q}_k$-allowable.

Conclusion

Therefore, we have the desired inclusion:

$$IC_\bullet^{\bar{p}}(X) \hookrightarrow IC_\bullet^{\bar{q}}(X)$$

□

Induced Morphism on Intersection Homology

The inclusion of intersection chain complexes induces a canonical morphism on the level of intersection homology:

$$IH_*^{\bar{p}}(X) \longrightarrow IH_*^{\bar{q}}(X)$$

We can use chain maps to construct linear maps of intersection homology groups.

Proof. Let X be a simplicial filtered space, and suppose $\bar{p}$ and $\bar{q}$ are general perversities on X such that $\bar{p} \preceq \bar{q}$. Let the i-simplex σ of X be $\bar{p}$-allowable. This implies

$$\dim(\sigma \cap X_{n-k}) \leq i - k + \bar{p}_k$$

for each codimension-k skeleton X_{n-k}.

Since $\bar{p} \preceq \bar{q}$, we have an inclusion of chain complexes

$$IC_\bullet^{\bar{p}}(X) \hookrightarrow IC_\bullet^{\bar{q}}(X)$$

The inclusion $i : IC_\bullet^{\bar{p}}(X) \hookrightarrow IC_\bullet^{\bar{q}}(X)$ is a chain map, meaning that it commutes with the boundary operators:

$$i(\partial c) = \partial(i(c)) \quad \text{for all } c \in IC_\bullet^{\bar{p}}(X).$$

Since i is a chain map, it induces a map on the homology groups. Specifically, if $[c] \in IH_*^{\bar{p}}(X)$ is a homology class represented by a cycle $c \in IC_\bullet^{\bar{p}}(X)$ (i.e., $\partial c = 0$), then $i(c)$ is a cycle in $IC_\bullet^{\bar{q}}(X)$ and represents a homology class in $IH_*^{\bar{q}}(X)$.

The map on homology induced by the inclusion i is the canonical morphism:

$$\Phi : IH_*^{\bar{p}}(X) \longrightarrow IH_*^{\bar{q}}(X)$$

For a homology class $[c] \in IH_*^{\bar{p}}(X)$, we have:

$$\Phi([c]) = [i(c)]$$

Thus, given $\bar{p} \preceq \bar{q}$, there is a natural inclusion of intersection chain complexes $IC_\bullet^{\bar{p}}(X) \hookrightarrow IC_\bullet^{\bar{q}}(X)$ which induces a canonical morphism on the level of intersection homology:

$$IH_*^{\bar{p}}(X) \longrightarrow IH_*^{\bar{q}}(X).$$

□

2.4 Persistent Intersection Homology

Persistent homology measures the scale or resolution of a topological feature using a geometric function on a topological space and an algebraic process that converts this function into measurements. First, a data set $\mathbb{X}$ is assigned a filtration of topological or combinatorial objects. From this filtered object, a persistence module is computed by calculating homology at each filtration level while keeping track of the functoriality of homology on the inclusions. This method is particularly useful for parameterized families of spaces, such as Vietoris-Rips complexes, that model point-cloud data sets, where features persisting over a larger parameter range are statistically significant [30].

The concept of persistence emerged independently around the turn of the century through the work of Frosini, Ferri, and collaborators in Bologna, Italy; the doctoral research of Robins in Boulder, Colorado; and a project led by Edelsbrunner at Duke University, North Carolina [23].

2.4.1 Persistent Homology

We now give a brief overview of persistent homology. For a more complete treatment of the subject, the reader is referred to [56] or [57]. We begin by recalling the definition of a filtration:

Filtration [56]

Let K be a simplicial complex; a **filtration** of K (of length n) is a nested sequence of subcomplexes of the form

$$F_1K \subset F_2K \subset \cdots \subset F_{n-1}K \subset F_nK = K$$

where i denotes the filtration step.

How we construct such a filtration depends on the underlying space. A commonly used filtration for analyzing such point clouds of data is the following. Consider a metric space M consisting of a finite collection of points in Euclidean space $\mathbb{R}^n$. For such point clouds, we can define a thickening process for any scale $\epsilon > 0$. Specifically, we define $M^{+\epsilon}$ as the union of ϵ-balls centered at each point in M within $\mathbb{R}^n$.

Vietoris-Rips Filtration [56]

Let (M, d) be a finite metric space. The **Vietoris-Rips filtration** of M is an increasing sequence of simplicial complexes $\mathbf{VR}_\epsilon(M)$, indexed by the real numbers $\epsilon \geq 0$. A subset $\{x_0, x_1, \ldots, x_k\} \subset M$ forms a k-dimensional simplex in $\mathbf{VR}_\epsilon(M)$ if and only if the pairwise distances satisfy $d(x_i, x_j) \leq \epsilon$ for all i, j.

In essence, the Vietoris-Rips filtration constructs simplicial complexes at varying scales by connecting points that are within a distance ϵ of each other. For the remainder of this section, however, we drop the F_i to denote a step in the filtration and instead use a filtration based on sublevel sets.

Given a monotonic function $f : K \to \mathbb{R}$ with values $r_1 < r_2 < \ldots < r_n$, we define the i-th sublevel set of f as $K_i = f^{-1}((-\infty, r_i])$. This means that K_i consists of all points in K where the value of f is less than or equal to r_i. Consider a filtration of a simplicial complex K:

$$\emptyset = K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n = K.$$

Rather than the sequence of complexes, we are particularly interested in the topological evolution indicated by the corresponding sequence of homology groups. For each $i \leq j$, there is an inclusion map from the underlying space of K_i to that of K_j, which induces a homomorphism $f_p^{i,j} : H_p(K_i) \to H_p(K_j)$ for each dimension p. This filtration results in a sequence of homology groups connected by these homomorphisms:

$$0 = H_p(K_0) \to H_p(K_1) \to \ldots \to H_p(K_n) = H_p(K),$$

with one such sequence for each dimension p [22]. This sequence is referred to as a persistence module. It can be expressed as a direct sum of indecomposable modules, each of which takes the form $\ldots \to 0 \to k \to \ldots \to k \to 0 \to \ldots$, where $k = \mathbb{Z}/2\mathbb{Z}$. In these indecomposable modules, the maps between the 1-dimensional vector spaces k are identity maps, while all other maps are zero maps. Each indecomposable module corresponds to a specific homological event: the birth and subsequent death of a homology class. As we move from K_{i-1} to K_i, new homology classes may emerge, while others might disappear if they become trivial or merge with existing classes.

Persistent Homology Groups [22]

The **p-th persistent homology groups** are the images of the homomorphisms induced by inclusion, $H_p^{i,j} = \operatorname{im} f_p^{i,j}$, for $0 \leq i \leq j \leq n$.

The persistent homology groups consist of the homology classes of K_i that persist until K_j, formally defined as $H_p^{i,j} = Z_p(K_i)/(B_p(K_j) \cap Z_p(K_i))$. We have such a group for each dimension p and each index pair $i \leq j$.

Consider a class γ in $H_p(K_i)$. We say γ is **born** at K_i if $\gamma \notin H_p^{i-1,i}$. Furthermore, if γ is born at K_i, it dies when entering K_j if it merges with an older class as we move from K_{j-1} to K_j. Specifically, γ **dies** at K_j if $f_p^{i,j-1}(\gamma) \notin H_p^{i-1,j-1}$ but $f_p^{i,j}(\gamma) \in H_p^{i-1,j}$. The difference in function value between birth at K_i and death at K_j is called the **persistence** of γ.

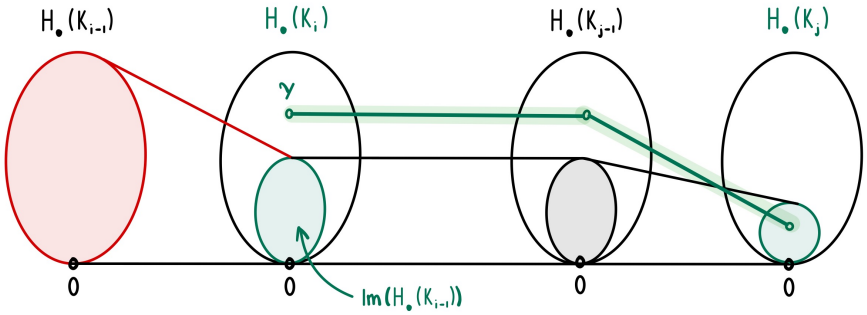


Figure 29: The figure above is due to [22]. The class γ is born at K_i (since it does not lie in the image of $\mathrm{H}_\bullet(K_{i-1})$) and dies entering K_j, where its image merges into the image of $\mathrm{H}_\bullet(K_{i-1})$.

We can conveniently represent the idea that γ is born at a_i and dies as it enters a_j by plotting the point (a_i, a_j) on a two-dimensional plane. When we gather these points for all p-dimensional classes, we create what's called the dimension p persistence diagram, denoted as $\mathrm{Dgm}_p(f)$. Since birth always occurs before death, all the points will lie above the diagonal in this diagram.

However, there's a case where a class γ is born at a_i but doesn't die because it represents a class in $\mathbb{X}_m = \mathbb{X}$. In such cases, we represent γ with the point (a_i, ∞) in the diagram. We get a diagram for each dimension corresponding to the homology groups, and we use $\mathrm{Dgm}(f)$ to refer to the entire infinite series of these diagrams.

An important property of persistence diagrams is their stability. Specifically, the bottleneck distance between the persistence diagrams of two functions $f, g : K \to \mathbb{R}$ is bounded above by the L_∞-distance between the functions [17].

$$W_\infty\left(\mathrm{Dgm}_p(f), \mathrm{Dgm}_p(g)\right) \leq \|f - g\|_\infty.$$

In a persistence diagram, the persistence of a point represents the vertical distance to the diagonal, given by $|f(r_j) - f(r_i)|$. The 1-norm of the diagram, denoted $\|\mathrm{Dgm}(f)\|_1$, is the sum of the persistences of all its points. To manage points at infinity, we introduce a cut-off value C, which effectively replaces ∞ for birth and death values that exceed this threshold.

The concepts of birth and death in persistent homology are not exclusive to homology groups. In fact, persistence can be applied to any sequence of vector spaces connected by homomorphism. Persistence modules were introduced in 2005 by G. Carlsson and A. Zomorodian as an algebraic framework to address persistent homology [76]. We can define persistence modules as follows.

$\mathbb{N}$-indexed Persistence Module [56]

An **$\mathbb{N}$-indexed persistence module over** $\mathbb{F}$ is a sequence $(V_\bullet, a_\bullet)$ of $\mathbb{F}$-vector spaces V_k and linear maps a_k defined for $k \geq 0$ which fit into a diagram

$$V_0 \xrightarrow{a_0} V_1 \xrightarrow{a_1} V_2 \xrightarrow{a_2} \cdots \xrightarrow{a_{k-1}} V_k \xrightarrow{a_k} V_{k+1} \xrightarrow{a_{k+1}} \cdots$$

Persistence modules can also be indexed by real numbers rather than natural numbers.

$\mathbb{R}_+$-indexed persistence module [56]

An **$\mathbb{R}_+$-indexed persistence module over** $\mathbb{F}$ is a pair $(V_\bullet, a_\bullet)$ consisting of an $\mathbb{F}$-vector space V_t for each real number $t \geq 0$ and a linear map $a_{s\leq t} : V_s \to V_t$ for each pair $s \leq t$ of non-negative real numbers satisfying

1. $a_{t\leq t}$ is the identity map on V_t for each $t \geq 0$, and
2. $a_{s\leq t} \circ a_{r\leq s} = a_{r\leq t}$ for every triple $0 \leq r \leq s \leq t$ of real numbers.

One must impose some finiteness constraints to guarantee certain invariants for $\mathbb{R}_+$-indexed persistence modules.

Tameness [56]

A persistence module $(V_\bullet, a_\bullet)$ is called **tame** if

1. the vector spaces V_t are finite-dimensional for all $t \geq 0$, and
2. there are only finitely many $t \geq 0$ for which the map $a_{t-\epsilon\leq t+\epsilon} : V_{t-\epsilon} \to V_{t+\epsilon}$ fails to be an isomorphism for arbitrarily small $\epsilon > 0$.

In particular, persistent homology provides invariants known as barcodes, which are collections of intervals within a fixed totally ordered set, such as $\mathbb{R}$ or $\mathbb{N}$. These intervals correspond to topological features in the data, where each interval's length corresponds to the feature's persistence (such as a connected component or a loop).

Let M be a persistence module indexed by a totally ordered set T. Then there exists a unique multiset $\mathcal{B}(M)$ of intervals in T such that

$$M \cong \bigoplus_{I\in\mathcal{B}(M)} k_I.$$

The multiset $\mathcal{B}(M)$ is called the barcode of M [9].

2.4.2 Persistent Intersection Homology

Given a stratified space equipped with a filtration, we define intersection homology persistence in direct analogy to the ordinary homology case. It has been shown that the persistent homology diagrams $\mathrm{Dgm}_r(f)$ and $\mathrm{Dgm}_r(g)$ for two similar functions are also similar [17], with the bottleneck distance between the diagrams bounded by the L_∞ distance between the functions. The proof can be adapted to yield the same result for persistent intersection homology diagrams [5].

IH Diagram Stability

Let f and g be two tame, real-valued functions on a stratified space X. Then, for each dimension r and each perversity $\bar{p}$:

$$d_B\left(I^{\bar{p}}\,\mathrm{Dgm}_r(f), I^{\bar{p}}\,\mathrm{Dgm}_r(g)\right) \leq \|f - g\|_\infty.$$

2.5 Kernel, Image, and Cokernel Persistence

In recent years, the theory of persistence modules has grown significantly within pure mathematics, leading to productive connections with function theory and symplectic ge-

ometry [62]. In this section, we extend persistence modules to study stratified spaces. In particular, we employ kernel, image, and cokernel persistence to investigate how the singularities of a space affect its homology. From another perspective, this can be interpreted as studying chromatic complexes, where singularities are assigned a different color than the rest of the space [21].

2.5.1 Analyzing Singularities

Suppose

$$\mathbb{X} = \mathbb{X}_n \supseteq \mathbb{X}_{n-1} \supseteq \cdots \supseteq \mathbb{X}_0$$

is a stratified space embedded in $\mathbb{R}^n$ and let $\Sigma_{\mathbb{X}} = \mathbb{X}_{n-1}$ denote the singular locus of $\mathbb{X}$ (that is, the set of all singular points). Suppose $P \subseteq \mathbb{X}$ is a finite point cloud sampled from $\mathbb{X}$, and let $\Sigma_{\mathbb{X}}^{+r}$ be an r-neighbourhood of $\Sigma_{\mathbb{X}}$. Set

$$P_{\mathcal{M}} = P \cap (\mathbb{X} - \Sigma_{\mathbb{X}}^{+r})$$

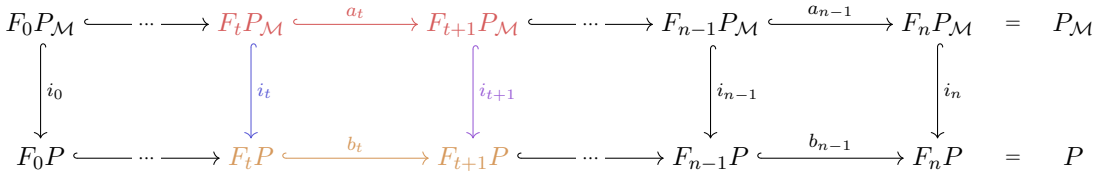
Suppose $\boldsymbol{F}_\bullet P$ is a filtration of P

$$\boldsymbol{F}_0 P \subset \boldsymbol{F}_1 P \subset \cdots \subset \boldsymbol{F}_{n-1} P \subset \boldsymbol{F}_n P = P$$

and denote the inclusion simplicial maps by $a_t : \boldsymbol{F}_t P \hookrightarrow \boldsymbol{F}_{t+1} P$. Let $\boldsymbol{F}_\bullet P_{\mathcal{M}}$ be the filtration restricted to the simplices in $P_{\mathcal{M}}$, so $\boldsymbol{F}_t P_{\mathcal{M}} = \boldsymbol{F}_t P - \Sigma_{\mathbb{X}}^{+r}$, and denote the inclusion simplicial maps by b_i. Then we have an inclusion of filtrations

$$i : \boldsymbol{F}_\bullet P_{\mathcal{M}} \hookrightarrow \boldsymbol{F}_\bullet P$$

Proof. It suffices to show the following diagram commutes.



Well-definedness

By definition, $\boldsymbol{F}_t P_{\mathcal{M}} = \boldsymbol{F}_t P - \Sigma_{\mathbb{X}}^{+r}$, where $\Sigma_{\mathbb{X}}^{+r}$ represents the set of simplices in $\boldsymbol{F}_t P$ that are not in $P_{\mathcal{M}}$. Since $\boldsymbol{F}_t P_{\mathcal{M}}$ is constructed by removing certain simplices from $\boldsymbol{F}_t P$, it is clear that $\boldsymbol{F}_t P_{\mathcal{M}} \subseteq \boldsymbol{F}_t P$ for each t, making the inclusion map i_t well-defined.

Preservation of Inclusion Property

It suffices to show

$$b_t \circ i_t = i_{t+1} \circ a_t$$

for all $t \geq 0$. We have

$$\begin{aligned} b_t \circ i_t(F_t P_{\mathcal{M}}) &= b_t \circ i_t(F_t P - \Sigma_{\mathbb{X}}^{+r}) \\ &= b_t \left(i_t(F_t P) - i_t(\Sigma_{\mathbb{X}}^{+r}) \right) \\ &= b_t(F_t P) \\ &= F_{t+1} P \end{aligned}$$

Similarly,

$$\begin{aligned} i_{t+1} \circ a_t(F_t P_{\mathcal{M}}) &= i_{t+1}(F_{t+1} P_{\mathcal{M}}) \\ &= i_{t+1}(F_{t+1} P - \Sigma_{\mathbb{X}}^{+r}) \\ &= i_{t+1}(F_{t+1} P) \\ &= F_{t+1} P \end{aligned}$$

Hence $i : \boldsymbol{F}_\bullet P_{\mathcal{M}} \hookrightarrow \boldsymbol{F}_\bullet P$ is indeed an inclusion of filtrations, as it is well-defined and respects the inclusion property of the filtrations at each step t.

□

Recall the definition of a simplicial map.

Simplicial Map [56]

Let K and L be simplicial complexes. A **simplicial map** $f : K \to L$ is an assignment $K_0 \to L_0$ of vertices to vertices, sending simplices to simplices. So for each simplex $\sigma = \{v_0, \ldots, v_k\}$ of K, the image $f(\sigma) = \{f(v_0), \ldots, f(v_k)\}$ must be a simplex of L.

Since $P_{\mathcal{M}} \subseteq P$ is a subcomplex, the inclusion map i sends each simplex of $P_{\mathcal{M}}$ to the same simplex in P. Recall further the definition of a chain map.

Chain Map [56]

A **chain map** $\phi_\bullet$ from chain complexes $(C_\bullet, d_\bullet)$ to $(C'_\bullet, d'_\bullet)$ is defined to be a sequence of $\mathbb{F}$-linear maps $\{\phi_k : C_k \to C'_k \mid k \geq 0\}$ which satisfy

$$d'_k \circ \phi_k = \phi_{k-1} \circ d_k$$

for each $k \geq 0$

We can use simplicial maps to produce chain maps:

Simplicial Maps induce Chain Maps [56]

For each dimension $k \geq 0$, and k-simplex σ in K, we have an equality $\partial_k^L \circ \mathbf{C}_k f(\sigma) = \mathbf{C}_{k-1} f \circ \partial_k^K(\sigma)$. Thus simplicial maps $f : K \to L$ induce chain maps

$$\mathbf{C}_\bullet f : \left(\mathbf{C}_\bullet(K), \partial_\bullet^K\right) \to \left(\mathbf{C}_\bullet(L), \partial_\bullet^L\right)$$

The inclusion of filtrations i is a well-defined simplicial map and thus induces a chain map

$$C_\bullet i : C_\bullet(\mathbf{F}_t P_{\mathcal{M}}, \partial_\bullet^{P_{\mathcal{M}}}) \hookrightarrow C_\bullet(\mathbf{F}_t P, \partial_\bullet^P)$$

for each $t \geq 0$, which implies the following diagram commutes for each k

$$\begin{array}{ccccccccccccc}
\cdots & \xrightarrow{\partial_{k+1}^{P_{\mathcal{M}}}} & C_k(\mathbf{F}_t P_{\mathcal{M}}) & \xrightarrow{\partial_k^{P_{\mathcal{M}}}} & C_{k-1}(\mathbf{F}_t P_{\mathcal{M}}) & \xrightarrow{\partial_{k-1}^{P_{\mathcal{M}}}} & \cdots & \xrightarrow{\partial_1^{P_{\mathcal{M}}}} & C_0(\mathbf{F}_t P_{\mathcal{M}}) & \xrightarrow{0} & 0 \\
 & & \downarrow{\scriptstyle C_k i} & & \downarrow{\scriptstyle C_{k-1} i} & & & & \downarrow{\scriptstyle C_0 i} & & \| \\
\cdots & \xrightarrow{\partial_{k+1}^{P}} & C_k(\mathbf{F}_t P) & \xrightarrow{\partial_k^{P}} & C_{k-1}(\mathbf{F}_t P) & \xrightarrow{\partial_{k-1}^{P}} & \cdots & \xrightarrow{\partial_1^{P}} & C_0(\mathbf{F}_t P) & \xrightarrow{0} & 0
\end{array}$$

Recall that chain maps induced well-defined linear maps on homology.

Chain Maps induce maps on Homology [56]

Let $\phi_\bullet : (C_\bullet, d_\bullet) \to (C'_\bullet, d'_\bullet)$ be a chain map. For each dimension $k \geq 0$, there is a well-defined $\mathbb{F}$-linear map

$$\mathbf{H}_k\phi := \mathbf{H}_k\left(C_\bullet, d_\bullet\right) \to \mathbf{H}_k\left(C'_\bullet, d'_\bullet\right)$$

induced by $\phi_\bullet$.

Thus, the chain map $\mathbf{C}_\bullet i$ induces a morphism of homology groups

$$\iota_t := \mathbf{H}_k C_\bullet i := \mathbf{H}_k\left(C_\bullet(\mathbf{F}_t P_{\mathcal{M}}), \partial_k^{P_{\mathcal{M}}}\right) \to \mathbf{H}_k\left(C_\bullet(\mathbf{F}_t P), \partial_k^P\right)$$

for each $k \geq 0$. Thus we have a map of quotient vector spaces $\ker \partial_k^{P_{\mathcal{M}}} / \operatorname{img} \partial_{k+1}^{P_{\mathcal{M}}} \to \ker \partial_k^P / \operatorname{img} \partial_{k+1}^P$, such that ι_t maps $\ker \partial_k^{P_{\mathcal{M}}}$ to $\ker \partial_k^P$ and $\operatorname{img} \partial_{k+1}^{P_{\mathcal{M}}}$ to $\operatorname{img} \partial_{k+1}^P$.

To simplify notation, we henceforth assume $\mathbb{Z}/2\mathbb{Z}$ coefficients. Recall there are induced linear maps on homology

$$\alpha_t := \mathbf{H}_k C_\bullet a_i : \mathbf{H}_k\left(\boldsymbol{F}_i P_{\mathcal{M}}\right) \to \mathbf{H}_k\left(\boldsymbol{F}_{i+1} P_{\mathcal{M}}\right)$$

and

$$\beta_t := \mathbf{H}_k C_\bullet b_i : \mathbf{H}_k\left(\boldsymbol{F}_i P\right) \to \mathbf{H}_k\left(\boldsymbol{F}_{i+1} P\right)$$

in every dimension $k \geq 0$ [56]). For a fixed k, these linear maps fit together into a sequence of vector spaces:

$$\mathbf{H}_k\left(\boldsymbol{F}_0 P_{\mathcal{M}}\right) \xrightarrow{\alpha_0} \mathbf{H}_k\left(\boldsymbol{F}_1 P_{\mathcal{M}}\right) \xrightarrow{\alpha_1} \cdots \xrightarrow{\alpha_{n-2}} \mathbf{H}_k\left(\boldsymbol{F}_{n-1} P_{\mathcal{M}}\right) \xrightarrow{\alpha_{n-1}} \mathbf{H}_k\left(\boldsymbol{F}_n P_{\mathcal{M}}\right)$$

and

$$\mathbf{H}_k\left(\boldsymbol{F}_0 P\right) \xrightarrow{\beta_0} \mathbf{H}_k\left(\boldsymbol{F}_1 P\right) \xrightarrow{\beta_1} \cdots \xrightarrow{\beta_{n-2}} \mathbf{H}_k\left(\boldsymbol{F}_{n-1} P\right) \xrightarrow{\beta_{n-1}} \mathbf{H}_k\left(\boldsymbol{F}_n P\right)$$

Define the associated persistence modules:

$$S_\bullet := t \mapsto H_*(\boldsymbol{F}_t P_{\mathcal{M}})$$
$$A_\bullet := t \mapsto H_*(\boldsymbol{F}_t P)$$

The morphisms on the level of homology

$$\iota_t : \mathbf{H}_k\left(\boldsymbol{F}_t P_{\mathcal{M}}\right) \to \mathbf{H}_k\left(\boldsymbol{F}_t P\right)$$

induce a morphism on the level of persistent modules

$$\iota_\bullet : S_\bullet \to A_\bullet$$

Proof. First, recall the definition of a morphism between persistence modules.

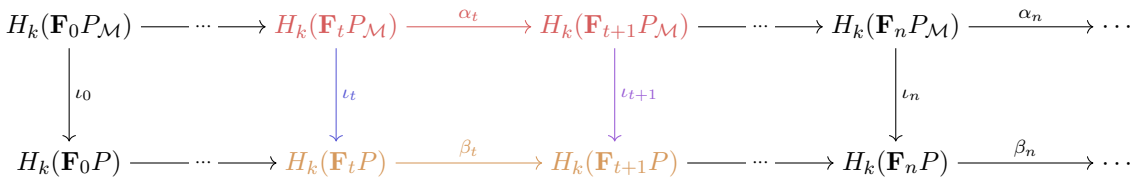
Morphism between Persistence Modules [56]

A **morphism** between persistence modules $(V_\bullet, a_\bullet)$ and $(W_\bullet, b_\bullet)$ is a family of linear maps $\phi_k : V_k \to W_k$ which satisfy

$$b_i \circ \phi_i = \phi_{i+1} \circ a_i$$

for every $i \geq 0$

This definition amounts to requiring the commutativity of all squares in the following diagram of vector spaces for each $k \geq 0$:



It suffices to show

$$\beta_t \circ \iota_t = \iota_{t+1} \circ \alpha_t$$

for each $t \geq 0$ and $k \geq 0$. We have:

- ι_t is the homomorphism induced by the inclusion $i_t : \boldsymbol{F}_t P_{\mathcal{M}} \hookrightarrow \boldsymbol{F}_t P$.
- α_t, the homomorphism induced by the inclusion $a_t : \boldsymbol{F}_t P_{\mathcal{M}} \hookrightarrow \boldsymbol{F}_{t+1} P_{\mathcal{M}}$.
- β_t, the homomorphism induced by the inclusion $b : \boldsymbol{F}_t P \hookrightarrow \boldsymbol{F}_{t+1} P$.

Since we have a well-defined inclusion of filtrations, we have that the following simplicial maps commute:

$$b_t \circ i_t = i_{t+1} \circ a_t$$

Furthermore, these simplicial maps induce well-defined chain maps. Since homology is functorial

Composition of Chain Maps [56]

Given chain maps $\phi_\bullet : (C_\bullet, d_\bullet) \to (C'_\bullet, d'_\bullet)$ and $\psi : (C'_\bullet, d'_\bullet) \to (C''_\bullet, d''_\bullet)$, we have

$$\mathbf{H}_k(\psi \circ \phi) = \mathbf{H}_k \psi \circ \mathbf{H}_k \phi$$

for each dimension $k \geq 0$

We have

$$\begin{aligned} H_k(C_\bullet b_t \circ C_\bullet i_t) &= H_k(C_\bullet i_{t+1} \circ C_\bullet a_t) \\ \implies H_k(C_\bullet b_t) \circ H_k(C_\bullet i_t) &= H_k(C_\bullet i_{t+1}) \circ H_k(C_\bullet a_t) \\ \implies \beta_t \circ \iota_t &= \iota_{t+1} \circ \alpha_t \end{aligned}$$

□

2.5.2 Kernel and Cokernel Persistence

We call the morphism of persistence modules $\iota_\bullet$ an isomorphism if every ι_t is an invertible linear map of vector spaces. We can measure the failure of $\iota_\bullet$ to be an isomorphism by its kernel and cokernel. Specifically, we're interested in the kernels, images, and cokernels of $\iota_\bullet$.

Kernel

The kernel of ι_t, denoted as $\ker \iota_t$, consists of all elements γ in the vector space $H_\bullet(\boldsymbol{F}_t P\mathcal{M})$ that are mapped to the zero element in the vector space $H_\bullet(\boldsymbol{F}_t P)$.

$$\ker \iota_t = \{\gamma \in H_\bullet(\boldsymbol{F}_t P_{\mathcal{M}}) \mid \iota_t(\gamma) = 0 \in H_\bullet(\boldsymbol{F}_t P)\}$$

In simpler terms, the kernel captures all the elements of homology that are "lost" when mapped through the inclusion ι_t because they end up being zero in the target space.

Image

The image of ι_t, denoted as $\operatorname{im} \iota_t$, is the set of all elements in the vector space $H_\bullet(\boldsymbol{F}_t P)$ that are mapped from some element in $H_\bullet(\boldsymbol{F}_t P\mathcal{M})$ through ι_t.

$$\operatorname{im} \iota_t = \{\iota_t(\gamma) \in H_\bullet(\boldsymbol{F}_t P) \mid \gamma \in H_\bullet(\boldsymbol{F}_t P_{\mathcal{M}})\}$$

The image represents the "output" of ι_t, showing us which elements of homology in the target space are actually hit by the inclusion.

Cokernel

The cokernel of ι_t, denoted as $\operatorname{cok} \iota_t$, is a measure of how much of the vector space $H_\bullet(\boldsymbol{F}_t P)$ is not covered by the image of ι_t. It is formally defined as the quotient space:

$$\operatorname{cok} \iota_t = H_\bullet(\boldsymbol{F}_t P) \,/\, \operatorname{im} \iota_t$$

In other words, the cokernel captures the "leftover" part of the target space that is not hit by any element of the homology of the domain via the inclusion ι_t. The following diagram illustrates the construction we proved above:

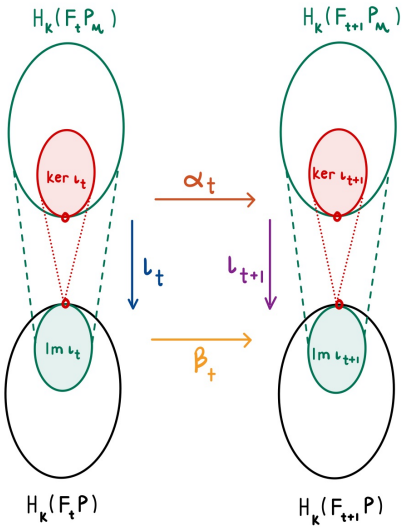


Figure 30: The square above commutes because all four maps are induced by inclusions [19]

From this, we see that the inclusion $\boldsymbol{F}_t P_{\mathcal{M}} \subseteq \boldsymbol{F}_{t+1} P$ induces a homomorphism from $\ker \iota_t$ to $\ker \iota_{t+1}$. Similarly, the inclusion $\boldsymbol{F}_t P \subseteq \boldsymbol{F}_{t+1} P$ induces a homomorphism from

$\operatorname{im} \iota_t$ to $\operatorname{im} \iota_{t+1}$, as well as another homomorphism from $\operatorname{cok} \iota_t$ to $\operatorname{cok} \iota_{t+1}$. This leads to sequences of kernels, images, and cokernels:

$$\begin{aligned}
\operatorname{Ker}(\alpha \to \beta): &\quad \ker \iota_0 \to \ker \iota_1 \to \ldots \to \ker \iota_n; \\
\operatorname{Im}(\alpha \to \beta): &\quad \operatorname{im} \iota_0 \to \operatorname{im} \iota_1 \to \ldots \to \operatorname{im} \iota_n; \\
\operatorname{Cok}(\alpha \to \beta): &\quad \operatorname{cok} \iota_0 \to \operatorname{cok} \iota_1 \to \ldots \to \operatorname{cok} \iota_n,
\end{aligned}$$

with each sequence connected from left to right by homomorphisms. Homology classes are born and die within these sequences just as they do in sequences of homology groups. Therefore, we can define persistent kernels, persistent images, and persistent cokernels, and construct corresponding persistence diagrams, denoted as $\operatorname{Dgm}(\ker \iota_\bullet), \operatorname{Dgm}(\operatorname{im} \iota_\bullet)$, and $\operatorname{Dgm}(\operatorname{cok} \iota_\bullet)$.

2.5.3 Stability

An important property of persistence diagrams in ordinary persistent homology is their stability.

Stability Theorem [17]

Let K be a simplicial complex. The bottleneck distance between the diagrams of $f, g : K \to \mathbb{R}$ is bounded from above by the L_∞-distance between the two maps:

$$W_\infty\left(\operatorname{Dgm}_k(f), \operatorname{Dgm}_k(g)\right) \leq \|f - g\|_\infty;$$

We discussed earlier that the stability of ordinary persistent homology can be extended to persistent intersection homology. This implies that small changes in the input functions f and g result in small changes in their corresponding persistence diagrams, as measured by the bottleneck distance. The stability proof from [17] can be further adapted to the context of kernel, image, and cokernel persistence by considering the maps

$$\iota_t : H_\bullet\left(\boldsymbol{F}_t P_{\mathcal{M}}\right) \to H_\bullet\left(\boldsymbol{F}_t P\right)$$

and

$$\iota''_{t+\varepsilon} : H_\bullet\left(\boldsymbol{F}_{t+\varepsilon} P_{\mathcal{M}''}\right) \to H_\bullet\left(\boldsymbol{F}_{t+\varepsilon} P''\right),$$

where ε is the maximum of the differences $\|f - f''\|_\infty$ and $\|g - g''\|_\infty$. To adapt the stability proof, it's necessary to show that the inclusions $\boldsymbol{F}_t P_{\mathcal{M}} \subseteq \boldsymbol{F}_{t+\varepsilon} P_{\mathcal{M}''}$ and $\boldsymbol{F}_t P \subseteq \boldsymbol{F}_{t+\varepsilon} P''$ map the kernel of ι_t into the kernel of $\iota''_{t+\varepsilon}$. This follows from the commutativity of the homology diagram:

$$\begin{array}{rcl}
H_\bullet\left(\boldsymbol{F}_t P\right) & \to & H_\bullet\left(\boldsymbol{F}_{t+\varepsilon} P''\right) \\
\uparrow \iota_t & & \uparrow \iota''_{t+\varepsilon} \\
H_\bullet\left(\boldsymbol{F}_t P_{\mathcal{M}}\right) & \to & H_\bullet\left(\boldsymbol{F}_{t+\varepsilon} P_{\mathcal{M}''}\right)
\end{array}$$

A similar argument shows that the inclusions $\boldsymbol{F}_t P_{\mathcal{M}''} \subseteq \boldsymbol{F}_{t+\varepsilon} P_{\mathcal{M}}$ and $\boldsymbol{F}_t P'' \subseteq \boldsymbol{F}_{t+\varepsilon} P$ map $\ker \iota''_t$ into $\ker \iota_{t+\varepsilon}$, ensuring the stability result holds. This reasoning also applies to the images and cokernels, allowing the original stability proof to extend to this setting:

Stability Theorem for Kernel, Image, and Cokernel Persistence [19]

Let $f, f'' : P \to \mathbb{R}$ and $g, g'' : P_{\mathcal{M}} \to \mathbb{R}$ be continuous tame functions, with the conditions that $f(x) \leq g(x)$ and $f''(x) \leq g''(x)$ for every $x \in P_{\mathcal{M}} \subseteq P$. Define ε as

the maximum difference between the functions:

$$\varepsilon = \max\left\{\left\|f - f''\right\|_\infty, \left\|g - g''\right\|_\infty\right\}.$$

Then, the bottleneck distance between the persistence diagrams of the kernel, image, and cokernel maps is bounded above by ε:

$$d_B\left(\mathrm{Dgm}(\mathrm{grp}\ g \to f), \mathrm{Dgm}\left(\mathrm{grp}\ g'' \to f''\right)\right) \leq \varepsilon,$$

where grp represents the kernel, image, or cokernel.

2.5.4 Interpreting Persistence Diagrams

Intuitively speaking, these derived persistence diagrams tell us how important the singularities of the space are in persistent homology. A useful concept is a collection of six related persistence diagrams, referred to as a "6-pack," introduced in [20]. This 6-pack quantifies how different point sets interact and can be defined for any pair of topological spaces $L \subseteq K$ with a filtration on K.

To be more specific, let $P_{\mathcal{M}} \subseteq P$ represent the subcomplex of nonsingular points in P, as defined earlier. Consider a radius function $f_k : P \to \mathbb{R}$, and denote its restrictions to $P_{\mathcal{M}}$ and $P \setminus P_{\mathcal{M}}$ as $f_{P_{\mathcal{M}}}$ and $f_{P,P_{\mathcal{M}}}$, respectively. The radius function, along with its restrictions, generates three persistence modules. Additionally, we derive three more persistence modules from the kernel, image, and cokernel of the homology map induced by the inclusion $P_{\mathcal{M}} \subseteq P$. The persistence diagrams in a six-pack are arranged to facilitate the comparison of information across these modules.

Kernel: $\mathrm{Dgm}\,(\ker \iota_\bullet)$	**Relative:** $\mathrm{Dgm}\,(f_{P,P_{\mathcal{M}}})$	**Cokernel:** $\mathrm{Dgm}\,(\mathrm{cok}\, \iota_\bullet)$
Domain: $\mathrm{Dgm}\,(f_{P_{\mathcal{M}}})$	**Image:** $\mathrm{Dgm}\,(\mathrm{im}\, \iota_\bullet)$	**Codomain:** $\mathrm{Dgm}\,(f_k)$

Recall that the inclusion $\boldsymbol{F}_t P_{\mathcal{M}} \subseteq \boldsymbol{F}_t P$ induces a map on homology $\iota_t : H_\bullet\left(\boldsymbol{F}_t P_{\mathcal{M}}\right) \to H_\bullet\left(\boldsymbol{F}_t P\right)$. This map has a component in each dimension, k, and we denote the kernel, image, and cokernel of ι_t in dimension k by $\ker_k \iota_t$, $\mathrm{im}_k\, \iota_t$, and $\mathrm{cok}_k\, \iota_t$, respectively. For notational convenience in the following sections, we drop the $\boldsymbol{F}_t$ as it is implied by ι_t.

Induced Short Exact Sequences [20]

For each dimension k, there are short exact sequences

$$0 \to \ker_k \iota_t \to H_k\left(P_{\mathcal{M}}\right) \to \mathrm{im}_k\, \iota_t \to 0$$
$$0 \to \mathrm{im}_k\, \iota_t \to H_k\left(P\right) \to \mathrm{cok}_k\, \iota_t \to 0$$
$$0 \to \mathrm{cok}_k\, \iota_t \to H_k\left(P, P_{\mathcal{M}}\right) \to \ker_{p-1} \iota_t \to 0$$

Proof. The first two exact sequences are derived directly from the definitions of the kernel, image, and cokernel, along with the application of the First Isomorphism Theorem.

First Sequence

Consider the sequence $0 \to \ker_k \iota_t \to H_k\left(P_{\mathcal{M}}\right) \to \mathrm{im}_k\, \iota_t \to 0$. This sequence is exact because the kernel of the map ι_t injects into the domain $H_k\left(P_{\mathcal{M}}\right)$, and the image of the

map is isomorphic to the quotient of the domain by the kernel, which surjects onto the codomain.

Second Sequence

Next, consider the sequence $0 \to \operatorname{im}_k \iota_t \to H_k(P) \to \operatorname{cok}_k \iota_t \to 0$. Here, the image $\operatorname{im}_k \iota_t$ injects into $H_k(P)$, and the cokernel $\operatorname{cok}_k \iota_t$ is defined as the quotient of $H_k(P)$ by the image. The exactness of this sequence is guaranteed by the definition of the cokernel, which ensures that the map from $H_k(P)$ to $\operatorname{cok}_k \iota_t$ is surjective.

Third Sequence

To derive the third sequence, we start by recalling the long exact sequence of homology associated with a pair of subcomplexes

$$\ldots \to H_k(P_{\mathcal{M}}) \to H_k(P) \to H_k(P, P_{\mathcal{M}}) \to H_{k-1}(P_{\mathcal{M}}) \to \ldots$$

Since all homology groups are vector spaces when working with field coefficients, they naturally decompose as:

$$H_k(P_{\mathcal{M}}) \cong \ker_k \iota_t \oplus \operatorname{im}_k \iota_t.$$

Substituting this decomposition into the long exact sequence, we obtain:

$$\operatorname{im}_k \iota_t \to H_k(P) \to H_k(P, P_{\mathcal{M}}) \to \ker_{k-1} \iota_t \to 0.$$

Given that $H_k(P) \cong \operatorname{im}_k \iota_t \oplus \operatorname{cok}_k \iota_t$, we can rewrite this as:

$$0 \to \operatorname{cok}_k \iota_t \to H_k(P, P_{\mathcal{M}}) \to \ker_{k-1} \iota_t \to 0.$$

This sequence is exact because $\operatorname{cok}_k \iota_t$ injects into $H_k(P, P_{\mathcal{M}})$, and the image of this injection is precisely $H_k(P, P_{\mathcal{M}})$ modulo the kernel $\ker_{k-1} \iota_t$. □

It follows that the ranks of these groups are related:

$$\begin{aligned} \operatorname{rank} \ker \iota_t + \operatorname{rank} \operatorname{im} \iota_t &= \operatorname{rank} H_\bullet(P_{\mathcal{M}}) \\ \operatorname{rank} \operatorname{im} \iota_t + \operatorname{rank} \operatorname{cok} \iota_t &= \operatorname{rank} H_\bullet(P) \end{aligned}$$

The 1-norm of the k-dimensional persistence diagram for a function $f_K : K \to \mathbb{R}$ is defined as the total difference between the deaths and births in the diagram, denoted by

$$\|\operatorname{Dgm}_k(f_K)\|_1$$

Assuming that every class dies at some threshold C unless it dies earlier, we have the following relations between the 1-norms of the corresponding persistence diagrams.

Relations between 1-Norms [20]

For each dimension, k, and any fixed cut-off for the 1-norms, $C > 0$,

$$\begin{aligned} \|\operatorname{Dgm}_k(f_{P_{\mathcal{M}}})\|_1 &= \|\operatorname{Dgm}_k(\ker \iota_\bullet)\|_1 + \|\operatorname{Dgm}_k(\operatorname{im} \iota_\bullet)\|_1 \\ \|\operatorname{Dgm}_k(f_{P,P_{\mathcal{M}}})\|_1 &= \|\operatorname{Dgm}_k(\operatorname{cok} \iota_\bullet)\|_1 + \left\|\operatorname{Dgm}_{k-1}(\ker \iota_\bullet)\right\|_1 \end{aligned}$$

which yields a vanishing alternating sum

$$\sum_{k \in \mathbb{Z}} (-1)^k \left[\|\operatorname{Dgm}_k(f_{P_{\mathcal{M}}})\|_1 - \|\operatorname{Dgm}_k(f_P)\|_1 + \|\operatorname{Dgm}_k(f_{P,P_{\mathcal{M}}})\|_1\right] = 0$$

Proof. We first consider the sequence $0 \leq r_1 < r_2, \ldots < r_n$, which are the critical values of the function f_P that are less than a given threshold C. Let $r_0 = -\infty$ and define a cut-off value $r_{n+1} = C$ for computing the 1-norms.

For each value r_t, let $P_{\mathcal{M}_t} = f_{P_{\mathcal{M}}}^{-1}([0, r_t])$. Notice that for any $r_t \leq r < r_{t+1}$, the preimage remains constant, so the ranks of the homology groups do not change between these consecutive values. Thus, the 1-norm of the persistence diagram $\mathrm{Dgm}_k(f_{P_{\mathcal{M}}})$ can be expressed as a sum of contributions over the intervals defined by the critical values:

$$\|\mathrm{Dgm}_k(f_{P_{\mathcal{M}}})\|_1 = \sum_{i=0}^{n} (r_{i+1} - r_i) \operatorname{rank} H_k(P_{\mathcal{M}_t}).$$

Similarly, the 1-norms of the persistence diagrams corresponding to the kernel and image of $\iota_\bullet$ can be written as:

$$\|\mathrm{Dgm}_k(\ker \iota_\bullet)\|_1 = \sum_{i=0}^{n} (r_{i+1} - r_i) \operatorname{rank} \ker_k \iota_t,$$

$$\|\mathrm{Dgm}_k(\operatorname{im} \iota_\bullet)\|_1 = \sum_{i=0}^{n} (r_{i+1} - r_i) \operatorname{rank} \operatorname{im}_k \iota_t.$$

Using the short exact sequence

$$0 \to \ker_k \iota_t \to H_k(P_{\mathcal{M}}) \to \operatorname{im}_k \iota_t \to 0,$$

we obtain:

$$\|\mathrm{Dgm}_k(f_{P_{\mathcal{M}}})\|_1 = \|\mathrm{Dgm}_k(\ker \iota_\bullet)\|_1 + \|\mathrm{Dgm}_k(\operatorname{im} \iota_\bullet)\|_1.$$

Applying a similar argument to the persistence diagram $\mathrm{Dgm}_k(f_P)$ and using the short exact sequence

$$0 \to \operatorname{im}_k \iota_t \to H_k(P) \to \operatorname{cok}_k \iota_t \to 0,$$

we obtain:

$$\|\mathrm{Dgm}_k(f_P)\|_1 = \|\mathrm{Dgm}_k(\operatorname{im} \iota_\bullet)\|_1 + \|\mathrm{Dgm}_k(\operatorname{cok} \iota_\bullet)\|_1.$$

Finally, considering the relative persistence diagram $\mathrm{Dgm}_k(f_{P,P_{\mathcal{M}}})$ and the exact sequence

$$0 \to \operatorname{cok}_k \iota_t \to H_k(P, P_{\mathcal{M}}) \to \ker_{k-1} \iota_t \to 0,$$

we obtain:

$$\|\mathrm{Dgm}_k(f_{P,P_{\mathcal{M}}})\|_1 = \|\mathrm{Dgm}_k(\operatorname{cok} \iota_\bullet)\|_1 + \left\|\mathrm{Dgm}_{k-1}(\ker \iota_\bullet)\right\|_1.$$

Putting the equations together directly yields the vanishing alternating sum. □

Similar equations do not hold for the 0-norm, which counts the points in the diagrams. However, further relationships among the diagrams in a 6-pack are explored in [19], on a case-by-case analysis of how births and deaths occur simultaneously in different groups. It is important to note that although there are relations between the diagrams, a single diagram is not necessarily determined by the others. In the following example, five of the 1-dimensional persistence barcodes are identical, while the barcodes of the codomain differ.

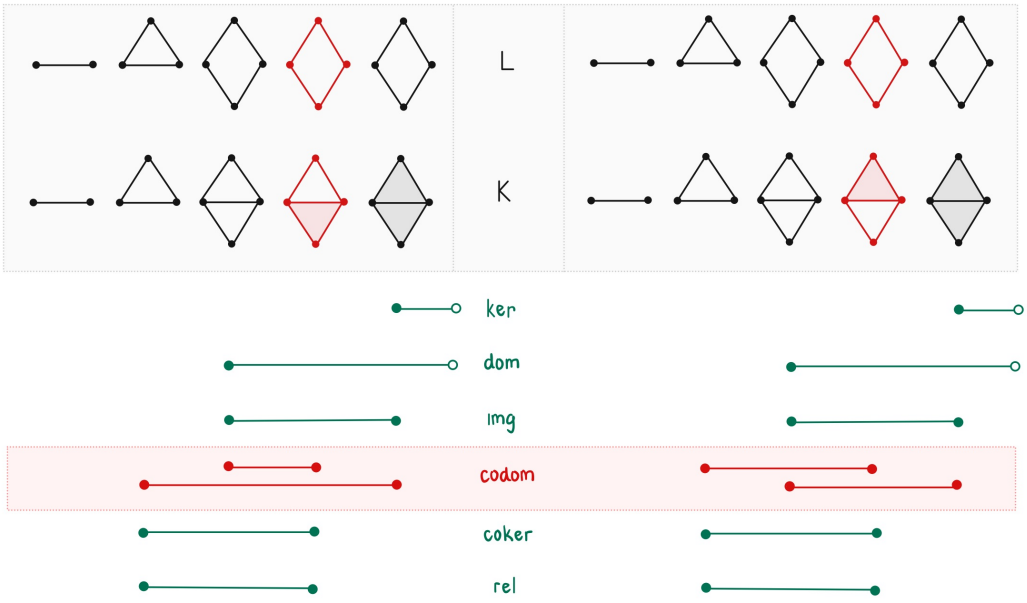


Figure 31: Example due to [20]; A single diagram is not necessarily determined by the others

2.5.5 The Manifold Hypothesis

The manifold hypothesis is a common assumption in multivariate data analysis. It proposes that data, despite being high-dimensional (with an intrinsic dimension D), actually lies on or near a lower-dimensional d-manifold M embedded within the high-dimensional space. This hypothesis is particularly supported by empirical evidence in domains like image analysis, where models often describe data in this way.

However, the practice of analyzing multivariate data suggests something more complex. For example, some real-world datasets are composed of a central "core" structure with various "flares" extending from it, indicating they may be composed of multiple manifolds, potentially with varying dimensions [13].

2.5.6 Sliding Window (Co)Kernel Diagrams

There are strategies available to test the manifold hypothesis, provided that a sufficient number of samples is available [26]. However, to analyze spaces that deviate from this hypothesis, we can use kernel, image, and cokernel persistence to quantify how much singularities interact with the topology of the space.

Suppose we have a nonempty $\mathrm{Dgm}_k(\ker \iota_\bullet)$ or $\mathrm{Dgm}_k(\operatorname{coker} \iota_\bullet)$. This implies that the singularities we have identified in our space give birth to or kill some homological features in our data. For example, consider the punctured pinched disk. Persistent points in the kernel and cokernel diagrams indicate that the singularities significantly affect the topology of the space. In the punctured pinched disk, the persistent point in the cokernel suggests that the singularity has a high level of interaction with the loop in our space (we omit the persistence diagrams associated with relative homology in the diagrams below due to limited computational resources):

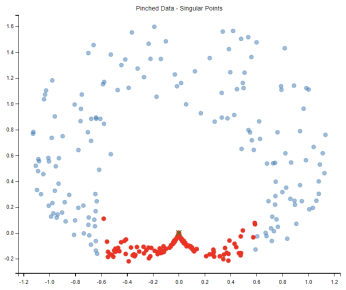


Figure 32: Punctured pinched disk with singularities

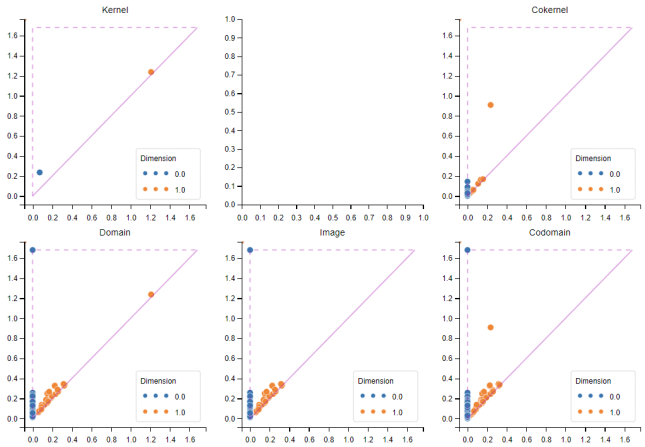


Figure 33: Punctured pinched disk 6-pack

Can we measure the amount of interaction the singular locus has with the topology of the space? To do so, we can construct a sliding window of (co)kernel diagrams by iteratively reducing the r-neighbourhood of $\Sigma_{\mathbb{X}}^{+r}$.

This allows us to distinguish between different types of singularities based on a shrinking neighborhood of points. For example, consider the punctured pointed disk. If we cut out a large enough neighborhood around the singular point, we get a relatively similar 6-pack.

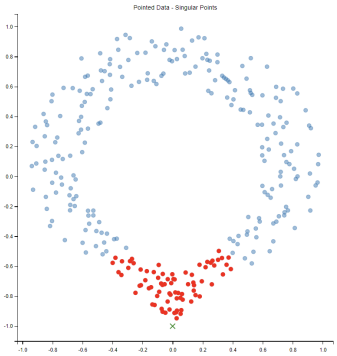


Figure 34: Punctured pointed disk with singularities

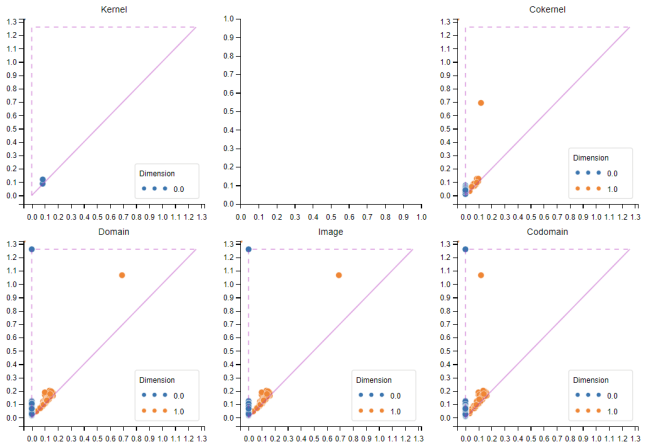


Figure 35: Punctured pointed disk 6-pack

Suppose we want to distinguish between these two singularities. We can do so by reducing the radius of the neighborhood around each point and studying the persistent points in the resulting sequence of cokernel diagrams.

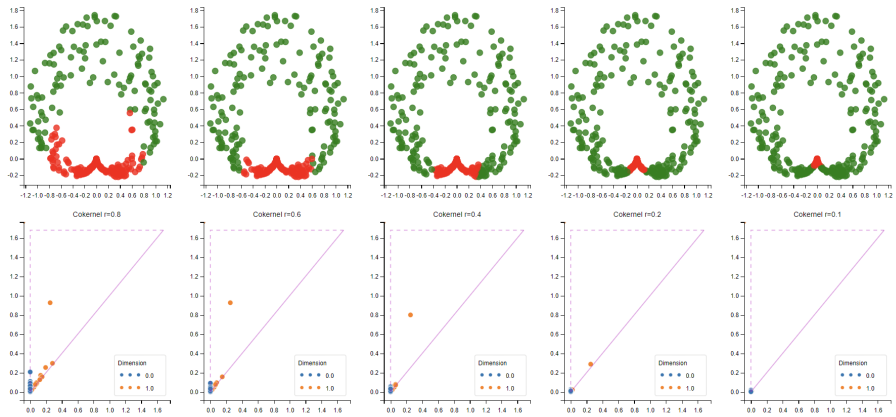


Figure 36: The sequence of persistent cokernel diagrams corresponding to shrinking the neighborhood around the singularity of the punctured pinched disk

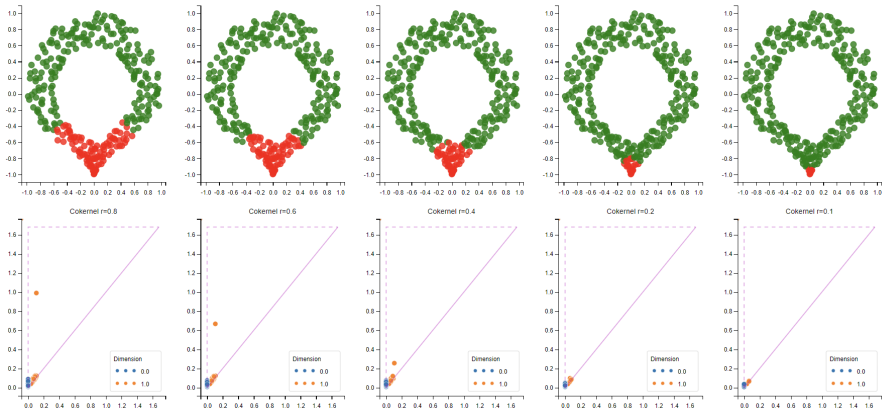


Figure 37: The sequence of persistent cokernel diagrams corresponding to shrinking the neighborhood around the singularity of the punctured pointed disk

We observe that the persistent point in the cokernel diagram of the punctured pinched disk has a significantly longer lifetime than the persistent point in the cokernel diagram of the punctured pointed disk. Therefore, we can conclude that it has significantly more interaction with the loop generating the 1-dimensional persistent homology group in our space.

Thus, we can use these techniques to quantify the effect of singularities on a space and characterize the singularities themselves.

2.5.7 ε-Clustering Assumption

We can leverage kernel persistence to test possible stratifications of our data as follows.

ε-Clustering Property

Let $P = \coprod_{i=1}^{n} M_i$ be the disjoint union of a point cloud P into n components. We say that P has the ε-clustering property if there exists $\varepsilon > 0$ such that for all $p \in M_i$

(a) $\exists q \in M_i$ with $p \neq q$ such that $d(p, q) < \varepsilon$

(b) $\forall r \in M_j$ with $j \neq i$ we have $d(p, r) > 2\varepsilon$

In other words, the point cloud P forms a reasonably nice space to cluster. We refer to the canonical decomposition of P into $P = \coprod_{i=1}^{n} M_i$ as the ε-clustering of P.

Now, let $P \subseteq \mathbb{X}$ be a finite point cloud sampled from a stratified space $\mathbb{X}$ and define $P_{\mathcal{M}} = P \cap (\mathbb{X} - \Sigma_{\mathbb{X}}^{+2\varepsilon})$. Suppose P is "connected", so for all points $i, j \in P$ we have that the pairwise distances satisfy $d(x_i, x_j) \leq \varepsilon$, and $P_{\mathcal{M}}$ satisfies the ε-clustering property.

We can describe the kernel of the bifiltration associated with $P_{\mathcal{M}} \hookrightarrow P$ as follows.

0-Dimensional Kernel Persistence

Suppose $P_{\mathcal{M}}$ decomposes into n distinct ε-clusters $P_{\mathcal{M}} = \coprod_{i=1}^{n} M_i$. Then, after computing a Vietoris-Rips filtration, the diagram $\operatorname{Dgm} k\left(\ker \iota_\bullet\right)$ will contain at least $n - 1$ distinct 0-dimensional points with persistence at least ε.

Proof. Recall for a finite metric space (P, d), the Vietoris-Rips filtration is an increasing sequence of simplicial complexes $\mathbf{VR}_r(P)$, indexed by the real numbers $r \geq 0$, where subset $\{x_0, x_1, \ldots, x_k\} \subset P$ forms a k-dimensional simplex in $\mathbf{VR}_r(P)$ if and only if the pairwise distances satisfy $d(x_i, x_j) \leq r$ for all i, j.

Let $\mathbf{VR}_r(P)$ and $\mathbf{VR}_r(P_{\mathcal{M}})$ be the VR-filtrations corresponding to P and $P_{\mathcal{M}}$ respectively. By the ε-clustering property, for all $\varepsilon \leq r < 2\varepsilon$, we have

$$\operatorname{rank} H_0(\mathbf{VR}_r(P_{\mathcal{M}})) \geq n$$

Furthermore, since all pairwise distances in P satisfy $d(x_i, x_j) \leq \varepsilon$ we have

$$\operatorname{rank} H_0(\mathbf{VR}_r(P)) = 1$$

Since we only consider 0-dimensional homology, denote

$$\ker f_{\mathbf{VR}_r(P_{\mathcal{M}})} \to f_{\mathbf{VR}_r(P)} = \{\gamma \in H_0\left(\mathbf{VR}_r(P_{\mathcal{M}})\right) \mid \iota_t(\gamma) = 0 \in H_0\left(\mathbf{VR}_r(P)\right)\}$$

Recall $\operatorname{rank}\ker \iota_t + \operatorname{rank}\operatorname{im} \iota_t = \operatorname{rank} H_0(P_{\mathcal{M}})$. Therefore, for all $\varepsilon \leq r < 2\varepsilon$ we have:

$$\begin{aligned}\operatorname{rank}(\ker f_{\mathbf{VR}_r(P_{\mathcal{M}})} \to f_{\mathbf{VR}_r(P)}) &= \operatorname{rank} H_0(P_{\mathcal{M}}) - (\operatorname{rank}\operatorname{im} f_{\mathbf{VR}_r(P_{\mathcal{M}})} \to f_{\mathbf{VR}_r(P)}) \\ &\geq n - (\operatorname{rank}\operatorname{im} f_{\mathbf{VR}_r(P_{\mathcal{M}})} \to f_{\mathbf{VR}_r(P)}) \\ &= n - 1\end{aligned}$$

Hence, the kernel will have $n-1$ components with persistence at least $2\varepsilon - \varepsilon = \varepsilon$. □

2.6 Multiparameter Persistence

Persistent homology is highly unstable to noise and outliers [8]. Many proposals have been made to deal with such issues within the framework of 1-parameter persistence. A natural solution to these problems is to consider 2-parameter persistent homology, where one of the parameters is a scale parameter, as in the Rips filtration, and the other is a density threshold. With this motivation, we briefly discuss multiparameter persistent homology.

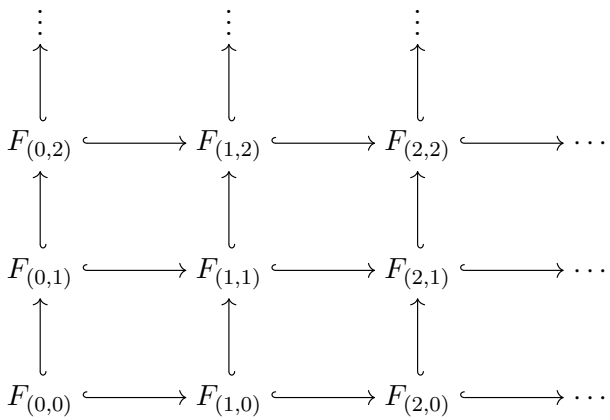
2.6.1 Multifiltrations

To develop multiparameter persistent homology, we generalize the persistent homology pipeline. Recall that persistent homology is defined on a filtration. We can extend this notion to multifiltrations, where we filter our space along multiple dimensions.

Multiparameter Filtration [10]

When the indexing set P of a filtration F is a product of totally ordered sets, $P = T_1 \times \cdots \times T_n$, we refer to F as a **multiparameter filtration** or n-parameter filtration.

Specifically, when $n = 2$, this is known as a **bifiltration**. A bifiltration is a structured collection of topological spaces arranged in a grid, where each space in the grid is nested according to the order in $\mathbb{N} \times \mathbb{N}$.



We focus on filtrations where the topological spaces F_x are simplicial complexes.

2.6.2 Degree-Rips Filtration

There are many ways of constructing multifiltrations from data. As one example, we have the following density-sensitive extension of the Rips filtration:

Degree-Rips Filtration [10]

For X a metric space, $r \geq 0$, and $d > 0$, let $\mathrm{DRips}(X)_{d,r}$ be the maximal subcomplex of $\mathrm{Rips}(X)_r$ whose vertices have degree at least $d-1$ in the 1-skeleton of $\mathrm{Rips}(X)_r$. Varying r and d, we obtain a bifiltration $\mathrm{DRips}(X)$, known as the **degree-Rips bifiltration**.

2.6.3 Multiparameter Persistence Modules

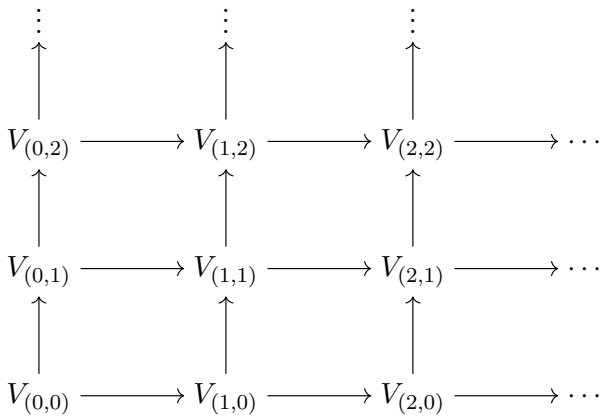
Applying homology to each space and each map in a multifiltration yields a multiparameter persistence module, exactly as in the 1-parameter case.

Let k be a fixed field, such as $k = \mathbb{Z}/2\mathbb{Z}$, and consider k-vector spaces:

Multiparameter Persistence Module

A multiparameter persistence module or an n-parameter persistence module indexed by a poset $P = T_1 \times \cdots \times T_n$ is a collection of k-vector spaces, denoted V_x for each $x \in P$, together with linear maps $V_{x,y} : V_x \to V_y$ for each pair $x \leq y$ in P. These maps satisfy the condition that for any $x \leq y \leq z$, the composition $V_{y,z} \circ V_{x,y}$ equals $V_{x,z}$.

In the specific case where $n = 2$, this is referred to as a bipersistence module. A bipersistence module can be visualized as a diagram of k-vector spaces arranged in a grid, where each space is connected to others through linear maps that respect the ordering in $\mathbb{N} \times \mathbb{N}$.



When applying the i-th homology functor with coefficients in k to each space and inclusion map in a multifiltration, the resulting persistence module captures the homological features of the filtration, structured according to the indexing poset P.

2.6.4 Invariants

Persistence modules in the single-parameter case are well understood in terms of their algorithmic and theoretical properties. However, the situation becomes considerably more complex when dealing with multi-parameter persistence modules. Unlike the single-parameter case, there is no decomposition theorem to break down any module into a direct sum of interval modules [14]. This added complexity makes it difficult to grasp their theoretical properties fully, but it also allows these modules to capture valuable information that their single-parameter counterparts might miss, making them useful in application.

Multiparameter persistence modules are challenging to analyze statistically, so we use certain algebraic properties to describe these modules instead [69]. For these properties to be reliable, they must be isomorphism invariants. Recall that an invariant is a function that maps a set of structures (such as modules) to another set, called the parameter set [14]. The key property of an invariant is that it assigns identical values to isomorphic structures.

Due to their potential, there has been a significant effort to develop practical and stable methods for working with multi-parameter persistence modules. Most existing approaches are still too computationally intensive for large-scale data. Below, we briefly describe some commonly used invariants. The most common topological multi-parameter persistence module invariants are the Hilbert function HF and rank invariant RI.

For each point in the parameter space, the Hilbert function returns the dimension of the corresponding vector space. In other words, it collects the dimensions of all vector spaces within the module, expressed as $\mathrm{HF} = x \mapsto \dim\left(H_*\left(F_x\right)\right)$.

The Hilbert Function

The **Hilbert function** $\mathrm{HF}_M : P \to \mathbb{N}$ of a persistence module M maps a point in the parameter space to the dimension of its corresponding vector space:

$$\mathrm{HF}_M : \mathbf{p} \mapsto \dim_{\mathrm{k}} M_{\mathbf{p}}$$

The rank invariant is used to estimate Betti numbers in a multifiltration, and in one dimension, it is equivalent to the barcode [14]. It works by collecting the ranks of all morphisms in the module, represented as $\mathrm{RI} = (x, y) \mapsto \mathrm{RI}\left(H_*\left(S_x\right) \to H_*\left(S_y\right)\right)$ when $x \leq_n y$. Essentially, the rank invariant counts the topological structures that are preserved as we move from x to y.

The Rank Invariant

The **rank function** $\mathrm{RI}_M : P \times P \to \mathbb{N}$ of a persistence module M takes an (ordered) pair of points in the parameter space to the rank of the morphism between them:

$$\mathrm{RI}_M : (\mathbf{p}, \mathbf{q}) \mapsto \mathrm{RI}\left(M_{\mathbf{p}} \to M_{\mathbf{q}}\right)$$

This invariant maps a pair of degrees to the rank of the map between them or assigns a value of 0 if no such map exists. The rank function recovers the homology functor if we study pairs $RI_M(\mathbf{p}, \mathbf{p})$ for all $\mathbf{p}$ [14]. However, it does not provide information about pairs of incomparable points. It cannot detect when a new homology class is created (born) or when it disappears (dies) in multipersistence modules.

The rank invariant involves calculating the persistence barcodes for every possible line in $\mathbb{R}^n$, a method known as the fibered barcode [45]. The fibered barcode is defined for a multi-parameter persistence module $\mathbb{M}$ as a map that takes a line l in $\mathbb{R}^n$ as input and outputs the persistence barcode corresponding to the single-parameter persistence module obtained by restricting $\mathbb{M}$ along l.

Fibered Barcode

Let $\mathbb{M}$ be a multiparameter persistence module. Given a discrete family of diagonal lines L (i.e., lines with direction vector $(1, \ldots, 1) \in \mathbb{R}^n$), the **fibered barcode** is

the collection of barcodes associated with the restrictions of $\mathbb{M}$ to the lines in L

$$\mathcal{FB}(\mathbb{M})_L = \{\mathcal{B}\left(\mathbb{M}|_l\right) : l \in L\}$$

3 Computation

Having covered the relevant theory, we now move on to our computational pipelines. We propose three computational approaches to handle data with singularities. More specifically, we introduce three computational approaches to address the following questions:

1. Which singularities significantly affect our space?
2. How do these singularities affect our space?
3. How do we compute more refined topological invariants of the resulting stratified space?

Our approaches leverage the basic idea behind persistence: if there is a parameter of whose value you are unsure about, rather than fix the value, vary the parameter, compute persistence, and look for intervals of stability.

We make a large underlying assumption that the singularities in our data have been identified beforehand. Singularity detection, especially when dealing with noise, is no trivial task. Fitting singular spaces to data is challenging due to the general scarcity of observations that lie precisely at singularities. Various approaches have been developed, some of which we briefly discuss below.

3.1 Singularity Detection

Stratification learning involves identifying and analyzing non-manifold points in data by modeling the data using stratified spaces. As discussed above, there exist datasets that do not satisfy the Manifold Hypothesis, which assumes that data lies on a smooth, n-dimensional surface. Instead, many datasets contain singularities—points where the local geometry does not resemble any n-dimensional Euclidean space or arises from a stratification of manifolds of possibly different dimensions. Singularities can be found in the parameter spaces of even the most basic machine learning architectures like the multilayer perceptron [2]. Thus, we seek an unsupervised representation learning framework capable of identifying singular regions in point cloud data. Various approaches have been developed to address this task, though many come with specific assumptions or are heuristic in nature [1].

Naturally, stratification learning has received significant attention from the topological data analysis community. For example, homology-based clustering approaches characterize intersection points by examining the topology of neighborhoods using persistent homology [7]. Although these methods effectively manage noise, they traditionally assume that intersecting curves meet transversely. Other approaches include local persistent cohomology [66] and sheaf theoretic approaches [55].

Recent methods, such as diffusion geometry, utilize Markov diffusion operators to extend the concept of heat flow on a manifold, enabling the development of Riemannian geometry theory for a wide range of probability spaces, including those underlying data [39]. On a Riemannian manifold $\mathcal{M}$ of dimension d, each point has a tangent space $T_x\mathcal{M}$, and the Riemannian metric is an inner product that defines this space. The tangent space $T_x\mathcal{M}$ is

d-dimensional, and the metric has exactly d positive eigenvalues, with their eigenvectors forming an orthonormal basis for $T_x\mathcal{M}$. Given a data sample from $\mathcal{M}$, we can calculate the metric at each point x and determine these eigenvalues. The k^{th} largest eigenvalue serves as an indicator of "at least k-dimensionality at x". If the data originate from a stratified space, the tangent space is well-defined where the space resembles a manifold but becomes degenerate where the manifold hypothesis fails. The largest eigenvalue of the metric can measure this local failure, remaining positive where the manifold hypothesis holds and approaching zero at degenerate points. This method's results are comparable to existing singularity detection approaches but are notably more robust to noise and outliers.

Other methods include multiple principal component analysis (PCA), which recovers cluster components formed by low-dimensional objects in high-dimensional spaces using generalized PCA [70]. When coping with non-linearity, the most studied methods are based on local PCA and kernelized versions of it [71] [16]. Local PCA methods estimate local tangent directions by analyzing covariance matrices. If a covariance matrix has at least $d + 1$ large singular values, it suggests the presence of a branching or intersection point in a local $(d + 1)$-dimensional space.

3.1.1 Method: HADES

When applying our methods, we will assume a method of singularity detection has been chosen and executed beforehand. The method of choice is up to the practitioner. Due to its computational efficiency, our experiments leverage HADES, an unsupervised algorithm that employs a kernel goodness-of-fit test to detect singularities in data [46]. The main idea of HADES is to perform the Uniformity Test at the neighborhood of each data point. Compared to topological approaches, this is significantly more computationally efficient and scalable.

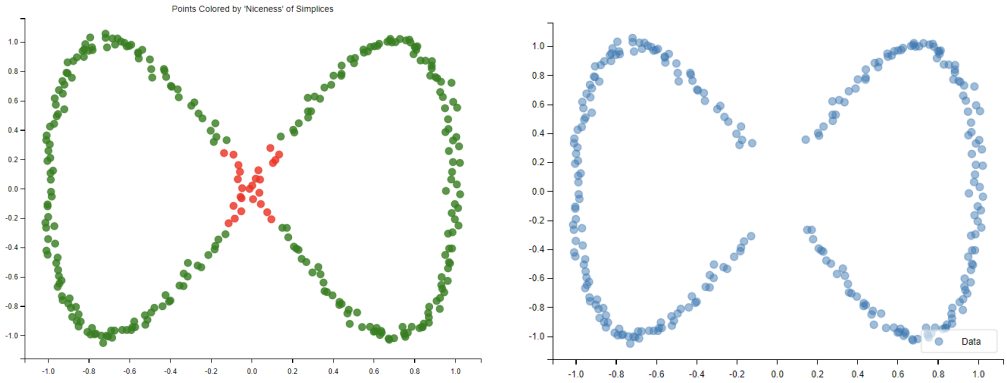
The Uniformity Test itself consists of two main steps: dimensionality reduction and a goodness-of-fit test against the uniform distribution over a disk. The dimensionality reduction is performed using PCA, where the estimated dimension of the data is determined by a threshold hyperparameter η. The local neighborhood of data points is then projected onto the principal components that account for η of the total variance.

The goodness-of-fit test involves calculating the Maximum Mean Discrepancy (MMD) between the projected data and a uniform distribution over a disk, followed by a p-value computation under a null hypothesis. The test produces two key outputs: the singularity score $\sigma(z)$, which measures the discrepancy between the data and the uniform distribution, and the singularity p-value $\tilde{\sigma}(z)$, which represents the probability that the observed discrepancy is due to chance.

The technical details of the method can be found in [46]. HADES employs an explicit formula for kernel MMD to compute the goodness-of-fit test, resulting in a linear time complexity relative to the data's dimensionality, which is a significant improvement over the exponential complexity found in existing topological methods.

3.2 Kernel, Image, and Cokernel Persistence

To begin with the question of "which singularities significantly affect our space," we turn to kernel, image, and cokernel persistence. Consider a point cloud P sampled from a stratified space $\mathbb{X}$, and let $\Sigma_{\mathbb{X}}$ represent the set of singular points identified by a given stratification learning algorithm. Define $P_{\mathcal{M}} = P - \Sigma_{\mathbb{X}}$ as the point cloud obtained after removing the singular points from P. As a running example, let P be the figure-8. After removing the singular points, we obtain $P_{\mathcal{M}}$:

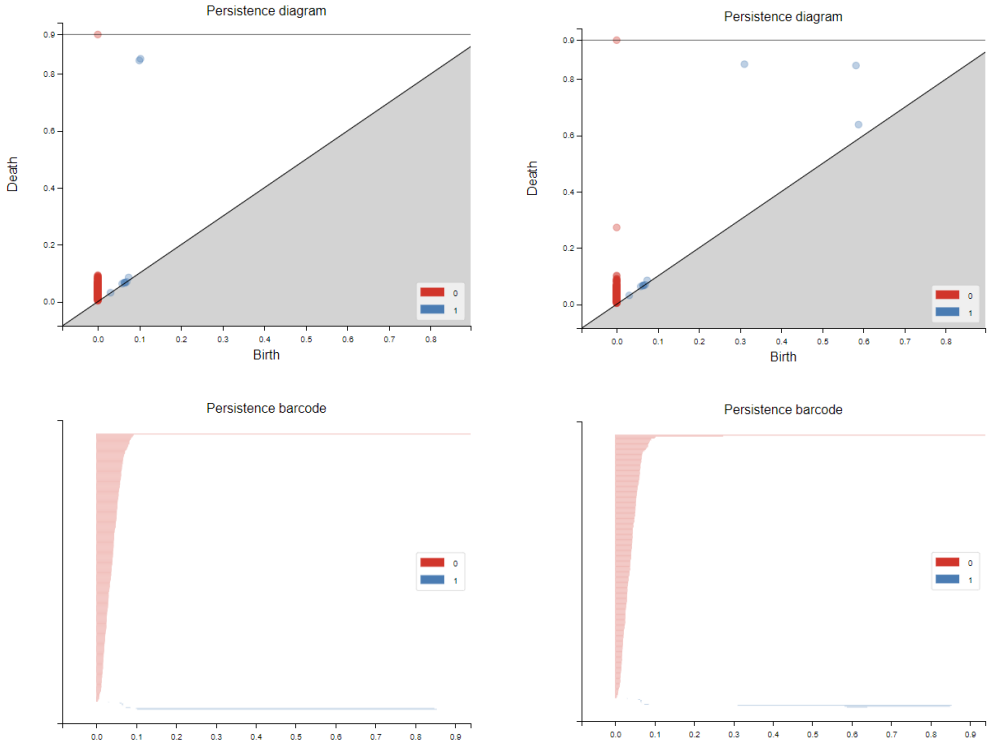


(a) Let P be the figure-8 (b) Define $P_{\mathcal{M}} \subseteq P - \Sigma_{\mathbb{X}}$

Figure 38: Comparison of the figure-8 singularities and pieces

3.2.1 Bifiltrations

We first construct two Vietoris-Rips filtrations of the original point cloud P and the point cloud with singularities removed $P_{\mathcal{M}} \subseteq P$. The persistence diagrams and barcodes corresponding to the individual sequences are as follows:



(a) Barcode corresponding to P (b) Barcode corresponding to $P_{\mathcal{M}}$

Figure 39: Comparison of the barcodes corresponding to the figure-8 P and $P_{\mathcal{M}}$

3.2.2 Image Persistence

We then compute kernel, image, and cokernel persistence and identify persistent points in the corresponding persistence diagrams. Persistent points indicate that the singularities

significantly affect the topology of the space.

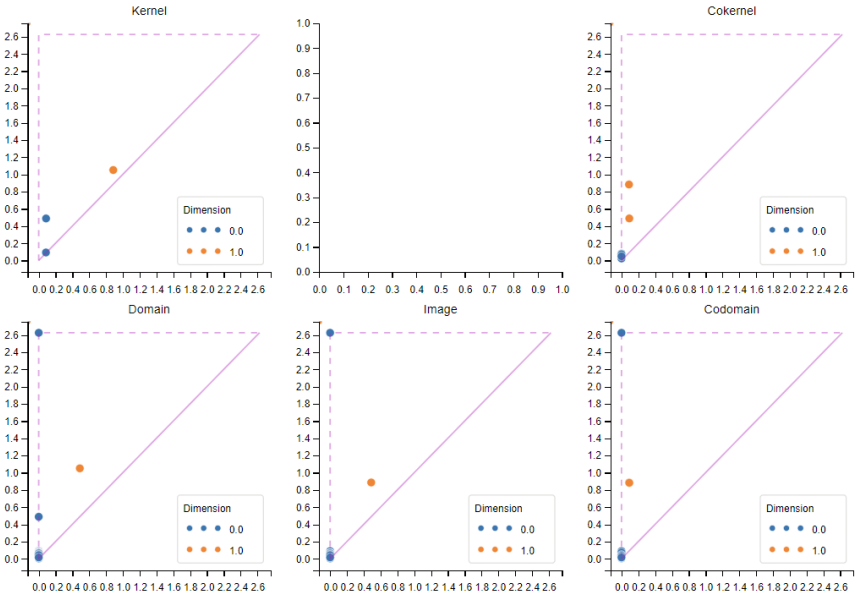


Figure 40: The 6-pack corresponding to the figure-8

In the figure-8 example, we observe one persistent point in the 0-dimensional kernel persistence diagram (which corresponds to one connected component) and two persistent points in the 1-dimensional cokernel persistence diagram (which correspond to loops). The 0-dimensional point in the kernel implies that the singularity is "killing" a component (since adding back the singularity connects two components). On the other hand, the two 1-dimensional points in the cokernel imply that the singularity gives birth to two loops in the space (since the act of "removing the singularity" kills two loops in the space). Thus, persistent points in the kernel indicate we have some critical singularities in our space that separate connected components, loops, etc, and persistent points in the cokernel tell us how these singularities affect the topology of the entire space.

If one wishes to study individual singularities, one can remove them one at a time and compute the corresponding 6-pack. Furthermore, as discussed above, one can also construct a sequence of persistence diagrams based on a shrinking neighborhood around the singularities to characterize the singularities themselves.

3.3 Multiparameter Persistence

Having identified which singularities affect our space, we turn to the question of "how do these singularities affect our space". To answer this, we leverage multiparameter persistence. We continue using the running example of the figure-8.

Having constructed parallel filtrations and applied image persistence, we can define a simple function to extend our construction to a multiparameter persistence module. We define $f : P \to \{0, 1\}$ by

$$f(p) = \begin{cases} 0 & \text{if } p \in P_{\mathcal{M}} \\ 1 & \text{otherwise} \end{cases}$$

We can then construct a bilfiltration along both f and the previous Vietoris-Rips filtration. This allows us to compute more general invariants of multiparameter persistence.

Furthermore, the function f can easily be extended to more dimensions, and the techniques described here continue to apply. For instance, one could order singularities by their dimension d and apply $f : P \to \{0, 1, \cdots, d\}$. Alternatively, one could define a perversity function $\bar{q}$ and define $f(p) = \bar{q}(p)$. We leave this to future work.

3.3.1 Candidate Decompositions

To make this task computationally feasible, we utilize a recently introduced family of stable invariants known as candidate decompositions for multi-parameter persistence modules [47]. Candidate decompositions are interval decomposable multi-parameter persistence modules that are controllable approximations, parameterized by a precision parameter $\delta > 0$. It is possible to quantify the approximation error between the candidate decompositions and the actual underlying module using standard interleaving and bottleneck distances. Analogous to the single parameter case, each component of the candidate decomposition

$$\tilde{\mathbb{M}}_\delta = \bigoplus_{i \in \tilde{\mathcal{I}}} \tilde{I}_i$$

is an interval summand in $\mathbb{R}^n$. For technical details, we refer the reader to [47].

3.3.2 Multiparameter Module Approximation (MMA)

Taking the figure-8 above as our running example, we can compute its candidate decomposition using MMA [43] To get an idea of the module's shape, we then visualize the module approximation as follows.

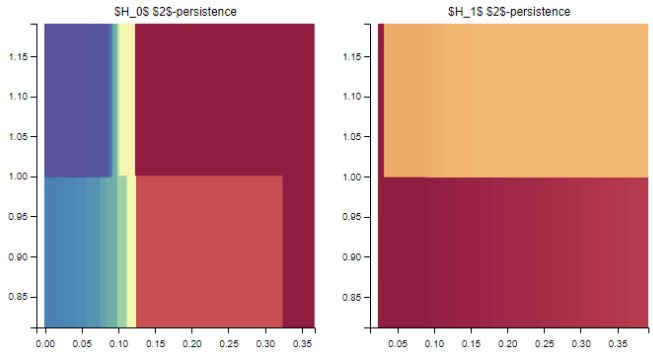


Figure 41: The figure-8 module approximation. The y-axis corresponds to our function f, and the x-axis corresponds to the Vietoris-Rips filtration

Each interval, represented by a colored block, visually corresponds to the lifetime of a cycle in the bifiltration. Conceptually, this is a two-dimensional barcode, where the y-axis corresponds to our function f and the x-axis corresponds to the Vietoris-Rips filtration.

For example, we observe a significant component in the 0-dimensional barcode of the figure-8 (the orange block) which is eliminated when we reintroduce the singularity (visualized by blocks with $y > 1$). This orange block corresponds to the two components we observe in the zero-dimensional persistent homology of $P_{\mathcal{M}}$ (which, for the same parameters in the Vietoris Rips filtration, are connected in P). In fact, we observed this difference between P and $P_{\mathcal{M}}$ as a persistent point in the kernel persistence diagram earlier. However, the module approximation allows us to easily identify the parameters of the Vietoris-Rips filtration for which this topological feature is significant. In addition, it

allows us to easily identify the parameters in the VR filtration for which the Betti numbers of P and $P_{\mathcal{M}}$ are identical (which corresponds to colored blocks lining up along the y-axis).

3.3.3 Signed Barcode Decompositions as Signed Measures

Another invariant we can compute is the signed barcode, which offers weaker topological invariants that can be applied in a machine learning context as signed measures [11]. For technical details, we refer the reader to [11]. Consider the Hilbert signed measure derived from the dimension vector of a 2-parameter module.

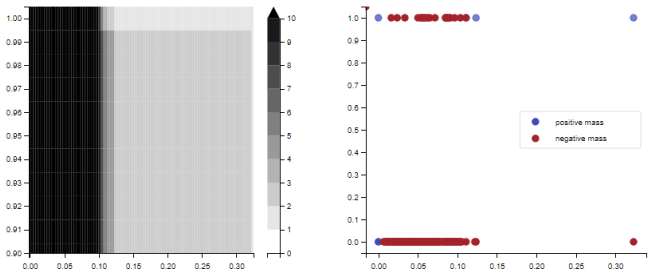


Figure 42: The figure-8 signed barcode

In the figure above, the left side shows the pointwise dimension of the bimodule (the Hilbert Function), while the right side displays the associated signed measure, representing the "changes" in the values of the Hilbert function. In the signed measure on the right, each blue dot indicates "something that appears in homology," and each red dot represents "something that disappears in homology."

Thus, equipping our space with a bilfiltration corresponding to f and Vietoris-Rips allows us to easily compute the invariants of the corresponding candidate decomposition. We leave the discussion of all these invariants and their potential applications to machine learning for future work.

3.4 Persistent Intersection Homology

Lastly, we turn to the question of how to compute more refined topological invariants of the resulting stratified space. To answer this, we identify singular points using stratification learning, use multiparameter persistence to identify manifold pieces in the resulting space, and then apply persistent intersection homology. Existing implementations of persistent intersection homology employ a simplex removal procedure known as ϕ-persistence [63]. This exists as a convenient abstraction that helps compute the more concrete notion of intersection homology persistence. One could, however, consider alternative simplex removal decision procedures other than those derived from a perversity.

Determining which simplices to allow when dealing with noisy data is not straightforward. The Vietoris-Rips complex is commonly used in topological data analysis to handle multivariate data sets, but when applied to persistent intersection homology, it can produce unexpected results. The following example is due to [63]. Consider the wedge sum of two circles $S^1 \vee S^1$. This space can be represented as a simplicial complex K, where the smallest stratification places the singular point x in its own subspace: $X_0 = \{x\}$ and $X_1 = K$. With $\bar{p} = (-1)$, the intersection homology of K gives $\beta_0 = 2$.

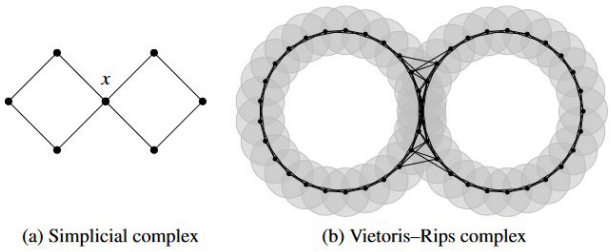


Figure 43: The Vietoris-Rips complex "smooths out" isolated singularities [63]

However, when calculating the persistent intersection homology of a point cloud representing this space, the result is always $\beta_0 = 1$. This occurs even if we make the triangulation flaglike by computing the first barycentric subdivision, ensuring the calculations are independent of the stratification. The discrepancy arises because the topological realization of the Vietoris-Rips complex is more closely aligned with regular neighborhoods than the actual homeomorphism type of $S^1 \vee S^1$. A regular neighborhood is always a manifold and so can be thought of as a "thickened" version of the space, where isolated singularities are smoothed out.

To address this issue, we would like to cluster the strata and forbid interaction between the resulting manifold pieces based on a perversity function. We outline our approach to exploring different clusterings and perversities, which we discuss in more detail below. Our general strategy is as follows. First, we remove singularities from our point cloud, either by manual inspection or using HADES [46]. Next, we cluster the remaining manifold pieces via Persistable [19] or AutToMATo [36] and compute intersection homology. We then reintroduce singularities one at a time by constructing a perversity sequence and recomputing intersection homology at each step. The resulting diagrams allow us to study the persistent intersection homology of our space and the effect of adding back singularities. Due to limited computational resources, we restrict ourselves to 0 and 1-dimensional homology.

3.4.1 Clustering with Multiparameter Persistence

Clustering techniques are central to understanding and interpreting data across different fields. The basic idea is to group objects together based on a defined notion of similarity, usually measured by a distance or metric within the dataset. There are many clustering methods to choose from, including hierarchical, centroid-based, and density-based techniques.

Density-based clustering is a method used to group data points based on their density, often through hierarchical clustering. Recall that degree-Rips is defined as a two-parameter filtration of simplicial complexes. In the context of clustering, only the underlying graphs are of interest. Degree-Rips clustering recovers well-known algorithms like DBSCAN but also faces stability issues [25]. To address this, [64] proposes a new, more stable method called kernel linkage, which varies both parameters in degree-rips clustering and introduces a metric called the correspondence-interleaving distance. The method is part of a broader pipeline for density-based clustering, called Persistable [65], which provides a more robust framework for clustering in topological data analysis.

Persistable allows one to interactively investigate possible clusterings based on a component counting function and prominence vineyard. This is particularly valuable when there may not be a clear-cut, "correct" solution for how a clustering should look. The

component counting function displays the number of clusters in the data based on a distance scale parameter and a density threshold, which are closely related to the parameters used in the DBSCAN algorithm. However, unlike DBSCAN, Persistable doesn't fix these parameters. Any line in the parameter space determines a one-parameter hierarchical clustering of the data. Persistable relies on a sequence of such "slices" of the parameter space, computing persistent homology along each line. A "prominence vineyard" is then constructed by taking a family of slices that interpolate between two user-selected slices. The curves ("vines") in the prominence vineyard represent clusters in the data that evolve with the choice of slice [18]. The larger their prominence value, the more likely they are to represent real structure in the data. We refer the reader to [64] for technical details.

Returning to our figure-8 example above, we plot the component counting function and a corresponding prominence vineyard below.

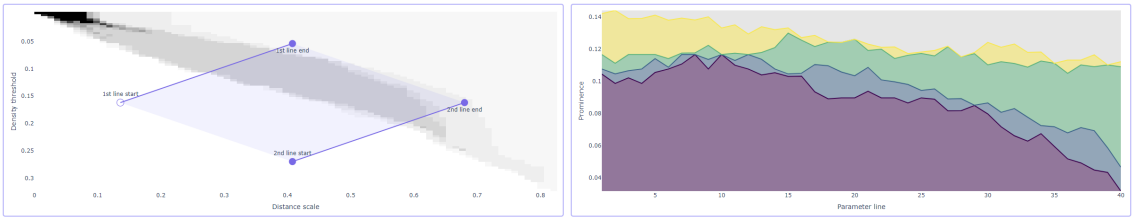


Figure 44: The figure-8 component counting function and prominence vineyard

We can then select a set of parameters and visualize the resulting clusters. For example, the two-component clustering corresponding to the initial line in the diagram above is visualized as follows.

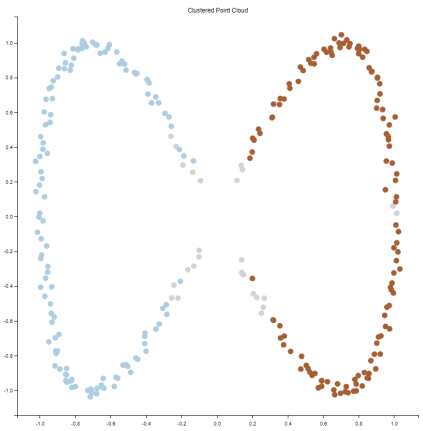


Figure 45: A possible clustering of the figure-8 corresponding to the Persistable output above

An alternative approach is the topological clustering algorithm ToMATo (Topological Mode Analysis Tool), which is designed to capture the prominences of peaks in a density function using persistence diagrams [15]. It works by combining a graph-based hill-climbing algorithm with a cluster-merging step guided by persistence. Hill-climbing can be highly unstable to small changes in an estimated density function $\tilde{f}$. By computing the persistence diagram of $\tilde{f}$, ToMATo quantifies the prominences of its peaks, allowing users to distinguish significant peaks of the true density f from less important ones. The user then selects a prominence threshold τ, retaining only those peaks with prominences

above this threshold, which results in the final clustering. The persistence diagram clearly illustrates how the choice of τ impacts the number of clusters obtained.

However, one of the challenges in clustering, especially in real-world applications, is selecting the right parameters—a process known as hyperparameter tuning. One way to avoid the complexity of choosing hyperparameters is through a bottleneck bootstrap on the persistence diagram produced by ToMATo, which is precisely what AuToMATo does [36]. After generating a persistence diagram from a point cloud using ToMATo, AuToMATo creates a confidence region for that diagram based on the bottleneck distance. This confidence region helps determine the optimal choice of τ, ultimately guiding a final (parameter-free) clustering.

Applying this to our figure-8 example results in the following clusters.

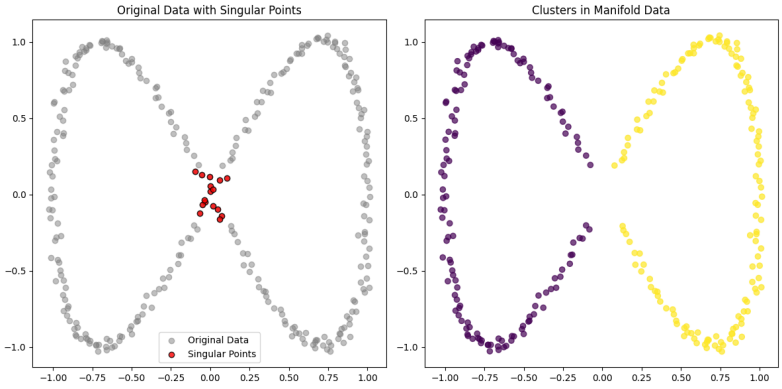


Figure 46: The AuToMATo clustering of the figure-8

3.4.2 Persistent Intersection Homology

After identifying singularities, we cluster the corresponding strata using AuToMATo, label singular points according to their nearest cluster, and plot the resulting intersection homology groups. For example, consider the following two overlapping circles.

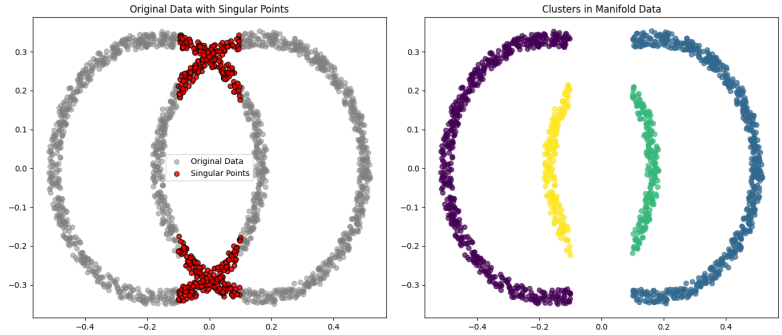


Figure 47: The AuToMATo clustering of the two overlapping circles

The persistent intersection homology groups which forbid all intersections with the singular locus are then computed as follows:

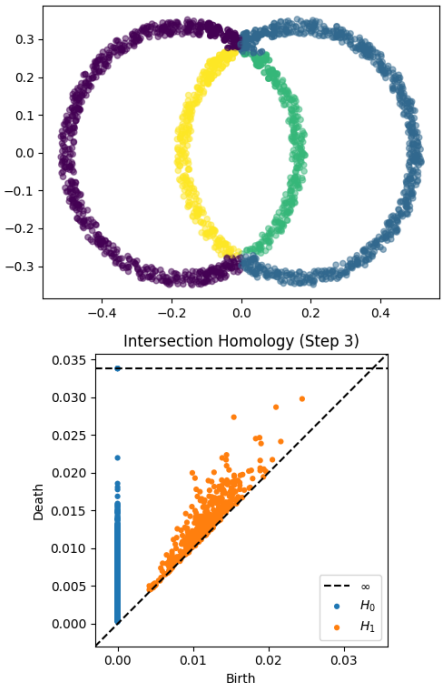


Figure 48: The persistent intersection homology groups of the two overlapping circles which forbid all intersections with the singular locus, where singular points are labeled according to their nearest cluster

Although this approach can be informative in some cases, we observe that we lose track of the loops in our space by forbidding all interaction with singularities. To address this issue, we construct a sequence of perversities to reintroduce allowable singularities one at a time.

3.4.3 Perversity Sequence

We first define a filtered space based on the singular locus. More specifically, we separate the singular locus into a disjoint union of singularities $\Sigma_{\mathbb{X}} = \coprod_{i=0}^{n} \Sigma_i$ and construct a filtration reintroducing each singularity one by one:

$$\{\Sigma_0\} \subseteq \{\Sigma_0 \cup \Sigma_1\} \subseteq \cdots \subseteq \{\Sigma_0 \cup \Sigma_1 \cup \cdots \cup \Sigma_n\} = \Sigma_{\mathbb{X}} \subseteq P$$

Although we don't specify an order, one could, for example, order the singularities by dimensionality. In practice, a filtration of the singular locus can be based on clustering individual singularities. For example, the filtration corresponding to a k-means clustering of the two singularities in the two overlapping circles is as follows.

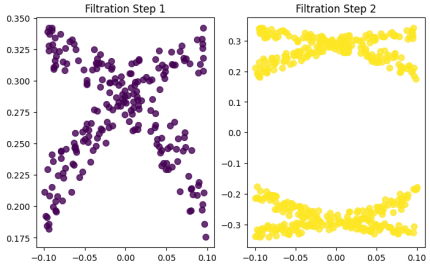


Figure 49: Filtration of the singularities in the two overlapping circles

We can define a sequence of perversity functions to "reallow" intersections with singularities one at a time. Let $\{\overline{p}\} = \overline{p^0}, \overline{p^1}, \ldots, \overline{p^n}$ be defined by

Algorithm 1 Perversity Sequence Initialization and Update

$\overline{p^0}_k \leftarrow k - \max\{i\} - 1$ ▷ Initialize perversity sequence disallowing all intersections, where $\max\{i\}$ is the maximum dimension of the i-chains considered.
for $t \in \{0, n\}$ do
  if $t = 0$ then
    $\overline{p^t} \leftarrow \overline{p^0}$
  else
    $\overline{p^t} \leftarrow \overline{p^{t-1}}$
  end if
  $k \leftarrow t$
  $\overline{p^t}_k \leftarrow k$ ▷ Update codimension $k = t$ to allow intersections.
end for

By starting with the sequence $\{0, -1, -2, \cdots, -n\}$, we effectively forbid all intersections since

$$\begin{aligned}(i-k) + (k - \max i - 1)\\ = i - \max i - 1\\ < 0\end{aligned}$$

From here, we enable codimensions one at a time since for any i-chain σ, we have

$$\begin{aligned}(i-k) + (k)\\ = i\\ \geq \dim(\sigma \cap X_{n-k})\end{aligned}$$

Furthermore, since $\overline{p^t} \leq \overline{p^{t+1}}$ at each step in the sequence above, we induce homomorphisms of the corresponding intersection homology groups. We can then track changes in intersection homology as we add back singularities. Consider the example of two overlapping circles above.

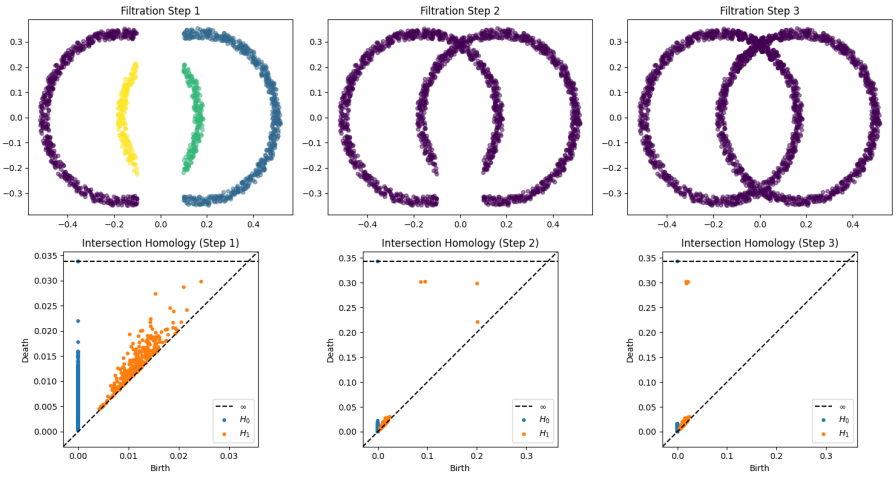


Figure 50: The sequence of persistent intersection homology diagrams of the two overlapping circles corresponding to reintroducing allowable singularities

We can observe how the singularities we add back interact with the topological features in our space by reintroducing singularities and tracking changes along the corresponding intersection homology groups. For example, we see that the singularity we add back at "Filtration Step 2" in the diagram above connects all of the components from "Filtration Step 1" and reintroduces three loops in our space.

4 Application

Current trends in Natural Language Processing (NLP) involve using highly complex models, such as neural networks and word embeddings [58]. Although these models achieve state-of-the-art performance across various tasks, their lack of interpretability is a major drawback.

We apply our computational approaches to the study of word embeddings. It has been argued that word vectors lie on a pinched manifold, that is, the quotient of a manifold obtained by identifying some of its points [38]. These singular points correspond to polysemous words, i.e., words with multiple meanings, which suggests that monosemous and polysemous words can be differentiated based on the topology of their neighborhoods.

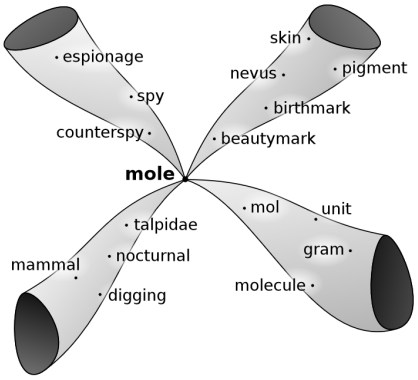


Figure 51: An idealized picture of the word "mole" [38]

For example, consider the word disambiguation problem, which refers to determining the number of possible meanings of a word. Recent work has employed local homology to study the word disambiguation problem [67], which has been extended to employ persistent homology [38]. We propose extending this to kernel/image/cokernel persistence, multiparameter persistence, and persistent intersection homology, all of which can detect more refined topological features in the presence of singularities.

4.1 Word Embeddings

Word embeddings represent words as vectors in a multi-dimensional space, where the distance and direction between these vectors indicate the similarity and relationships between the words. This approach contrasts with traditional methods like one-hot encoding, which fails to capture semantic information or relationships between words. In contrast, word embeddings are dense vectors with continuous values, typically learned through machine learning techniques, often involving neural networks. These embeddings are trained on large text datasets, allowing the model to adjust the vectors based on the contexts in which words appear.

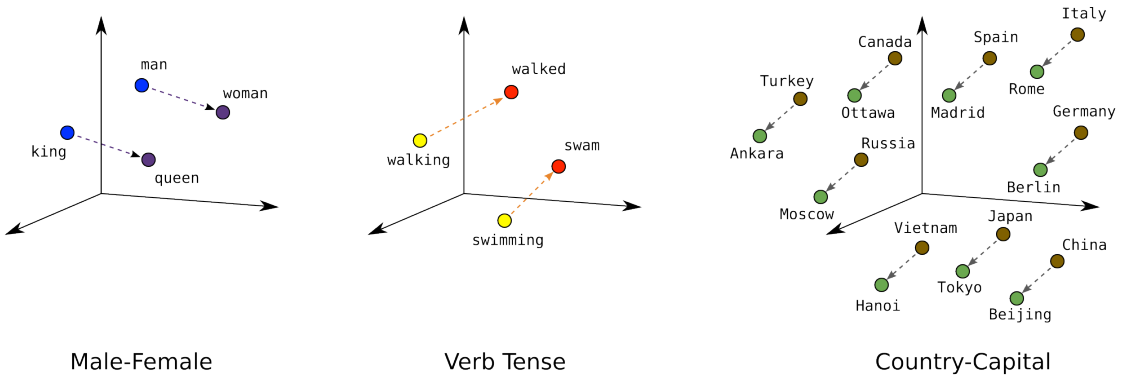


Figure 52: "Embeddings can produce remarkable analogies"; Image from developers.google.com

Word embeddings primarily aim to create numerical representations of words that reflect their semantic relationships and contextual meanings. Popular methods for training word embeddings include Word2Vec [53], which uses a neural network to predict surrounding words, and GloVe (Global Vectors for Word Representation) [59], which creates embeddings based on global text statistics.

The success of methods like Word2Vec and GloVe has led to the development of more advanced language models, such as FastText, BERT, and GPT. These models incorporate techniques like subword embeddings, attention mechanisms, and transformers to handle more complex embeddings. However, for simplicity in computational experiments, we will focus on GloVe embeddings and leave the exploration of these advanced models for future work. We briefly discuss Word2Vec and GloVe below, omitting technical details. For a more thorough exposition, the reader is referred to [24].

4.1.1 Word2Vec

Word2Vec, a method developed by a team of Google researchers in 2013 [53], includes two primary models for generating vector representations of words: Continuous Bag of Words (CBOW) and Continuous Skip-gram. The CBOW model predicts a target word based on its surrounding context words within a specified window, capturing the semantic relationships between them. In contrast, the Continuous Skip-gram model works in the opposite direction, taking a target word as input and predicting its surrounding context words.

However, Word2Vec has some limitations. One of the main challenges is handling polysemy; the model tends to average or blend the representations of different senses [12]. More advanced models like FastText, GloVe, and transformer-based models have been developed to address these issues.

4.1.2 GloVe

GloVe (Global Vectors for Word Representation) is a word embedding model that captures global statistical patterns of word co-occurrences within a corpus. Developed by Jeffrey Pennington, Richard Socher, and Christopher D. Manning in 2014 [59], GloVe differs from Word2Vec by leveraging global information instead of concentrating only on local context. The core idea behind GloVe is that understanding word semantics requires considering

the overall statistics of word co-occurrence across the entire corpus. Rather than just examining the immediate context around individual words, GloVe accounts for how often words appear together throughout the whole dataset. The model aims to minimize the difference between the predicted co-occurrence probabilities and the actual probabilities derived from corpus statistics.

4.2 Dimensionality Reduction

Word embeddings are usually represented in high-dimensional spaces, such as 100, 300, or even 768 dimensions, which makes them hard to interpret or visualize. To address this, it is common to apply various dimensionality reduction techniques, allowing one to visualize how words relate. This approach helps uncover semantic clusters, word analogies, and patterns within the data. Additionally, reducing dimensions helps lower the computational cost of processing high-dimensional data. We employ Uniform Manifold Approximation and Projection (UMAP) [52] for dimensionality reduction.

4.2.1 Method: UMAP

Uniform Manifold Approximation and Projection (UMAP) is a visualization and dimensionality reduction technique based on manifold learning. We use UMAP for dimensionality reduction for two key reasons. First, it is designed to preserve the topological structure of the data. Second, UMAP effectively preserves both local structure (larger semantic groupings) and global structure (capturing broader semantic groupings). This makes it a good choice for balancing the representation of local neighborhoods with an accurate representation of the overall data structure.

UMAP leverages local manifold approximation, combining them through their local fuzzy simplicial set representations to form a topological representation of high-dimensional data. This process is mirrored in a lower-dimensional space, where UMAP constructs a corresponding topological representation based on an initial low-dimensional representation of the data. The optimization phase of UMAP then adjusts this low-dimensional representation to reduce the cross-entropy between the high-dimensional and low-dimensional topological representations. The reader is referred to [52] for technical details.

4.3 Experimental Results

We study the "gensim/glove-twitter-200" pre-trained word embedding model available in the Gensim library [77]. It consists of pre-trained GloVe vectors of dimension 200 that are derived from 2 billion tweets, encompassing 27 billion tokens and a vocabulary of 1.2 million words, all in an uncased format.

4.3.1 Image Persistence

We begin by comparing words and embedding dimensions using image persistence. In particular, we use image persistence to investigate how the persistent kernel diagram changes as the dimension of the embedding space is reduced.

Our pipeline is as follows. We get the local neighborhood of a word by selecting the 50 most similar words based on cosine similarity. We project the selected neighborhood to 25, 10, 5, and 2 dimensions using UMAP, constraining the size of the local neighborhood to 5. We then further remove a hypothetical "singular neighborhood" around the word by selecting and removing the ten most similar words. Finally, we compute the 6-pac ks of the resulting spaces, using a maximum edge length of 2 to track local changes. It's

important to note that the size of the neighborhoods and chosen projection dimensions are arbitrary. In the future, we will need to systematically explore various nonlinear dimensionality reduction approaches, neighborhood sizes, and other factors.

As an example, consider the words pound' and purple.' According to WordNet, a large lexical database of the English language, words are grouped into synonyms called "synsets." Each synset represents a unique concept [54]. The word "pound" has 22 senses, and the word "purple" has 7. Because of the difference in the number of senses, one might conjecture that there are more 0-dimensional components in the kernel persistence diagram of "pound" than in that of "purple."

We begin by listing some words in the neighborhoods of "pound" and "purple" according to cosine similarity below. The most similar words to "purple" are all colors. In contrast, we interpret "dollar" and "lb" as having different senses.

	word	similarity
0	pound	1.000000
1	pounds	0.651422
2	dollar	0.613808
3	lb	0.546180
4	million	0.527864

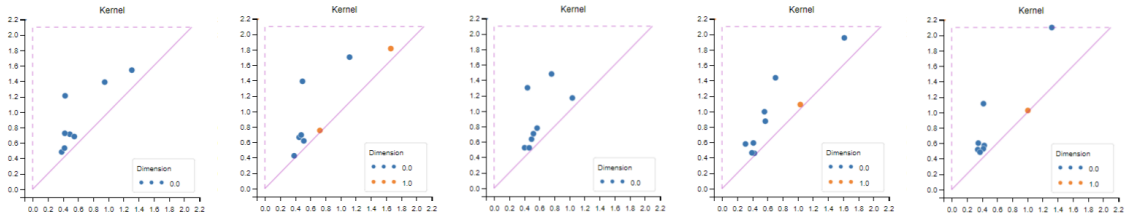
(a) A 5 element neighborhood of "pound"

	word	similarity
0	purple	1.000000
1	pink	0.831380
2	blue	0.820810
3	yellow	0.772634
4	green	0.755454

(b) A 5 element neighborhood of "purple"

Figure 53: Comparison of 5 element neighborhoods for the words "pound" and "purple"

Next, we plot the persistence kernel diagrams corresponding to the 25, 10, 5, and 2-dimensional UMAP projections of the word "pound" below.



(a) 25 dimensions (b) 10 dimensions (c) 5 dimensions (d) 3 dimensions (e) 2 dimensions

Figure 54: Kernel persistence of the UMAP projections of the word "pound" in different dimensions

These persistent kernel diagrams enable us to observe how the topology of the space changes as we perform dimensionality reduction. For example, we have at least approximately three persistent points in the persistent kernel of the 25-dimensional embedding, which implies that when we assume the word "pound" is a singularity, it may be connected to at least four components. In contrast, we only observe two such persistent points in the 10-dimensional embedding. The 6-pack corresponding to the 10-dimensional UMAP projection of the word "pound" is illustrated as follows.

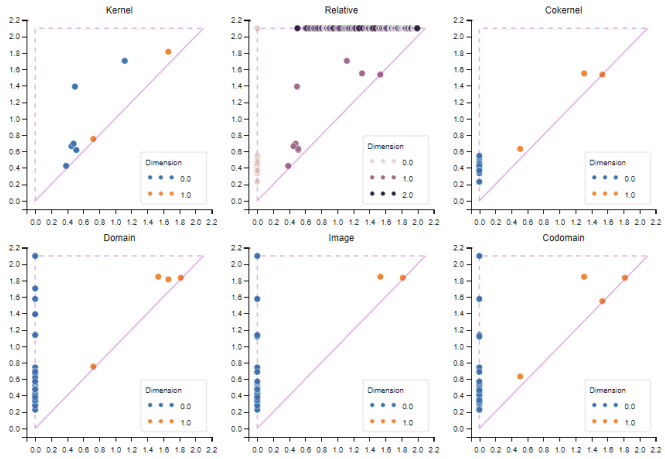
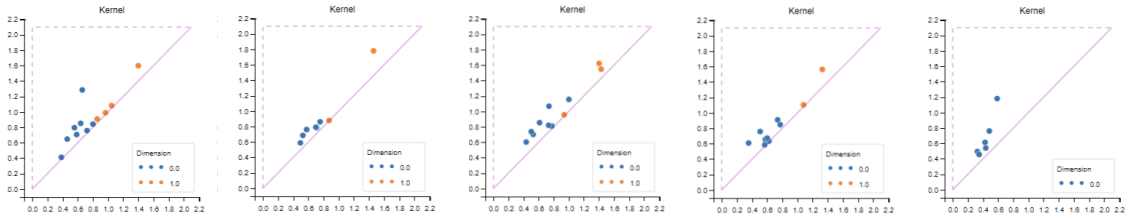


Figure 55: 10-dimensional UMAP projection of the word “pound”

Next, consider the word “purple”. We plot the corresponding persistence kernel diagrams below.



(a) 25 dimensions (b) 10 dimensions (c) 5 dimensions (d) 3 dimensions (e) 2 dimensions

Figure 56: Kernel persistence of the UMAP projections of the word “purple” in different dimensions

We observe significantly fewer zero-dimensional points in the persistent kernel diagrams of the word “purple”. The 6-pack corresponding to the 10-dimensional UMAP projection of “purple” is illustrated as follows.

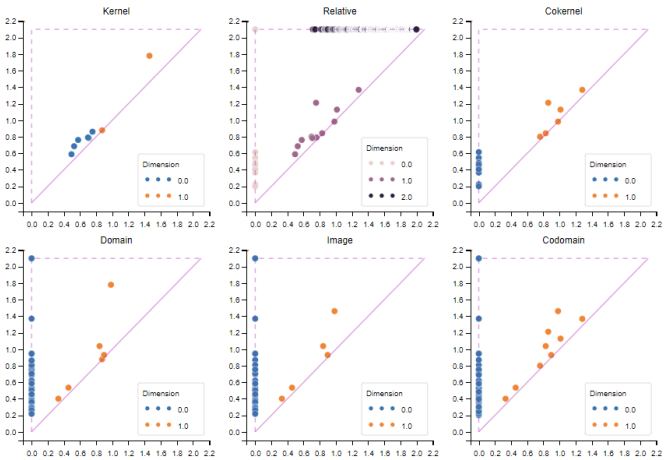
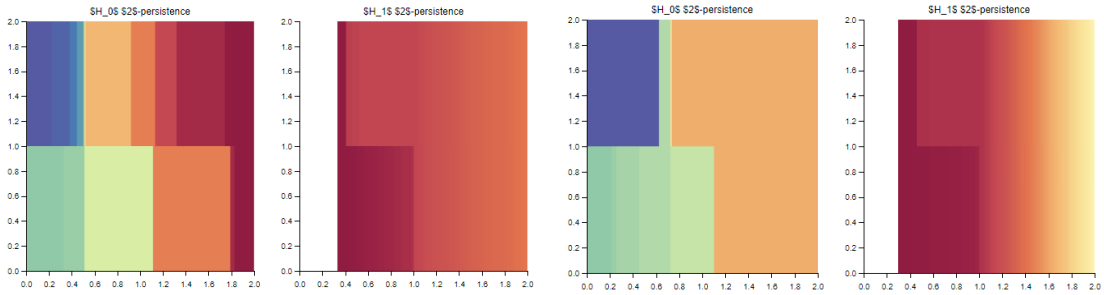


Figure 57: 10-dimensional UMAP projection of the word “purple”

4.3.2 Multiparameter Persistence

We turn to multiparameter persistence to further investigate the 10-dimensional UMAP projections of the words "pound" and "purple." Using the same sequence of nested filtrations described above, we can construct a multiparameter persistence module and compute a candidate decomposition using multiparameter module approximation. We then visualize the module approximations as follows.

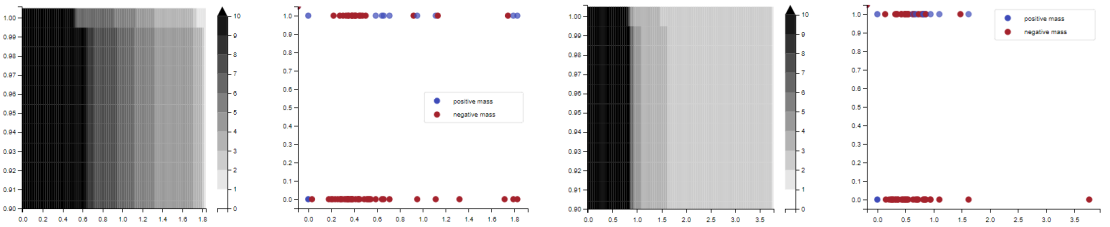


(a) 10-dimensional UMAP projection of a neighborhood of the word "pound"

(b) 10-dimensional UMAP projection of a neighborhood of the word "purple"

Figure 58: Side-by-side comparison of multiparameter module approximations for "pound" and "purple".

We are also able to compute various invariants, such as the signed barcode with the Hilbert signed measure.



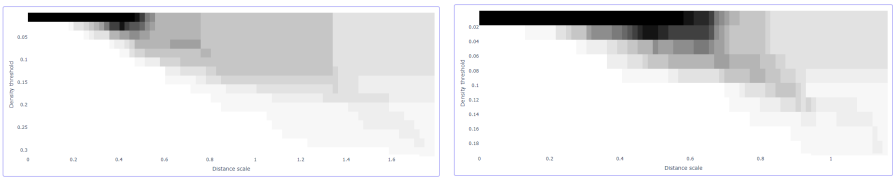
(a) 10-dimensional UMAP projection of a neighborhood of the word "pound"

(b) 10-dimensional UMAP projection of a neighborhood of the word "purple"

Figure 59: Side-by-side comparison of Hilbert signed measures for "pound" and "purple".

4.3.3 Intersection Homology

Lastly, we compute more refined topological invariants of the resulting space. We begin with multiparameter density-based clustering using Persistable [65]. The component counting functions are illustrated below.



(a) The component counting function corresponding to “pound”

(b) The component counting function corresponding to “purple”

Figure 60: Comparison of component counting functions for “pound” and “purple”.

The persistent dark grey block in the component counting function for the word ”pound” represents five components, while the lighter grey block on the right represents three components. Thus, multiparameter density-based clustering indicates that there are likely either three or five components in this space. In contrast, applying the same method to the word “purple ” seems to show fewer intervals of stability, with only one block corresponding to three persistent components. We can choose any slice and visualize the corresponding prominence vineyard as follows.

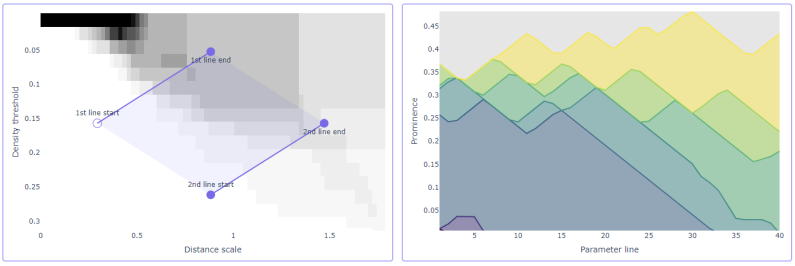


Figure 61: A prominence vineyard of “pound”

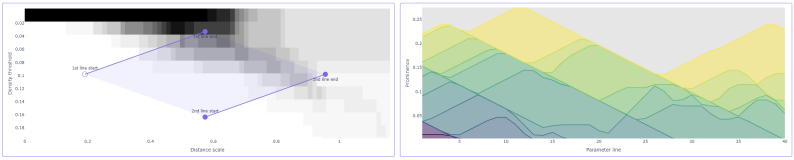


Figure 62: A prominence vineyard of “purple”

Below are examples of possible clusterings generated by Persistable. We visualize them by plotting the first two coordinates of the corresponding 10-dimensional UMAP embeddings.

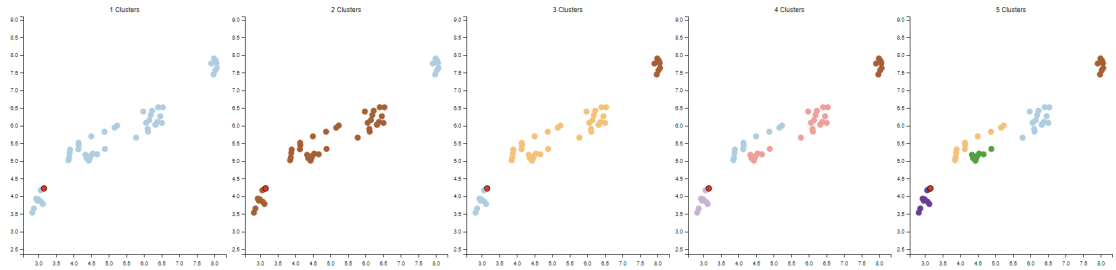


Figure 63: Clusterings of the word “pound”, plotted against the first two dimensions of the 10-dimensional UMAP projection

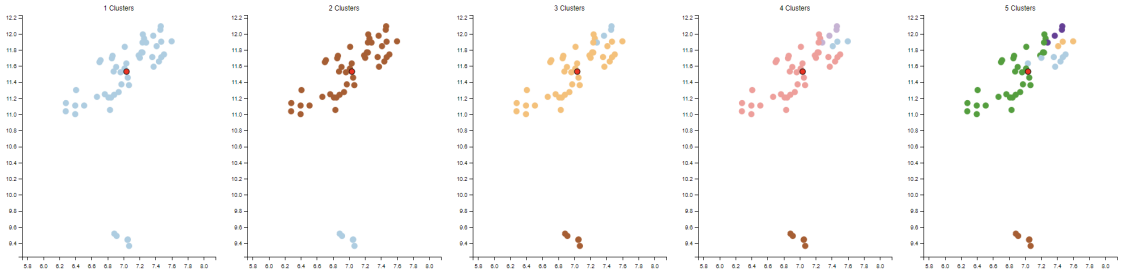


Figure 64: Clusterings of the word "pound", plotted against the first two dimensions of the 10-dimensional UMAP projection

Based on the component counting functions, we select a clustering of 5 components for the word "pound" and 3 for the word "purple". We then construct a one-step filtration that "reallows" intersections with the neighborhood of the word "pound", merge clusters that are within distance $\epsilon = 2$, and compute the corresponding intersection homology groups. Note that the clustering and the distance ϵ at which we merge clusters is arbitrary, and comparing various choices is a direction we leave for future research. We visualize the persistent intersection homology groups associated with "pound" and "purple" below by plotting the first two coordinates of the 10-dimensional UMAP embedding.

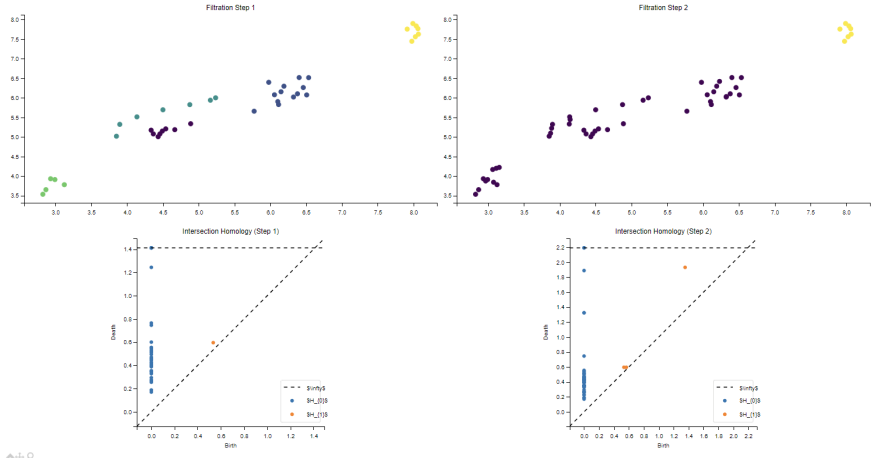


Figure 65: Persistent intersection homology groups of the singularity filtration associated with the word "pound"

We observe that the neighborhoods of both "pound" and "purple" introduce a component and cycle in the corresponding intersection homology groups. However, the persistent homology of the word "pound" seems to have more 0-dimensional components.

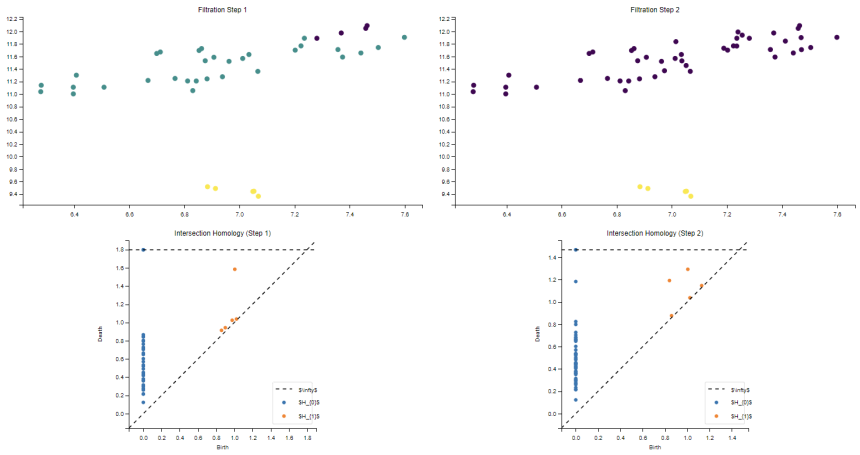


Figure 66: Persistent intersection homology groups of the singularity filtration associated with the word "purple"

These preliminary observations indicate that the topological structures and components within the word embeddings of "pound" and "purple" differ, reflecting the complexity and multiplicity of senses associated with each word. Further exploration and systematic analysis could help clarify these relationships.

5 Conclusion

We have explored the application of persistent homology to the analysis of stratified spaces, particularly in the context of word embeddings. We have seen that by introducing kernel, image, and cokernel persistence, it is possible to analyze the singularities within stratified data more effectively. The use of bifiltrations and sliding window (co)kernel diagrams has provided novel insights into the topology of the data, offering a means to differentiate between various types of singularities, such as those associated with polysemous words. We demonstrated the practical applications of these theoretical advancements by conducting computational experiments using word embeddings. Our findings suggest that the topology of local neighborhoods within word embeddings is influenced by factors such as the choice of embedding method and dimensionality reduction techniques.

There are many directions for future research with regard to the theoretical, computational, and application-oriented aspects of this work, more than can reasonably fit in this section. We mention some possible directions below.

5.1 Future Work

Theory

One potential direction is to extend the study of bifiltrations in kernel/image/cokernel persistence to multifiltrations that incorporate perversities and ideas from chromatic complexes. Additionally, further work is needed to investigate the invariants associated with these multifiltrations and their potential applications in machine learning.

Computation

Future research could focus on exploring different ways of ordering singularities when constructing perversity sequences, such as based on dimensionality. Moreover, investigat-

ing invariants of multiparameter persistence, like the component counting function, could lead to deeper insights.

Application

On the application side, these computational techniques have significant potential for studying the training process of word embeddings and tracking how the topology of neighborhoods of words changes. Extending these techniques to study attention mechanisms could further enhance our understanding of word embeddings. Additionally, we propose applying these techniques to other settings with stratified data, such as the parameter spaces of multilayer perceptrons or biological materials.

In conclusion, this dissertation has laid the groundwork for future exploration at the intersection of topological data analysis and stratified spaces. The methods developed here open new avenues for research and applications, promising to yield further insights into the complex structures that underlie word embeddings and other high-dimensional data representations.

References

[1] E. Aamari and C. Berenfeld. A theory of stratification learning. *arXiv preprint arXiv:2405.20066*, 2024.

[2] S.-i. Amari. *Information geometry and its applications*, volume 194. Springer, 2016.

[3] M. Banagl. *Topological invariants of stratified spaces*. Springer Science & Business Media, 2007.

[4] P. Bendich. *Analyzing stratified spaces using persistent versions of intersection and local homology*. Duke University, 2008.

[5] P. Bendich and J. Harer. Persistent intersection homology. *Foundations of Computational Mathematics*, 11(3):305–336, 2011.

[6] P. Bendich, D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and D. Morozov. Inferring local homology from sampled stratified spaces. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 536–546. IEEE, 2007.

[7] P. Bendich, B. Wang, and S. Mukherjee. Local homology transfer and stratification learning. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 1355–1370. SIAM, 2012.

[8] 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*, 14:745–789, 2014.

[9] M. Botnan and W. Crawley-Boevey. Decomposition of persistence modules. *Proceedings of the American Mathematical Society*, 148(11):4581–4596, 2020.

[10] M. B. Botnan and M. Lesnick. An introduction to multiparameter persistence. *arXiv preprint arXiv:2203.14289*, 2022.

[11] M. B. Botnan, S. Oppermann, and S. Oudot. Signed barcodes for multi-parameter persistence via rank decompositions and rank-exact resolutions. *arXiv preprint arXiv:2107.06800*, 2021.

[12] J. Camacho-Collados and M. T. Pilehvar. From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, 63: 743–788, 2018.

[13] G. Carlsson. Topological pattern recognition for point cloud data. *Acta Numerica*, 23:289–368, 2014.

[14] G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. In *Proceedings of the twenty-third annual symposium on Computational geometry*, pages 184–193, 2007.

[15] F. Chazal, L. J. Guibas, S. Y. Oudot, and P. Skraba. Persistence-based clustering in riemannian manifolds. *Journal of the ACM (JACM)*, 60(6):1–38, 2013.

[16] G. Chen and G. Lerman. Spectral curvature clustering (scc). *International Journal of Computer Vision*, 81:317–330, 2009.

[17] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 263–271, 2005.

[18] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the twenty-second annual symposium on Computational geometry*, pages 119–126, 2006.

[19] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and D. Morozov. Persistent homology for kernels, images, and cokernels. In *Proceedings of the twentieth annual ACM-SIAM symposium on Discrete algorithms*, pages 1011–1020. SIAM, 2009.

[20] S. C. di Montesano, O. Draganov, H. Edelsbrunner, and M. Saghafian. Chromatic alpha complexes, 2024.

[21] S. C. di Montesano, O. Draganov, H. Edelsbrunner, and M. Saghafian. Chromatic topological data analysis. *arXiv preprint arXiv:2406.04102*, 2024.

[22] H. Edelsbrunner and J. L. Harer. *Computational topology: an introduction*. American Mathematical Society, 2022.

[23] H. Edelsbrunner, J. Harer, et al. Persistent homology-a survey. *Contemporary mathematics*, 453(26):257–282, 2008.

[24] J. Eisenstein. *Introduction to natural language processing*. MIT press, 2019.

[25] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.

[26] C. Fefferman, S. Mitter, and H. Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049, 2016.

[27] G. Friedman. Stratified fibrations and the intersection homology of the regular neighborhoods of bottom strata. *Topology and its Applications*, 134(2):69–109, 2003.

[28] G. Friedman. An introduction to intersection homology with general perversity functions. In *Proceedings of the Workshop on the Topology of Stratified Spaces at MSRI*, 2008.

[29] G. Friedman. *Singular intersection homology*, volume 33. Cambridge University Press, 2020.

[30] R. W. Ghrist. *Elementary applied topology*, volume 1. Createspace Seattle, 2014.

[31] M. Goresky and R. MacPherson. Intersection homology theory. *Topology*, 19(2): 135–162, 1980.

[32] M. Goresky and R. MacPherson. Intersection homology 11. *Inc. Mat*, 71:77–129, 1983.

[33] M. Goresky, R. MacPherson, M. Goresky, and R. MacPherson. *Stratified morse theory*. Springer, 1988.

[34] N. Habegger and L. Saper. Intersection cohomology of cs-spaces and zeeman's filtration. *Inventiones mathematicae*, 105(1):247–272, 1991.

[35] A. Hatcher. *Algebraic topology*. 2005.

[36] M. Huber, S. Kalisnik, and P. Schnider. Automato: A parameter-free persistence-based clustering algorithm. *arXiv preprint arXiv:2408.06958*, 2024.

[37] B. Hughes and S. Weinberger. Surgery and stratified spaces. *Surveys on surgery theory*, 2:319–352, 2000.

[38] A. Jakubowski, M. Gašić, and M. Zibrowius. Topology of word embeddings: Singularities reflect polysemy. *arXiv preprint arXiv:2011.09413*, 2020.

[39] I. Jones. Diffusion geometry. *arXiv preprint arXiv:2405.10858*, 2024.

[40] H. C. King. Topological invariance of intersection homology without sheaves. *Topology and its Applications*, 20(2):149–160, 1985.

[41] F. Kirwan and J. Woolf. *An introduction to intersection homology theory*. CRC Press, 2006.

[42] S. L. Kleiman. The development of intersection homology theory. *Pure Appl. Math. Q*, 3(1):225–282, 2007.

[43] D. Lapous, H. Schreiber, L. Scoccola, and M. Carrière. multipers: Multiparameter persistence for machine learning. URL `https://github.com/DavidLapous/multipers`.

[44] S. Lefschetz. *Topology*, volume 12. American Mathematical Soc., 1930.

[45] M. Lesnick and M. Wright. Interactive visualization of 2-d persistence modules. *arXiv preprint arXiv:1512.00180*, 2015.

[46] U. Lim, H. Oberhauser, and V. Nanda. Hades: Fast singularity detection with local measure comparison. *arXiv preprint arXiv:2311.04171*, 2023.

[47] D. Loiseaux, M. Carriere, and A. J. Blumberg. Fast, stable and efficient approximation of multi-parameter persistence modules with mma. *arXiv preprint arXiv:2206.02026*, 2022.

[48] R. MacPherson. Intersection homology and perverse sheaves. In *Unpublished AMS Colloquium Lectures, San Francisco*, volume 4, page 15, 1991.

[49] R. MacPherson and K. Vilonen. Elementary construction of perverse sheaves. *Inventiones mathematicae*, 84(2):403–435, 1986.

[50] L. Maxim. *Intersection Homology & Perverse Sheaves*. Springer, 2019.

[51] C. McCrory. Cone complexes and pl transversality. *Transactions of the American Mathematical Society*, 207:269–291, 1975.

[52] L. McInnes, J. Healy, and J. Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

[53] T. Mikolov. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[54] G. A. Miller. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[55] V. Nanda. Local cohomology and stratification. *Foundations of Computational Mathematics*, 20:195–222, 2020.

[56] V. Nanda. *Computational Algebraic Topology: Lecture Notes*. 2023.

[57] S. Y. Oudot. *Persistence theory: from quiver representations to data analysis*, volume 209. American Mathematical Soc., 2017.

[58] R. Patil, S. Boit, V. Gudivada, and J. Nandigam. A survey of text representation and embedding techniques in nlp. *IEEE Access*, 11:36120–36146, 2023.

[59] J. Pennington, R. Socher, and C. D. Manning. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014. URL `http://www.aclweb.org/anthology/D14-1162`.

[60] H. Poincaré. *Analysis situs*. Gauthier-Villars Paris, France, 1895.

[61] H. Poincaré. Complémenta l'analysis situs. *Rendiconti del Circolo matematico di Palermo*, 13(285-343):10–1007, 1899.

[62] L. Polterovich, D. Rosen, K. Samvelyan, and J. Zhang. *Topological persistence in geometry and analysis*, volume 74. American Mathematical Soc., 2020.

[63] B. Rieck, M. Banagl, F. Sadlo, and H. Leitte. Persistent intersection homology for the analysis of discrete data. In *Topological Methods in Data Analysis and Visualization V: Theory, Algorithms, and Applications 7*, pages 37–51. Springer, 2020.

[64] A. Rolle and L. Scoccola. Stable and consistent density-based clustering via multiparameter persistence. *arXiv preprint arXiv:2005.09048*, 2020.

[65] L. Scoccola and A. Rolle. Persistable: persistent and stable clustering. *Journal of Open Source Software*, 8(83):5022, 2023. doi: 10.21105/joss.05022. URL `https://doi.org/10.21105/joss.05022`.

[66] B. J. Stolz, J. Tanner, H. A. Harrington, and V. Nanda. Geometric anomaly detection in data. *Proceedings of the national academy of sciences*, 117(33):19664–19669, 2020.

[67] T. Temčinas. Local homology of word embeddings. *arXiv preprint arXiv:1810.10136*, 2018.

[68] R. Thom. Ensembles et morphismes stratifiés. 1969.

[69] A. L. Thomas. *Invariants and metrics for multiparameter persistent homology*. PhD thesis, Duke University, 2019.

[70] B. S. Thomas, L. Lin, L.-H. Lim, and S. Mukherjee. Learning subspaces of different dimension. *arXiv preprint arXiv:1404.6841*, 2014.

[71] R. Tinarrage. Recovering the homology of immersed manifolds. *Discrete & Computational Geometry*, 69(3):659–744, 2023.

[72] J. Von Rohrscheidt and B. Rieck. Topological singularity detection at multiple scales. In *International Conference on Machine Learning*, pages 35175–35197. PMLR, 2023.

[73] S. Weinberger. *The topological classification of stratified spaces*. University of Chicago Press, 1994.

[74] H. Whitney. Tangents to an analytic variety. In *Hassler Whitney Collected Papers*, pages 537–590. Springer, 1964.

[75] H. Whitney. Local properties of analytic varieties. In *Hassler Whitney Collected Papers*, pages 497–536. Springer, 1965.

[76] A. Zomorodian and G. Carlsson. Computing persistent homology. In *Proceedings of the twentieth annual symposium on Computational geometry*, pages 347–356, 2004.

[77] R. Řehůřek, P. Sojka, et al. Gensim—statistical semantics in python. *Retrieved from genism. org*, 2011.