Persistent Intersection Homology*

Paul Bendich† and John Harer‡

November 5, 2009

Abstract

The theory of intersection homology was developed to study the singularities of a topologically stratified space. This paper incorporates this theory into the already developed framework of persistent homology. We demonstrate that persistent intersection homology gives useful information about the relationship between an embedded stratified space and its singularities. We give, and prove the correctness of, an algorithm for the computation of the persistent intersection homology groups of a filtered simplicial complex equipped with a stratification by subcomplexes. We also derive, from Poincaré Duality, some structural results about persistent intersection homology.

AMS Subject Classifications $55N33, 68$.

1 Introduction

Over the last several years, computational topology has grown into a flourishing area. A cornerstone of the subject is persistent homology, introduced in [16]. Persistent homology is a method for measuring features of a filtered topological space. The theory has expanded to include a number of topological tools including Poincaré and Lefschetz duality [11], Mayer-Vietoris sequences [10] and Morse Theory [15].

Persistent Intersection Homology. In [18], Goresky and MacPherson developed intersection homology theory as a tool for the study of stratified spaces. Briefly, these are spaces obtained by piecing together manifolds of different dimensions in a controlled and coherent fashion. The goal of this paper is the incorporation of intersection homology into the persistence framework, with several types of potential applications which we describe below.

After defining persistent intersection homology we give an algorithm for its computation in a simplicial setting. While the full definition and topological meaning of intersection homology is rather involved, its algebraic description is quite simple. One defines intersection homology chains by starting with an ordinary simplicial chain complex and removing certain simplices from consideration to obtain a new chain complex, and then one simply computes the homology of this chain complex.

Here we will abstract this algebraic process into a definition of what we call $\phi-$homology, where any arbitrarily chosen binary function ϕ represents the simplex removal decision procedure. We then define $\phi-$persistence and give an algorithm for its computation. Persistent intersection homology will then be described as one case of ϕ-persistence, but one in which the function ϕ has an obvious topological meaning based on the singularity structure of the space.

A key fact about intersection homology, proven in [18], is that it provides Poincaré Duality for stratified spaces; using ordinary homology theory, duality fails. We will use Poincaré Duality to prove a series of symmetry results for persistent intersection homology on a embedded stratified space which is filtered by the sublevel sets of a function. These symmetries closely mirror those held by the usual persistent homology on an embedded and filtered *manifold* ([11]).

In [1] manifold symmetries were used to define an elevation function on an embedded 2-manifold. The critical points of this function help to locate important features,

such as "pockets" and "protrusions" on the manifold. The theory has subsequently been used to study protein docking in [27]. In an upcoming paper [6], we follow a similar pattern: the symmetry properties for persistent intersection homology are used to define a series of elevation functions which will be more sensitive to the singularities of an embedded 2-dimensional stratified space. The most immediate application envisioned is in the study of the protein-protein interface surface ([2]).

We also expect that persistent intersection homology will play a role in a more ambitious program: the analysis of high-dimensional datasets (point clouds). In particular, in certain cases point clouds appear to be samples of stratified spaces. Since the study of datasets is plagued by the *curse of dimensionality*, methods have been developed to reduce the dimension of the dataset. We can think of this as searching for dependency among variables, thereby recognizing that the dataset is really a noisy version of a smaller dimensional manifold. This process is called *manifold learning* (for some techniques, see, for example [3] and [26]). Implicit in most of this work is the assumption that the dataset lies on a manifold, often a hyperplane, and therefore that the dimensionality of the dependence is constant. Of course it is well known in the manifold learning community that this is not true in general, but so far there are no formal methods for discovering the structure of the dimension changes that may occur.

Local homology groups are a traditional algebraic tool for analyzing stratified spaces; briefly, they provide a homological description of the neighborhood of a (possibly) singular point on the space. A persistent version of these groups was defined in [5] in order to infer the possible stratifications of the underlying spaces from which the dataset might be sampled. Persistent intersection homology aides this analysis since intersection homology groups are designed to capture structures of interest in singular spaces.

Outline. In Sec. 3, we recall the definition of persistent homology and briefly review the algorithm for its computation. We also give an example which shows its limitations when applied to an embedded stratified space. Sec. 4 defines ϕ-persistence and gives an algorithm for its computation, although a proof is deferred until the Appendix.

The paper then switches gears into a discussion of in-

tersection homology. Sec. 5 discusses the category of topologically stratified spaces. The intersection homology groups of a stratified space, as well as the concept of intersection homology persistence, are then described in Sec. 6. As these concepts are rather involved on a first reading, we work an extended example in Sec. 7. This example illustrates the Duality and Symmetry properties enjoyed by intersection homology persistence. We formally describe these properties in Sec. 8. Finally, Sec. 9 describes how to filter a simplicial complex, via subdivided stars, in such a way as to effectively mimic the intersection homology persistence on an actual stratified space.

2 Results and Relation to Prior Work

2.1 Persistence.

The history of persistence is surveyed in [14]. Versions of persistence emerged independently in [17], [7], [25] and [16]. Later work on the theory and some of its applications can be found in e.g. [12], [11], [9], [10] and [8],.

2.2 Intersection Homology.

The intersection homology groups of a triangulated topologically stratified space X were first defined by Goresky and MacPherson in [18]. The same authors gave a different definition using sheaf theory in [19], while King put the theory into a singular chain context in [22]. Our version of intersection homology differs slightly from those above, for reasons that we now briefly discuss.

The authors in [18] impose the condition that their stratified spaces have no strata of codimension one. Since their primary application was to complex algebraic varieties, this was not a troublesome requirement. They also imposed a condition on perversities (explained below) to ensure that their definition was independent of the choice of stratification. We, on the other hand, envision applying this theory to all types of stratified spaces, and hence wish to make no assumptions on stratum-codimension. Dropping this assumption has two consequences, however:

1. Poincaré Duality no longer holds.

2. Intersection homology groups are no longer independent of choice of stratification.

The first problem is easily dealt with by a slight alteration of the original definition. Instead of working with ordinary chains, we use the relative chain groups $C_*(X, \Sigma)$, where Σ is the singular set of X. As proven in [4], this restores Poincaré Duality. The second problem, on the other hand, is not a problem at all but a feature that we desire since our goal is to apply the theory developed here to point cloud data where the stratification is in fact the very structure that we seek.

2.3 Results.

The main results of this paper are the following.

- The definition of ϕ-persistence (with persistent intersection homology as its main example), and an algorithm for its computation.

- Duality and Diagram Symmetry results for intersection homology.

- The Subdivided Star Filtration.

3 Persistence

We begin this section by reviewing the definitions of ordinary and extended persistent homology for a topological space equipped with two filtrations by closed subspaces. We then present two examples: a one-dimensional manifold and a two-dimensional stratified space, each filtered by the sublevel and superlevel sets of a height function. The latter example motivates our extension of the theory to persistent intersection homology. Along the way, we define the persistence diagram of a function and its subdiagrams. The section concludes with a brief description of the persistence algorithm ([16]).

3.1 Definition of Persistence.

Suppose that X is a topological space and that

$$\varnothing = A_0 \subset A_1 \subset \ldots \subset A_{n-1} \subset A_n = X$$

is a filtration of X by closed subspaces. Suppose further that $h : \{0, \ldots, n\} \to \mathbb{R}$ is a monotone increasing function that assigns a height to each level. Fix a dimension r and let $H_r^i = H_r(A_i; \mathbb{Z}/2\mathbb{Z})$, $0 \leq i \leq n$. When $i < j$, the inclusion $A_i \hookrightarrow A_j$ induces $f_r^{i,j} : H_r^i \to H_r^j$, whose image consists of all homology classes that live at least from H_r^i to H_r^j.

For a particular class α we define its (ordinary) *persistence* to be the difference between the height values of its birth and death. Specifically, α is *born* at A_i if $\alpha \in H_r^i - \text{im} f_r^{i-1,i}$ and *dies* at A_j if $f_r^{i,j-1}(\alpha) \notin \text{im} f_r^{i-1,j-1}$ but $f_r^{i,j}(\alpha) \in \text{im} f_r^{i-1,j}$. The *persistence* of α is then $|h(j) - h(i)|$.

The classes that are born at stage i and die at stage j form the *pair group*;

$$\mathsf{P}_r^{i,j} \quad = \quad \frac{\text{im} f_r^{i,j-1} \cap \ker f_r^{j-1,j}}{\text{im} f_r^{i-1,j-1} \cap \ker f_r^{j-1,j}}. \tag{1}$$

Extended Persistence A class that is born at some level and never dies is called an *essential class*. Essential classes represent the actual homology of the space and are not paired by ordinary persistence. But there are geometric reasons that we might wish to assign them a persistence, and we can do this when we have a second filtration of X by closed subspaces:

$$\varnothing = D_0 \subset D_1 \subset \ldots \subset D_n = X$$

We ascend using A and descend using D. Specifically, put $H_r^{n+j} = H_r(X, D_j)$ for $0 \leq j \leq n$ and use the following sequence to extend the notions of birth, death, persistence, and pair group.

$$0 = H_r^0 \to H_r^1 \to \ldots \to H_r^n \to$$
$$H_r^{n+1} \to H_r^{n+2} \ldots \to H_r^{2n} = 0$$

Since the sequence is bracketed by zero groups, all classes born must eventually die. A class whose birth and death occurs during the top half of the sequence is called an *ordinary class*, while a class whose birth and death occurs during the bottom half is called a *relative class*, Finally, a class whose birth and death straddles both halves of the sequence is an *extended class*.
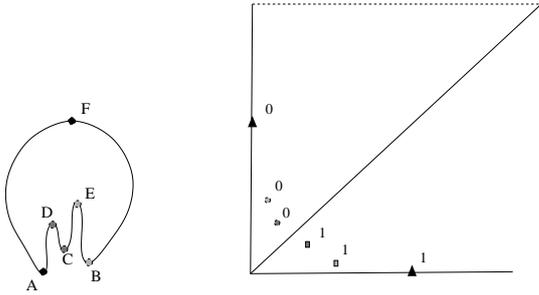
Figure 1: A manifold with a height function, along with the extended persistence diagram. In the diagram, points are labelled by dimension. Circles, squares, and triangles represent ordinary, relative, and extended pairs, respectively.

3.2 Height Function Examples.

To illustrate the definitions above, as well as to provide partial justification for the development of intersection homology persistence, we briefly present two examples. Along the way, we will see a type of duality for pair groups which holds as long as X is a manifold, but can fail otherwise.

One-Manifold. Consider the space X, topologically a circle, but embedded in the plane as shown in Fig. 1. Let $f : X \to \mathbb{R}$ measure height in the vertical direction. We filter X by the sublevel sets of f, $X_a = f^{-1}((-\infty, a])$, and the superlevel sets $X^a = f^{-1}([a, \infty))$.

As one may easily see, the topology of these sets will only change when we pass one of the critical points labeled in Fig. 1, a general fact from Morse Theory ([24]). Interleaving regular values $a_0 \leq f(A) \leq a_1 \leq f(B) \leq \ldots \leq f(F) \leq a_6$, we define $A_i = X_{a_i}$ and $D_j = X^{a_6-j}$. Moving up along the ascending filtration, a component is born at each of the points A, B and C. The component born at C dies as we pass point D, while the one born at B dies at E. Finally, at point F, an essential one-cycle is born. Descending, the essential component born at A become trivial passing F, since $H_0(X, F)$ is trivial. Relative one-cycles born at E and D die at B and C, respectively. Finally the essential one-cycle becomes trivial at A.

In terms of Equation 1, the nonzero pair groups are $P_0^{1,7}, P_0^{2,5}, P_0^{3,4}, P_1^{6,12}, P_1^{8,11}$, and $P_1^{9,10}$. Each has rank

one.

Persistence Diagrams and Subdiagrams Both ordinary and extended persistence pairs can be encoded compactly in a persistence diagram. For each nonzero basis element of a pair group, we locate the critical point that created the class and the critical point that destroyed it. These two points are paired together, and we plot their height values as x, y coordinates in the plane. For example, the ordinary pair represented by $P_0^{2,5}$ is plotted as the point $(f(B), f(E))$, while the extended pair that measures the essential cycle gives the point $(f(F), f(A))$. The diagram for this example is shown on the right in Fig. 1.

In general, the persistence diagram associated to the filtrations given by a height function f is denoted $Dgm(f)$, with $Dgm_r(f)$ being the restriction to dimension r. Within each persistence diagram are overlaid several important sub-diagrams which we label with the symbols $Ord_r(f), Rel_r(f), Ext_r(f)$ to stand for, respectively, ordinary, relative, and extended pairs in dimension r.

Pair Group Duality The reader will note the obvious symmetry in the persistence diagram above. This holds because X is a manifold and is a consequence of Poincaré Duality. If we assume that our function f has n critical points and that X is a d-manifold, then duality can be expressed in the following way ([11]),:

1 (Pair Group Duality) *For $0 \leq i < j \leq 2n$, $0 \leq r \leq d$, intersection of homology classes induces a perfect pairing*

$$P_r^{i,j} \otimes P_{d-r}^{2n-j+1,2n-i+1} \to \mathbb{Z}/2\mathbb{Z}$$

Stratified Spaces. Next consider the space X shown in Fig. 2. Topologically, X consists of a pinched torus with a disc attached. This disc, not pictured, is attached along the dotted circle drawn passing through points B, D and F. X is an example of a 2-dimensional stratified space, an object described in greater detail in Sec. 5. Imagine that the function f measures height in the vertical direction and that we filter X by sublevel and superlevel sets as before.
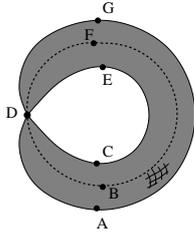
Figure 2: A disc (not pictured) is attached along the drawn curve.

Although a rigorous argument requires Stratified Morse Theory ([20]), it should be clear that the only possible homological changes occur upon passing the points labeled with letters in Fig. 2. Actually, nothing whatsoever happens at the point B since the sublevel set just after $f(B)$ deformation retracts onto the sublevel set just before it. Similarly, nothing happens upon passing the point F. This means that persistent homology does not measure the height difference between the minimum of the torus and the minimum of the attached disc which is an important piece of information about this space and its singularities. Intersection homology persistence, on the other hand, will capture this information, as it is more finely tuned to shape changes caused by the presence of singularities.

Note also that Pair Group Duality fails for this example. This is not surprising, since Poincaré Duality does not hold for non-manifolds. On the other hand, intersection homology restores a version of Poincaré Duality to stratified spaces. As we explain below, this means that we can recover a version of Pair Group Duality.

3.3 Algorithm.

We now briefly review the algorithm for the computation of the persistent homology pairings on a simplicial complex K equipped with a filtration $\{K^i\}$ and a height function h as above. For the details on the extended persistence algorithm for a complex equipped with two filtrations, we refer the reader to [11].

The first step is to refine the input filtration to one in which one simplex is added at each level. We do this by ordering the simplices arbitrarily within each $K^i - K^{i-1}$

while ensuring that a simplex precedes all of its cofaces in the ordering. The height values coming from the function h are maintained, and so the reordering does not affect the persistence values defined above.

Positive and Negative Simplices. Suppose K has n simplices $\sigma_1, \ldots, \sigma_n$, with $K^i = K^{i-1} \cup \{\sigma_i\}$. Fix an i and assume $\dim(\sigma_i) = r$. Let α be the $r-1$ dimensional homology class represented by $\partial \sigma_i$. The addition of σ_i to the filtration will have one of two possible homological effects:

- If α is nontrivial in H_r^{i-1}, adding σ_i kills it so $\beta_{r-1}(K^i) = \beta_{r-1}(K^{i-1}) - 1$, while all other Betti numbers remain unchanged. We say that σ_i is a **negative** r-simplex.

- If α is already 0 in H_r^{i-1}, there is a chain $\gamma \in C_r(K^{i-1})$ such that $\partial \gamma = \partial \sigma_i$. We see then that the cycle $\gamma + \sigma_i$ represents an r-dimensional class born at the ith level. Therefore $\beta_r(K^i) = \beta_r(K^{i-1}) + 1$ and once again all other Betti numbers are unchanged. We say that σ_i is a **positive** r-simplex.

If the positive σ_i creates a class which is subsequently destroyed by the addition of the negative σ_j, we say informally that σ_j kills σ_i, and pair these two simplices together. If the addition of σ_i creates an essential class, we leave σ_i unpaired.

There is then a one-to-one correspondence between the pairs (σ_i, σ_j) and the nonzero, and thus necessarily rank one, pair groups $P_r^{i,j}$, where $r = \dim(\sigma_i)$. Similarly, there is a one-to-one correspondence between the unpaired r-simplices and the rank of $H_r(K)$. The simplicial persistence algorithm, described below, computes these pairs of simplices and also identifies the unpaired ones.

Fig. 3 illustrates this correspondence for a filtered triangle. The simplices of the triangle are added in increasing numerical order. The addition of edge 4 merges the component formed by vertex 3, so these two simplices are paired. Similarly, we pair vertex 2 and edge 5. Edge 6 and triangle 7 are paired since the 1-cycle created by the edge is immediately filled in by the triangle. Vertex 1, representing the entire component, goes unpaired, although it would be paired later by extended persistence if we had another filtration.
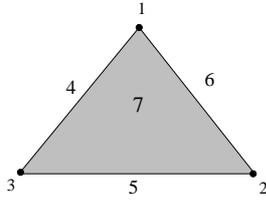
Figure 3: The only nonzero pair groups are $P_0^{2,5}$, $P_0^{3,4}$, and $P_1^{6,7}$.

Figure 4: Filled-in rectangles indicate the non-zero entries.

Boundary Matrix. We form the $n \times n$ binary incidence matrix D by setting $D[i,j] = 1$ iff σ_i is a codimension-one face of σ_j. For example, the matrix corresponding to the filtered triangle in Fig. 3 is shown in Fig. 4, although we omit the vertex-indexed columns, since vertices have no faces.

The algorithm performs column operations to transform D into a matrix of simpler form; the paired and unpaired simplices are then read off the simpler matrix.

Reduced Matrices. Let M be an arbitrary $n \times n$ binary matrix. Define the "lowest-one" function $low_M : \{1, 2, \ldots, n\} \to \{0, 1, 2, \ldots, n\}$ by setting $low_M(j)$ to the index of the lowest nonzero entry in the jth column, if it exists. If the column is all 0s, set $low_M(j) = 0$. A matrix M is said to be *reduced* if low_M is injective on the complement of the preimage of 0.

Reduction Process and Interpretation. The algorithm reduces D by performing column operations left-to-right:

    for j = 1 to n **do**
        while $\exists j' < j$ with $low(j') = low(j) \neq 0$ **do**
            add column j' to column j

Figure 5: The lighter shaded rectangles indicate pairings.

    end while
  end for.

This produces a reduced matrix R, and the paired simplices are then given directly by the associated function low_R:

- if $low_R(k) = 0$, then σ_k is a positive simplex. It will either be paired later or remain unpaired.

- if $low_R(j) = i$, then we pair the positive simplex σ_i with the negative simplex σ_j.

The reduced matrix for the filtered triangle example is shown in Fig. 5.

The column operations each correspond to multiplication by an elementary matrix, and their product produces a matrix V with $R = DV$. The columns of V give additional information:

- If $low_R(k) = 0$, then the entries of the kth column of V give a cycle representing an element in the coset of classes born when σ_k is added.

- If $low_R(j) = i$, then the entries in the jth column of V give a representative for one of the classes in the coset that dies after the addition of σ_j.

4 ϕ - Persistence: Definition and Algorithm

In this section, we introduce the idea of ϕ - persistence for a simplicial complex equipped with an ordering on its simplices, and give an algorithm for its computation. The proof of correctness of the algorithm is contained in the Appendix.

6
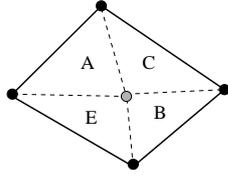
Figure 6: The interior vertex and incident edges are improper. All other simplices are proper.

4.1 ϕ - Homology

Given a simplicial complex K, let $\phi : K \rightarrow \{0, 1\}$ be a function on the simplices of K. If $\phi(\sigma) = 1$, we say that σ is a proper simplex; otherwise, it is improper. Let $P(K)$ be the set of all proper simplices, and $P_i(K)$ the $\mathbb{Z}/2\mathbb{Z}$- vector space with basis the proper i-dimensional simplices.

Allowable Chains. We would like to replace the $C_i(K)$ by the $P_i(K)$ but this does not work directly because there is no guarantee that the boundary of a proper i-simplex will be the sum of proper $(i-1)$ simplices.

As an example, consider Fig. 6. Suppose that all triangles and the outside edges and vertices are proper, but the central vertex and all incident edges are improper. Then triangle $E \in P_2(K)$ but $\partial E \notin P_1(K)$, since the boundary of E contains two improper 1-simplices. On the other hand, the 2-chain $\Delta = A+B+C+E$ is a sum of proper 2-simplices *and* its boundary is also a sum of proper edges; by adding the triangles together, we have cancelled all improper boundary edges.

This picture illustrates the general definition: a chain $\xi \in C_i(K)$ is allowable if both ξ and $\partial \xi$ can be written as sums of proper simplices. In the above example, E would not be allowable, but Δ would be. Note that the set of all allowable i-chains forms a $\mathbb{Z}/2\mathbb{Z}$-vector space $I^\phi C_i(K)$.

Now suppose ξ is an allowable i-chain. Since $\partial\partial\xi = 0$, $\partial\xi$ is itself an allowable $(i-1)$-chain. Therefore, the boundary maps ∂_i give a sequence of well-defined homomorphisms $\partial_i : I^\phi C_i(K) \rightarrow I^\phi C_{i-1}(K)$ with $\partial_i \circ \partial_{i+1} = 0$, so we have a chain complex. Define $I^\phi H_i(K)$ to be the ith homology group of this complex.

ϕ - **Persistence.** Given a filtration $\{K^i\}$ of K, and a binary function ϕ, we restrict ϕ to K^i to define $I^\phi H_r^i = I^\phi H_r(K^i)$. For $i < j$, the inclusions $K^i \hookrightarrow K^j$ induce maps on ϕ-homology. Using these maps, we define birth and death, persistence, and the pair groups $I^\phi P_r^{i,j}$ in exact analogy with their definitions in the case of standard homology persistence (3.1). Given another filtration, we could also define extended persistence in the obvious way.

4.2 Active and Neutral Simplices.

The ϕ-persistence algorithm is quite similar in form to the ordinary persistence algorithm, in that it reduces a boundary matrix until the "lowest-one" function is injective. The main differences are the interpretation of this lowest-one function and the initial ordering of the columns and rows of the boundary matrix. These changes are necessitated by the fact that we can no longer partition the set of simplices into positive and negative, as we did for ordinary homology. Instead, there is a third category, "neutral" , which requires special attention. Before describing the algorithm, we first address this distinction.

Case Analysis. For a chain $\gamma \in P_i(K)$, let $I(\gamma)$ be the set of improper simplices in its boundary. Adding an improper simplex σ to a simplicial complex has no effect on its ϕ−homology as σ can not form part of an allowable chain. Adding a proper σ, however, can change ϕ−homology, although it need not. Suppose that the complex so far is called L and that we add the i-dimensional σ to L. Then one of three things occurs:

- There exists $\gamma \in P_i(L)$ such that $I(\gamma) = I(\sigma)$. In this case, the sum $\alpha = \gamma + \sigma$ is an *allowable* chain since the addition cancels out all improper simplices along the boundaries of γ and σ. One of two things then happens:

  1. $\partial\alpha$ was not the boundary of an allowable i-chain in L, but is in $L \cup \{\sigma\}$. This means the addition of σ lowered the $(i-1)$st ϕ−Betti number by one. In this case, we call σ **negative**.

  2. $\partial\alpha$ was already the boundary of an allowable i-chain β in L. Then $\alpha + \beta$ represents a new nonbounding allowable i-cycle. The ith ϕ−Betti

7

number increases by one and we say that σ is **positive**.

- For each $\gamma \in P_i(L)$, we have $I(\gamma) \neq I(\sigma)$. In this case, the addition of σ cannot create any new allowable chains. All $\phi-$Betti numbers remain the same and we think of σ as **neutral**. Note that σ may later aid in the creation of an allowable chain.

Sometimes we will wish to stress only that a particular simplex is not neutral, without specifying whether it is positive or negative. In this case, we call the simplex **active**.

Example. Referring again to Fig. 6, suppose that we filter this simplicial complex by first adding all vertices and edges, and then the triangles with E coming last in the ordering. Then every triangle other than E will be neutral. On the other hand, the set $I(E)$ consists of the pair of its boundary edges which are incident on the central vertex. These two edges also form $I(\gamma)$, where γ is the sum of all the other previously-added triangles. The 2-chain $E + \gamma$ is then allowable, as its boundary, consisting of all the external edges, is a sum of proper 1-simplices. Hence E is active. Since this boundary was previously a non-bounding 1-cycle, E is in fact negative.

4.3 ϕ-Persistence Algorithm.

Values. The input to the algorithm is the ordered set $\tau_1, \cdots \tau_m$ of simplices of K and the function $\phi : K \rightarrow \{0, 1\}$.

Re-Ordering. Recall that $P(K), I(K)$ are the subsets of proper and improper simplices; assume they are of size $s, m - s$, respectively. We reorder the input simplices so that those in $P(K)$ come first, while otherwise preserving the input ordering. The proper simplices are then renamed $\sigma_1, \sigma_2, \ldots, \sigma_s$, and the improper simplices are $\sigma_{s+1}, \sigma_{s+2}, \ldots, \sigma_m$. We will later need to refer back to the original ordering when formulating the proof of correctness for our algorithm. To make this easier, we define an order-preserving bijection g on $\{1, 2, \ldots, m\}$ by $\sigma_i = \tau_{g(i)}$.
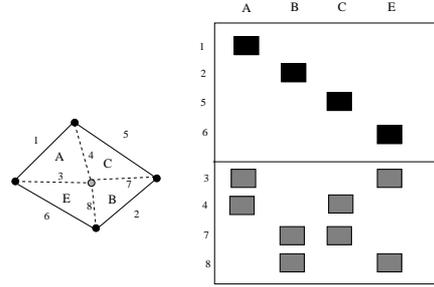
Figure 7: The simplices are added in increasing numerical and then increasing alphabetical order. The columns are indexed by proper triangles, while the rows are indexed first by proper edges, then by improper edges. The horizontal line divides proper from improper edges.

Reduction Algorithm. The $m \times s$ binary matrix D (Fig. 7) is constructed as follows. The s proper simplices index the columns, while the rows are indexed first by the s proper simplices and then by the $m - s$ improper ones. We define $D[i, j]$ to be 1 iff σ_i is a codimension one face of σ_j. We then define the "lowest-one" function and the concept of a *reduced* matrix exactly as in SubSec. 3.3; we also use an identical reduction procedure.

The part of the D matrix corresponding to the triangle-indexed columns for the complex in Fig. 6 is shown in Fig. 7.

Interpretation and Pairings The algorithm above produces a reduced $m \times s$ matrix R. We read the pairings from low_R as follows:

- if $low_R(j) = 0$, then σ_j is active and positive.

- if $low_R(j) = i \leq s$, then σ_j is active, negative, and paired with σ_i.

- if $low_R(j) = k > s$, then σ_j is neutral.

Intuition. Consider the original simplex ordering and imagine adding one simplex at a time in sequence. An *allowable* chain, whether it is a cycle or not, must necessarily be a sum of *proper* simplices. Hence the addition of an *improper* simplex can neither create an allowable cycle nor destroy one via an allowable chain. For this reason,

8
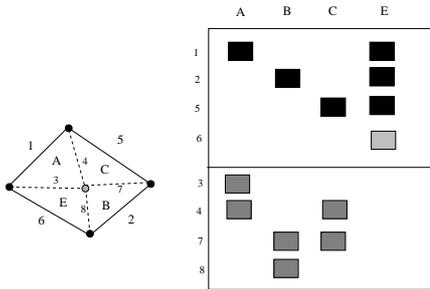
Figure 8: A lowest rectangle must be *above* the horizontal line in order to give a pairing.

we imagine that we are adding only the proper simplices in sequence and we index the columns accordingly. However, the boundary of a proper simplex σ need not itself consist of a sum of proper simplices. In this case, the simplex is not, by itself, an allowable chain. This does not mean that σ is neutral, since we might hope to add older proper simplices to σ in an attempt to cancel off the improper simplices in its boundary. We include the improper simplices at the bottom of the row listing and separate them from the proper rows by a horizontal line. The cancellation of the improper simplices along the boundary of σ raises the lowest one in the σ−indexed column; σ will create an allowable chain if and only if this lowest one ends up above the horizontal line.

For example, consider Fig. 8, which shows the reduced matrix from the input matrix in Fig. 7. The only lowest-one above the horizontal line is in row 6 of column E. This illustrates the fact that triangle E created an allowable 2-chain whose boundary is the 1-cycle created by edge 6.

5   Stratified Spaces

As stated earlier, the main motivating example of ϕ-persistence is persistent intersection homology. In this section, we give a description of topologically stratified spaces, the spaces on which intersection homology theory is most naturally defined. These objects come with many different definitions (for a survey, see [21]). We give the one that works most naturally with intersection homology below.

Intuitively, a stratified space is a topological space decomposed into manifold pieces of possibly different dimensions, which "fit together nicely." More precisely,

Definition. A *d-dimensional topologically stratified space* is a topological space $X \subset \mathbb{R}^n$ together with a descending chain of closed subsets:

$$X = X_d \supseteq X_{d-1} \supseteq X_{d-2} \supseteq$$
$$\ldots X_1 \supseteq X_0 \supseteq X_{-1} = \varnothing$$

so that $X_d - X_{d-1}$ is dense in X and so that the following condition is satisfied:
For each $x \in X_i - X_{i-1}$ there is a stratified space V_x

$$V_x = V_d \supseteq \ldots \supseteq V_i = \{point\}$$

where $V_k - V_{k-1}$ has dimension $i - k$, and a map

$$\psi_x : B^i \times V_d \to X$$

such that $B^i \times V_k$ maps PL-homeomorphically onto a closed neighborhood of x in X_k, for all $k \geq i$. Here B^i is a closed i-dimensional ball.

A few remarks on this definition may help to clarify.

1. By taking $k = i$ in the above condition, we see that $X_i - X_{i-1}$ must be a (possibly empty, possibly disconnected) i-manifold. We denote this subspace S_i and call it the ith stratum of X. The connected components of the strata are called pieces. The union of lower strata X_{d-1} is also called Σ, the "singular set" of X.

2. The existence of ψ in the above definition is often referred to as "local normal triviality"; indeed, the space V_x in the above condition may be thought of as a "normal slice" at $x \in S_i$. To make this more precise, let N be a subspace of X which is transverse to each stratum and intersects S_i in the single point x, and let B_δ be a small ball in X centered at x. Then V_x will be homeomorphic to $N \cap B_\delta$, which we denote by N_x. One can show([20]) that the homeomorphism type of the normal slice N_x depends neither on choice of δ nor of N, nor indeed on the choice of x within a particular piece of S_i. Hence the pieces, themselves manifolds, fit uniformly into the larger space. An example of this construction is shown in Fig. 9.
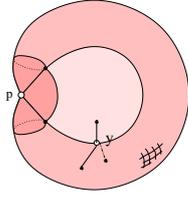
Figure 9: A stratified space with more than one singular stratum: the normal slices at $p \in S_0$ and at $y \in S_1$ are highlighted.

By a *stratified simplicial complex*, we will mean a simplicial complex K which triangulates X so that all the X_i are subcomplexes. Finally, by a *stratified subspace Y* of X, we will mean a closed subspace $Y \subseteq X$ which is itself a stratified space under the stratification inherited from that of X. This means that the ith strata of Y is $S_i \cap Y$. One way to ensure this is to demand that Y intersect each S_i transversely. For example, the normal slice at a point is a stratified subspace. When Y is also a subcomplex, we call it a *stratified subcomplex*.

Example: Pinched Torus with Disc As a first example, let Y be the example drawn in Fig. 9. It is a torus which has had one of its boundary circles pinched to a point (which we'll call p) with a disc stretched across the hole. Let us call the boundary circle of the disc C.

If we remove C from Y, we obtain the disconnected 2-manifold S_2. Note that C itself is a one-manifold. However, not all points on C are singularities of the same kind. If $y \in C, y \neq p$, then y has a neighborhood homeomorphic to three sheets glued together along a line; in terms of the definition, this neighborhood is the product of a 1-ball in C and a cone on three points, one from the disc and two from the torus. On the other hand, p has no such neighborhood; in fact all of its sufficiently small neighborhoods consist of a cone on two circles (in the torus) joined by a line (in the disc). Hence the "local normal triviality" condition demands that we place p in its own individual stratum, leading to the following stratification of X:

$$Y = Y_2 \supseteq Y_1 = C \supseteq Y_0 = \{p\}.$$

Example: Suspended Torus. We include this next example so that we may later use it to illustrate the definition of intersection homology. Let ΣT denote the suspended torus, defined to be the result of collapsing each end of the product $T \times [-1, 1]$ to a point. This space does not embed in $\mathbb{R}^3$, so we picture it in $\mathbb{R}^4$ as the union of two cones. The middle section $T \times 0$ is the usual embedding of the torus in $\mathbb{R}^3$. The cone points a and b are the points $(0, 0, 0, \pm 1)$. The cones are then the collection of straight line segments in $\mathbb{R}^4$ from the torus to the cone points. ΣT is a three-dimensional stratified space, with stratification:

$$\Sigma T = X^3 \supseteq X^2 = X^1 = X^0 = \{a, b\}.$$

We now compute the homology of ΣT. Since ΣT is connected, we have $\beta_0 = 1$. Now T itself had two non-bounding one-cycles, represented by the two boundary circles C_1, C_2. Within ΣT, these cycles become boundaries: for example, C_1 bounds the cone $C_1 * a$. Thus, $\beta_1 = 0$. On the other hand, we also see that $C_1 = \partial(C_1 * b)$. Thus, we obtain a 2-cycle, represented by $C_1 * a + C_1 * b$, which we will denote by ΣC_1. Similarly, ΣC_2 is a 2-cycle, and we find $\beta_2 = 2$. Finally, $\beta_3 = 1$, a three-cycle formed by suspending the fundamental 2-cycle of T.

Note that Poincaré Duality fails here, as the Betti numbers in complementary dimensions are not even equal.

6 Intersection Homology

In this section, we first give a definition of the intersection homology groups for a stratified simplicial complex. We then compute these groups for the two stratified spaces above: the suspended torus and the pinched torus with attached disc. As promised, intersection homology will be an example of ϕ−homology; the key lies in a topologically meaningful definition of ϕ. We then give a quick discussion of the effect that choice of stratification and/or triangulation can have on the intersection homology of a space. The section concludes with the definition of intersection homology persistence.

6.1 Definition

Perversities. A *perversity* is a sequence of integers $\bar{p} = (p_1, p_2, \ldots p_d)$. We impose no restrictions on these integers, although later it will become apparent that the re-

striction $-1 \le p_k \le k - 1$ will still in fact lead to all possible intersection homology groups. We use these perversities to provide a measure of how much intersection between simplices and lower-dimensional strata we will accept.

The top perversity is $\bar{t} = (-1, 0, 1, \ldots, d - 2)$. Two perversities $\bar{p}, \bar{q}$ are called *dual* if $\bar{p} + \bar{q} = \bar{t}$.

Proper Simplices and Intersection Homology Groups.
Given a stratified space X, we choose a triangulation K to get a stratified simplicial complex. An i-simplex σ in K is said to be $\bar{p}$-proper if the following condition holds for all $k = 0, \ldots, d$:

$$\dim(\bar{\sigma} \cap X_{d-k}) \le i - k + p_k$$

where $\bar{\sigma}$ denotes the closure of the open simplex σ. Here we are intentionally confusing σ and $\bar{\sigma}$ with their underlying topological spaces. The intuition behind this inequality is as follows: if an i-dimensional subspace intersects a codim-k subspace *transversely*, the dimension of the intersection will be $i - k$. A non-transverse intersection will result in a higher dimension. Thus, if $p_k = 0$, we are requiring that for σ to be proper, $\bar{\sigma}$ must intersect the codim-k stratum transversely. Higher values of p_k give more tolerant intersection conditions.

As we are permitting codimension-one strata, we will need to work within the relative chain group $C_i(K, \Sigma) = C_i(K)/C_i(\Sigma)$. Thus, an i-chain ξ will be a sum of i-simplices which do not lie entirely within Σ; furthermore, the boundary $\partial \xi$ of this i-chain will be the sum of those $(i-1)$-simplices in the boundary of ξ which also do not lie entirely within Σ. This distinction will be illustrated below, when we compute the intersection homology of the pinched torus with a disc attached.

The ith intersection homology group with perversity $\bar{p}$, $I^p H_i(K)$ is then defined exactly as in Sec. 4, where we use the perversity $\bar{p}$ to define the required binary function.

Singular Intersection Homology. If X is a stratified space, we may wish to make reference to its intersection homology without bothering to triangulate it. This can be done by considering its singular intersection homology groups ([22]), defined as follows.

Let $\sigma \in S_i(X)$ be a singular i-simplex. This means that $\sigma : \Delta^i \to X$ is a continuous map from the standard i-

simplex into X. One says that σ is $\bar{p}$-proper if $\sigma^{-1}(X_{d-k})$ is contained within the $i - k + p_k$ skeleton of Δ^i, for each k, and then proceeds exactly as above to define the singular intersection homology groups.

For any "good" triangulation of a stratified space, the simplicial intersection homology groups will match up with the singular ones. We make this notion of "good" precise below (SubSec 6.2).

Example: Suspended Torus. To illustrate the definition, we now calculate the intersection homology groups of the suspended torus using the two perversities $\bar{p} = (-1, 0, 0)$ and $\bar{q} = (0, 0, 1)$. Any edge whose closure contains the codim-three singularity a (or b) cannot be proper for either perversity, since we this would require $\dim(\bar{e} \cap X_0) \le 1 - 3 + q_1 = -1$. Thus, no single point in ΣT is a boundary. On the other hand, any two vertices in the smooth part of ΣT can be connected via a path which entirely avoids the two singular points. Hence, $I^p H_0(\Sigma T) = I^q H_0(\Sigma T) = \mathbb{Z}/2\mathbb{Z}$.

The sum ξ of all three-simplices in any triangulation of ΣT necessarily contains the singular points. If σ is one such three-simplex, then from the computation $\dim(\bar{\sigma} \cap X_0) = 0 \le 3 - 3 + p_3 = 0$, we see that ξ is a sum of proper simplices. Since $\partial \xi = 0$ and thus trivially a sum of proper simplices, ξ is allowable. Hence we have: $I^p H_3(\Sigma T) = I^q H_3(\Sigma T) = \mathbb{Z}/2\mathbb{Z}$.

In dimensions 1 and 2, the two perversities give different answers. For $\bar{p}$, the 2-simplices which we obtain by coning the boundary circles of the torus to either one of the singular points are not proper: for example, $\dim((C_1 * a) \cap X_0) = 0 > 2 - 3 + p_3 = -1$. Hence the boundary circles C_1 and C_2 are allowable 1-cycles which are not the boundary of an *allowable* 2-chain, from which we see that $I^p H_1(\Sigma T) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$, with basis elements the homology classes of C_1 and of C_2. On the other hand, $I^p H_2(\Sigma T) = 0$.

Replacing $p_3 = 0$ with $q_3 = 1$ in the above discussion shows that $I^q H_1(\Sigma T) = 0$, while $I^q H_2(\Sigma T) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$, with basis elements the homology classes of ΣC_1 and ΣC_2.

Example: Pinched Torus with Disc. Let Y be the pinched torus with disc stretched across the hole (see Fig. 9). Consider the two perversities $\bar{p} = (-1, 0)$ and

$\bar{q} = (0,0)$.

For the first perversity, there are two distinct components, represented by points on the interior of the torus and the disc, respectively. Note that the two points are not $\bar{p}$−homologous, as any 1-chain which crosses the boundary of the disc will necessarily contain an improper edge. Hence $I^p H_0(Y) = \mathbb{Z}/2\mathbb{Z} \bigoplus \mathbb{Z}/2\mathbb{Z}$. In dimensions one and two, there are no $\bar{p}$-cycles.

For the latter perversity, there are no components: any point on Y can be connected to Σ by an allowable 1-chain. Since we are computing mod Σ, this point becomes a boundary. There is also no $\bar{q}$-homology in dimension one. On the other hand, the group $I^q H_2(Y)$ has rank two. As representatives, one may take the attached disc (whose boundary is in Σ), and the pinched torus without the disc, which has empty boundary.

Duality   The last two examples illustrate the following theorem ([18]):

2 (Poincaré Duality) *Let X be a d-dimensional stratified space with $\bar{p}, \bar{q}$ dual perversities. Then, for all r, there is a perfect pairing given by intersection of chain representatives:*

$$I^p H_r(X) \otimes I^q H_{d-r}(X) \to \mathbb{Z}/2\mathbb{Z}$$

In the suspended torus example, note that the chains C_1 and ΣC_2 intersect in precisely one point; the former represents a one-dimensional $\bar{p}$-class, the latter a two-dimensional $\bar{q}$-class. We also have:

3 (Lefschetz Duality) *Let X be a d-dimensional stratified space with boundary ∂X and let $\bar{p}, \bar{q}$ be dual perversities. Then, for all r, there is a perfect pairing:*

$$I^p H_r(X) \otimes I^q H_{d-r}(X, \partial X) \to \mathbb{Z}/2\mathbb{Z}$$

6.2   Stratification and Triangulation Effects

Stratification Dependence. A natural question is whether the intersection homology groups of a stratified space X depend on the stratification. Certain assumptions ([18]) on both space and perversity guarantee independence. Specifically, one requires that $S_{d-1} = \varnothing$ and
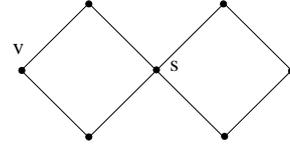
Figure 10: A wedge of two circles.
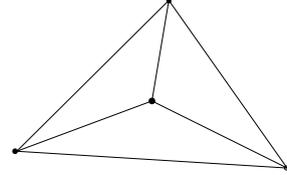
Figure 11: A two-sphere, stratified to have four isolated singular points, with a non-flaglike triangulation.

$p_i \le p_{i+1} \le p_i + 1$. In our more general context, however, the intersection homology groups will depend on the stratification.

As an example, consider the wedge of two circles, X, shown in Fig. 10. The coarsest stratification of X simply places the wedge point s into the 0-stratum. If we compute using perversity $\bar{p} = (-1)$, we find two components, since any allowable edge must exclude s. On the other hand, nothing stops us from placing both s and some smooth point v in into the 0-stratum. This choice creates an extra $\bar{p}$-component.

Triangulation Dependence. In addition to stratification, intersection homology groups also depend on the choice of triangulation. For example, suppose that X is a two-sphere stratified to have four isolated singular points. Triangulate X as the boundary of a tetrahedron with the singular points as vertices (Fig. 11) and attempt to compute its intersection homology using the perversity $\bar{q} = (0,0)$. Unfortunately, we get a ludicrous answer: there are no components because there are no allowable vertices!

On the other hand, if we take the barycentric subdivision, we create allowable vertices and thus a component. Hence two triangulations of the same stratified space give different intersection homology groups.

12

Flaglike Triangulations. Fortunately, dependence on triangulation is not very strong. 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 $\bar{\sigma}$. Note that the triangulation in Fig. 11 is not flaglike: in fact, all edges violate the condition.

It can be shown ([23]) that simplicial intersection homology groups, computed using a flaglike triangulation, are isomorphic to singular intersection homology groups. Furthermore, the first barycentric subdivision of *any* triangulation will always be flaglike.

6.3 Intersection Homology Persistence

Given a stratified space equipped with an ascending and descending filtration, we define the notions of intersection homology persistence and extended persistence in exact analogy to the ordinary homology case. The pair groups with perversity $\bar{p}$ are denoted by $I^p H_r^{i,j}$. If the filtrations come from a function f, then we use $I^p Dgm(f)$ to refer to the associated persistence diagram, and we also make the obvious adjustments to the notation for the subdiagrams.

In [12], the authors prove that the persistent *homology* diagrams $Dgm_r(f), Dgm_r(g)$ for two similar functions are themselves similar, in the sense that the bottleneck distance between the diagrams is bounded by the L_∞ distance between the functions. Their proof can be adapted, with only minor notational changes, to give:

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

$$d_B(I^p Dgm_r(f), I^p Dgm_r(g)) \leq ||f - g||_\infty$$

7 Stratified Morse Example

As stated before, one motivation behind the study of intersection homology persistence is to gain information about an embedded stratified space that would not be obtainable using standard homology persistence. Consider the 2-dimensional stratified space in Fig. 2, filtered as before

Figure 12: A portion of a disc (not pictured) is attached along the dotted line.
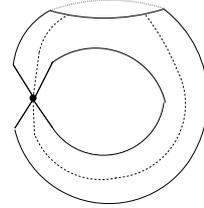
Figure 13: There are no allowable 1-cycles, as it is forbidden to touch the one-stratum with an edge

by height in the vertical direction. Recall that standard persistence did not detect the points B and F. On the other hand, if we compute persistence and extended persistence of intersection homology using this filtration, we will see it gives more information.

Ascending Past the Critical Points We fix our perversity $\bar{p} = (-1, 0)$, recall that this choice of perversity forbids edges to touch the one-stratum.

At A, a component is born which survives all the way to the top. At B (Fig. 12), two new components are born, one from the disc and one from the portion of the torus cut off by the one-stratum. This latter component is merged at point C, while the former is essential. At the pinch point D, a component is again born, which dies upon passing point E (Fig. 13). At point F (Fig 14), a one-cycle is born, represented by the circle which forms the boundary of the sublevel set; note that this circle is not trivial: since 2-simplices cannot touch the one-stratum along an edge, the chain formed by triangulating the entire sublevel set is not allowable and hence the boundary circle does not in fact bound. This one-cycle is subsequently capped off by the global maximum point G.

Descending with Relative Intersection Homology Now we begin the descent, where we quotient out by superlevel sets as we pass each critical point in turn. At

13
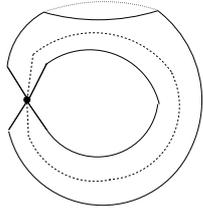
Figure 14: The circle on top represents a nonzero intersection homology class.
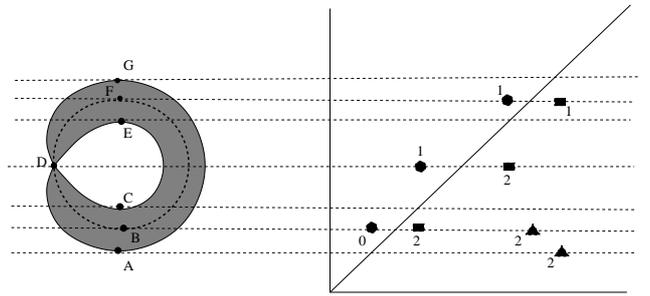
Figure 16: Points are labelled by dimension. Circles, squares, and triangles represent ordinary, relative, and extended pairs, respectively.
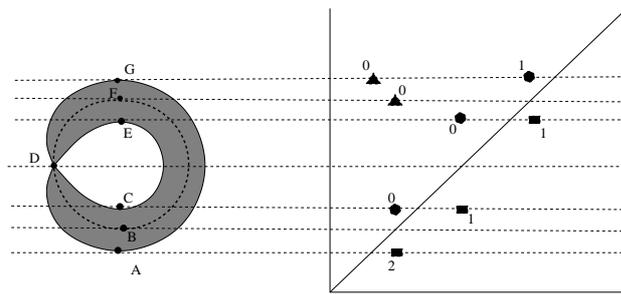
Figure 15: Points are labelled by dimension. Circles, squares, and triangles represent ordinary, relative, and extended pairs, respectively.

point G, the component represented by the torus becomes trivial, and so we get the extended pairing of A with G. Passing point F (Fig. 13), two things happen: the component represented by the disc becomes trivial, and a relative one-cycle is born and subsequently killed at point E. D and C form a relative pair, as the former creates a relative one-cycle and the latter kills it. B and A are paired for the same reason in dimension two.

The results of this analysis are summarized in the persistence diagram $I^p Dgm(f)$, shown in Fig. 15 along with the space itself. Note the lack of symmetry in this diagram, reflecting the failure of Poincaré duality using a single perversity, even for intersection homology.

On the other hand, suppose we compute persistence for the same filtration using the dual perversity $\bar{q} = (0,0)$. Then, for example, there will be an ordinary pair $(f(A), f(B))$, representing the component born at the global minimum which becomes trivial upon first touching the one-stratum. This pair is dual to the relative 2-

dimensional pair $(f(B), f(A))$ which we computed in the $\bar{p}$-diagram. We omit any further calculations, but the complete diagram $I^q Dgm(f)$ is shown in Fig. 16. Comparing this to Fig. 15, we see an obvious symmetry. In the next section, we explain where this comes from.

8 Duality and Symmetry

In this section, we explain why the diagrams above are symmetric by proving that a new version of Pair Group Duality (Res. 1) holds in this context. From this result, we can derive several other symmetry results for our diagrams.

Pair Group Duality   Suppose we have a d-dimensional stratified space X endowed with a stratified Morse function f. Assuming that f has n ordered critical points $\{t_1, \ldots, t_n\}$, we define the ascending filtration A_i and the descending filtration D_{n-i} as above (SubSec 3.2): $A_i = f^{-1}((-\infty, a_i]$) and $D_{n-i} = f^{-1}([a_i, \infty))$, where a_i is a regular value with $t_i < a_i < t_{i+1}$. Assume also that $\bar{p}, \bar{q}$ are dual perversities.

For each i, the space A_i will itself be a d-dimensional stratified space with boundary ([20]). Hence, for each dimension r, Lefschetz Duality (3) gives a perfect pairing:

$$I^p H_r(A_i) \otimes I^q H_{d-r}(A_i, \partial A_i) \to \mathbb{Z}/2\mathbb{Z}.$$

On the other hand, $\partial A_i = f^{-1}(a_i) = \partial D_{n-i}$. Combining this fact with excision, we find a perfect pairing:

$$I^p H_r^i \otimes I^q H_{d-r}^{2n-i} \to \mathbb{Z}/2\mathbb{Z}.$$

This pairing can be further refined in the following sense. Suppose that we have a $\bar{p}$-class $\alpha \in I^p P_r^{i,j}$, so α is born at level i and dies entering level j. For each level of its lifetime, α is paired with some $(d-r)$-dimensional $\bar{q}$-class. But an inspection of the following diagram:

$$I^p H_r^{i-1} \longrightarrow I^p H_r^i \longrightarrow I^p H_r^{j-1} \longrightarrow I^p H_r^j$$

$$\otimes \qquad\qquad \otimes \qquad\qquad \otimes \qquad\qquad \otimes$$

$$I^q H_{d-r}^{2n-i+1} \longleftarrow I^q H_{d-r}^{2n-i} \longleftarrow I^q H_{d-r}^{2n-j+1} \longleftarrow I^q H_{d-r}^{2n-j}$$

$$\downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow \qquad\qquad \downarrow$$

$$\mathbb{Z}/2\mathbb{Z} \qquad \mathbb{Z}/2\mathbb{Z} \qquad \mathbb{Z}/2\mathbb{Z} \qquad \mathbb{Z}/2\mathbb{Z}$$

shows that this class must have been born at level $2n - j + 1$ and died entering level $2n - i + 1$. This proves:

5 (IH Pair Group Duality) *For each r and for $0 \le i \le j \le 2n$, whenever $\bar{p}, \bar{q}$ are dual perversities, there is a perfect pairing:*

$$I^p P_r^{i,j} \otimes I^q P_{d-r}^{2n-j+1,2n-i+1} \to \mathbb{Z}/2\mathbb{Z}$$

Diagram Symmetries The algebraic result above can be made more concrete in terms of symmetries of persistence diagrams. The proofs are almost identical to those for ordinary homology found in [11], and we omit them here.

To state these results, we define three involutions of the plane: $(x, y)^T = (y, x), (x, y)^R = (-y, -x)$, and $(x, y)^O = (-x, -y)$. Then, given f and dual perversities $\bar{p}, \bar{q}$ we have:

6 (Diagram Symmetry I)

$$I^p Dgm_r(f) = [I^q Dgm_{d-r}(f)]^T.$$

Furthermore,:

- $I^p Ord_r(f) = [I^q Rel_{d-r}(f)]^T$

- $I^p Rel_r(f) = [I^q Ord_{d-r}(f)]^T$

- $I^p Ext_r(f) = [I^q Ext_{d-r}(f)]^T$

An example of these relations can be in seen in Figs. 15 and 16. On the other hand, suppose we also filter our space with the function $-f$. Then we have:

7 (Diagram Symmetry II)

$$I^p Ord_r(f) = [I^q Ord_{d-r-1}(-f)]^R$$

$$I^p Rel_r(f) = [I^q Rel_{d-r+1}(-f)]^R$$

$$I^p Ext_r(f) = [I^q Ext_{d-r}(-f)]^O$$

Hence if all we care about is *persistence* of classes, rather than order or dimension, the information gained by the two filtrations will be identical.

Example. Recall the stratified space in Fig. 2 and let $\bar{p} = (-1, 0), \bar{q} = (0, 0)$. When we calculated I^p-persistence for the filtration defined by vertical height f, there was an ordinary 1-dimensional class born at point F which was then capped off at point G; this corresponds to the point $(f(F), f(G)) \in I^p Ord_1(f)$.

Suppose we now compute I^q-persistence using the filtration defined by $-f$. There will then be a component born at G which dies as soon as we pass F because any point which can be allowably connected to the singular set will become a boundary when using perversity $\bar{q}$. In other words, there will be a point $(-f(G), -f(F)) \in I^q Ord_0(-f)$.

Similarly, the essential $\bar{p}$-component represented by the point $(f(A), f(G)) \in I^p Ext_0(f)$ is partnered with the point $(-f(A), -f(G)) \in I^q Ext_2(-f)$, which represents the essential $\bar{q}$-void formed by the torus itself.

The complete persistence diagrams for the two functions are shown in Fig. 17.

9 Subdivided Star Filtration

The algorithm given in Sec. 4 computes the intersection pair groups for a simplicial complex equipped with an ordering on its simplices. On the other hand, we also have a
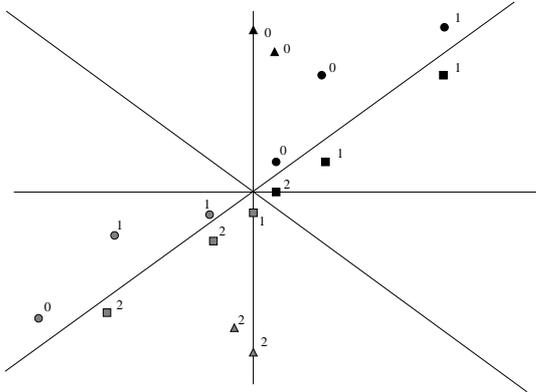
Figure 17: The diagrams $I^p Dgm(f), I^q Dgm(-f)$, in all three relevant dimensions, are superimposed, with the former diagrams in darker shading. Circles, boxes, and triangles indicate ordinary, relative, and extended points, respectively. Each point is labelled by dimension.

notion of intersection homology persistence for an actual stratified space equipped with a height function (see Secs. 7 and 8). In this section, we connect the two concepts. In fact, we demonstrate that the latter type of persistence can be approximated, with as much precision as we like, by the former.

Vertex Ordering. Suppose that we have a simplicial complex K embedded in Euclidean space and an injective, real-valued function h defined on the vertices of K. Interpolate to get $h : K \to \mathbb{R}$, then order the vertices of K so that $v < w$ iff $h(v) < h(w)$. Let $h(v_i) = r_i$, pick some t_i just larger than r_i, and consider the sublevel sets $K_{\leq t_i}$ and superlevel sets $K_{\geq t_i}$. These sets provide ascending and descending filtrations of $|K|$, the topological realization of K, and we can compute intersection homology persistence along these filtrations.

Simplicial Analogue. We may also filter K in the following manner. Letting K' denote the first barycentric subdivision, and assuming K has n vertices ordered by h-value, we set:

$$\bar{K}^i = \bigcup_{j \leq i} \bar{st}(v_j, K') \tag{2}$$

$\bar{K}^i$ consists of the complete star within K' of the first i

original vertices from K. We give each simplex in $\bar{K}^i - \bar{K}^{i-1}$ the value $h(v_i)$, set:

$$\bar{L}^{n-i} = \bigcup_{s > i} \bar{st}(v_s, K') \tag{3}$$

and give each simplex in $\bar{L}^{n-i} - \bar{L}^{n-i-1}$ the value $h(v_i)$. These two collections then give ascending and descending filtrations of K and we can compute intersection homology persistence in the sense of Sec. 4.

For each i there exist stratum-preserving deformation retractions between $K_{\leq t_i}$ and K^i, and also between $K_{\geq t_i}$ and L^{n-i}. These retractions induce intersection homology isomorphisms which commute with the homomorphisms induced by inclusions along the two sets of ascending and descending filtrations. In other words, the two persistence diagrams will be identical.

Note also that $\bar{K}_i$ and $\bar{L}_{n-i}$ are stratified subcomplexes of K. Furthermore, their boundaries are the same, equal to the full subcomplex of K' spanned by the barycentres of the simplices in K which are spanned by at least one vertex lower than or equal to v_i and at least one vertex higher than v_i. Hence the duality results derived above also go through perfectly in this context.

Persistence Diagram Approximation Given a stratified space X along with a function f, we might wish to actually compute the persistence diagrams associated to the filtrations provided by f. The discussion above shows us that we can by choosing a triangulation K of X and defining $\tilde{f}$ via linear approximation from the values of f on the vertices of K. Then we compute persistence using the subdivided star filtrations above. By choosing a fine enough triangulation K of X, we can make the function f and $\tilde{f}$ arbitrarily close. And so by Diagram Stability (4), the persistence diagrams for the two functions will also be arbitrarily close.

10 Discussion

We list here some further thoughts and questions:

- At the present moment, ϕ-persistence exists only as a convenient abstraction that helps to explain and com-

pute the more concrete notion of intersection homology persistence. Are there examples of simplex removal decision procedures, other than those derived from a perversity, that might be of interest?

- The assumption that the top stratum be dense in X seems to be necessary for the definition of intersection homology. On the other hand, this seems an unfortunate requirement, as one might easily imagine datasets where this might not hold. Can the theory be redeveloped to deal with this issue?

- Given a dataset, one might envision fitting to it a family of stratified spaces, with a changing stratification, each based on different uncertainty levels. Although the structure of this family is not yet clear, it seems likely that intersection homology might aid in its analysis: if the family can be structured in such a way so that intersection homology homomorphisms are induced between different spaces in the family, then we could do persistence along these maps. This fits in with the basic governing paradigm of persistence: if there is a parameter of whose value you are not certain, don't fix the parameter. Rather, vary the parameter, compute persistence in some way, and look for islands of stability.

References

[1] P.K. Agarwal, H. Edelsbrunner, J. Harer, and Y. Wang. Extreme elevation on a 2-manifold. In *Proc. 20th Ann. Sympos. Comput. Geom.*, pages 357–365, 2004.

[2] Y.-H. A. Ban, H. Edelsbrunner, and J. Rudolph. Interface surfaces for protein-protein complexes. In *Proc. 8th Intl. Conf. Res. Comput. Mol. Bio.*, pages 205–212, 2004.

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. In *Neural Computation*, pages 347–356, 2003.

[4] P. Bendich. *Analyzing Stratified Spaces Using Persistent Versions of Intersection and Local Homology*. PhD thesis.

[5] P. Bendich, D. Cohen-Steiner, H. Edelsbrunner, J. Harer, and D. Morozov. Inferring local homology from sampled stratified spaces. In *Proc. 48th Ann. Sympos. Found. Comp. Sci.*, pages 536–546, 2007.

[6] P. Bendich and J. Harer. Extreme elevation on a stratified surface. In Preparation.

[7] F. Cagliari, M. Ferri, and P. Pozzi. Size functions from the categorical viewpoint. *Acta Appl. Math.*, **67**:225–235, 2001.

[8] G. Carlsson, A. Collins, L. Guibas, and A. Zomorodian. Persistence barcodes for shapes. *Internat. J. Shape Modeling*, 2005.

[9] G. Carlsson and A. Zomorodian. Computing persistent homology. In *Proc. 20th Ann. Sympos. Comput. Geom.*, pages 347–356, 2004.

[10] G. Carlsson and A. Zomorodian. Localized homology. In *Shape Modeling International*, 2007.

[11] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Extending persistence using Poincare and Lefschetz duality. *Found. Comput. Math.* to appear.

[12] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. In *Proc. 21st Ann. Sympos. Comput. Geom.*, pages 263–271, 2005.

[13] D. Cohen-Steiner, H. Edelsbrunner, and D. Morozov. Vines and vineyards by updating persistence in linear time. In *Proc. 22nd Ann. Sympos. Comput. Geom.*, pages 119–126, 2006.

[14] H. Edelsbrunner and J. Harer. Persistent homology-a survey. In J.E. Goodman, J. Pach, and R. Pollack, editors, *Twenty Years After*. AMS. to appear.

[15] H. Edelsbrunner, J. Harer, and A. Zomorodian. Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds. *Discrete Comput. Geom.*, **30**:87–107, 2003.

[16] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete Comput. Geom.*, **28**:511–533, 2002.

[17] P. Frosini and C. Landi. Size theory as a topological tool for computer vision. *Pattern Recognition and Image Analysis*, **9**:596–603, 1999.

[18] M. Goresky and R. MacPherson. Intersection homology I. *Topology*, **19**:135–162, 1980.

[19] M. Goresky and R. MacPherson. Intersection homology II. *Inv. Math.*, **71**:77–129, 1983.

[20] M. Goresky and R. Macpherson. *Stratified Morse Theory*. Springer-Verlage, Berlin, 1987.

[21] B. Hughes and S. Weinberger. Surgery and stratified spaces. *Ann. of Math. Stud.*, **2**:319–352, 2001.

[22] H. King. Topological invariance of intersection homology without sheaves. *Topology Appl.*, **20**:149–160, 1985.

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

[24] J. Milnor. *Morse Theory*. Princeton Univ. Press, Princeton, New Jersey, 1963.

[25] V. Robins. Toward computing homology from finite approximations. *Topology Proceedings*, **24**:503–532, 1999.

[26] J. Tenenbaum, V. De Silva, and J. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, **290**:2319–2323, 2000.

[27] Y. Wang, P.K. Agarwal, P. Brown, H. Edelsbrunner, and J. Rudolph. Coarse and reliable geometric alignment for protein docking. In *Proc. Pacific Sympos. Biocomput.*, pages 65–75, 2005.

Appendix: Proof of Correctness

Here we give the proof that the ϕ-persistence algorithm is correct. We start by providing a precise statement of correctness. Recall that the bijection g tracks the reordering of the input simplices into proper and improper sets.

8 (Correctness of ϕ-persistence Algorithm) *For $0 < i, j \leq s$, (σ_i, σ_j) is computed by our algorithm iff $I^\phi P_r^{g(i),g(j)} = \mathbb{Z}/2\mathbb{Z}$, where $r = dim(\sigma_i)$.*

To prove this statement we will construct another reduction algorithm for D for which the associated *low*-function clearly and provably computes the correct persistence pairs. We then employ the Pairing Uniqueness Lemma ([13]) which states that any such *low* function must depend only on D and hence that our given algorithm computes the same pairs as the provably correct ones. This new reduction algorithm is built on two procedures $Make - Active_i$ and $Pair - Simplices$. $Make - Active_i$ decides if a given simplex σ_i is or is not active. The resulting active simplices are then input into the $Pair - Simplices$ procedure. We will prove that the analogous statement to Statement 8 above is true for this new algorithm.

Before giving these procedures, we first define a function $n : P_i(K) \to \{s + 1, s + 2, \ldots, m\}$ via $n(\gamma) = i$ where σ_i is the youngest (most recently added) simplex in $I(\gamma)$. If $I(\gamma) = \varnothing$, which means that γ is an allowable chain, we set $n(\gamma) = 0$.

Finding Neutral Simplices. Here is pseudocode for the recursive procedure $Make - Active_i$:

$\gamma_i = \sigma_i$
for $j = 1$ to $i - 1$ **do**
$\quad \gamma_j = Make - Active_j(\sigma_j)$
end for
while $\exists j < i$ such that $n(\gamma_j) = n(\gamma_i)$ **do**
$\quad \gamma_i \leftarrow \gamma_j + \gamma_i$
end while
Return γ_i

For each i, we define

$$a(\sigma_i) = \partial(\gamma_i) \tag{4}$$

Notice that σ_i is neutral iff $n(a(\sigma_i)) \neq 0$. If σ_i is indeed neutral, we define $low_M(\sigma_i) = n(a(\sigma_i))$. Otherwise, we leave $low_M(\sigma_i)$ undefined for the moment.

The rest of the proof now very closely follows the proof of correctness for the usual homology persistence algorithm given in [16]. As a few details are different, we will write a full description here, while maintaining some of the notation found therein.

Basis Construction. Recall (see e.g. [14]) that for each positive r-simplex σ_i there is an r-cycle c_i which contains σ_i as its only positive simplex. Let h_i denote the homology class of c_i. Now suppose $\alpha \in I^q H_r^{g(i)}$. Then α was born at some level $g(k) \leq g(i)$. Hence α can be written as a sum

$$\alpha = h_k + \sum_{j \in I(\alpha)} h_j, \tag{5}$$

where $I(\alpha)$ is a set of indices all less than i. In other words, for some subset of indices $I(i)$ taken from $\{1, 2, \ldots i\}$, the classes h_j, or more precisely $f_r^{g(j),g(i)}(h_j)$, for $j \in I(i)$, form a basis for the intersection homology group $I^q H_r^{g(i)}$. Using this fact, we define a function, $y : I^q H_r^{g(i)} \to \{1, 2, \ldots i\}$ by $y(\alpha) = k$, where k is defined as in Eqn. 5.

Pair-Simplices Algorithm. We now give an algorithm which pairs some of the active simplices. These pairings will complete the definition of the low_M function, the values of which have already been given for the neutral simplices. The algorithm maintains, for each r and k, a list P_r^k of the paired simplices at stage k of the algorithm. Here is the pseudocode:

$\forall r, P_r^0 = \varnothing$
for $j = 1$ to s **do**
$\quad$ **if** σ_j non-negative **then**
$\quad\quad \forall r, P_r^{g(j)} = P_r^{g(j-1)}$
$\quad$ **else**
$\quad\quad i = y([a(\sigma_j)])$
$\quad\quad k = \dim(\sigma_j)$
$\quad\quad P_k^{g(j)} = P_k^{g(j-1)} \cup \{(\sigma_i, \sigma_j)\}$
$\quad\quad \forall r \neq k, P_r^{g(j)} = P_r^{g(j-1)}$
$\quad$ **end if**
end for

18

We then finish the definition of low_M by defining $low_M(j) = i$ iff (σ_i, σ_j) is produced by $Pair - Simplices$.

Matrix Formulation. As with our original algorithm, the procedures above can all be accomplished by performing column operations on the original boundary matrix D, or alternatively, by multiplying on the right by a product of elementary matrices V.

Let us call a simplex σ_k *potentially neutral* iff $I(\sigma) \neq \varnothing$. The $Make - Active_k$ procedure manifests itself by adding columns from the left which correspond to actually neutral simplices, to column j. If the procedure succeeds in raising the lowest one in column j above the proper/improper demarcation line, then σ_j is in fact active. Otherwise, it is actually neutral and we never use its corresponding column again in the reduction. The rest of the algorithm just completes the reduction of the matrix, starting all over again from left to right, but this time only employing columns corresponding to active simplices.

At the end of the process, we have a reduced matrix M and a corresponding lowest one function low_M. If $low_M(j) = i > s$, then column j of V stores a chain of neutral simplices; in other words, a non-allowable chain. If column j of M is empty, then column j of V stores a cycle consisting of positive, negative, and neutral simplices; this cycle is precisely the representative of the basis element h_j described above. Finally, if $low_M(j) = i \leq s$, then column j of V stores an allowable chain which destroys the class created by σ_i. Hence we have one matrix D and two reduced matrices R, M which result from performing column operations on D. By the Pairing Uniqueness Lemma ([13]) then, we conclude $low_M = low_R$.

The correctness proof will therefore be complete after we prove the following, which is the analogue to Statement 8 for our algorithm.

9 *Let $r = dim(\sigma_i)$. Then $low_M(j) = i < s$ iff $I^q P_r^{g(i),g(j)} = \mathbb{Z}/2\mathbb{Z}$*

PROOF. We prove the forward direction; the other direction is essentially just a restatement of what follows below. Let c_i be the cycle containing σ_i as constructed above (5), and let h_i be its homology class in $I^q H_r^{g(i)}$. We show that h_i is born at level $g(i)$ and dies at level $g(j)$:

First, we show $h_i \notin \text{Im} f_r^{g(i-1),g(i)}$: Suppose it was, then $\exists \alpha \in I^p H_r^{g(i-1)}$ such that $f_r^{g(i-1),g(i)}(\alpha) = h_i$. Writing α as in (5), we arrive at a contradiction.

Next we show $f_r^{g(i),g(j)}(h_i) \in \text{Im} f_r^{g(i-1),g(j)}$: By construction, $y(a(\sigma_j)) = i$. This means that, dropping the maps induced by inclusion for the moment, we can write:

$$a(\sigma_j) = c_i + \sum_{k \in I} c_k \qquad (6)$$

where I is some set of indices less than i. We then pass this equation to homology and push it forward to level $g(j)$, where $[a(\sigma_j)] = 0$, since by construction (see (4)), $a(\sigma_j) = \partial(\gamma_j)$. Hence at level $g(j)$, we see that the image of h_i is equal to the image of the sum of classes on the right. But all of these classes existed at levels lower than $g(i)$. Hence h_i died at least by level $g(j)$.

Finally, we show that $f_r^{g(i),g(j-1)}(h_i) \notin \text{Im} f_r^{g(i-1),g(j-1)}$. Suppose it were. Then at level $g(j-1)$, the image of h_i is homologous to the image of a class coming from before level $g(i)$. Hence, using the basis defined in (5) and dropping maps, we can find an allowable chain $\eta \in I^q H_{r+1}^{g(j-1)}$ such that

$$\partial(\eta) = c_i + \sum_{t \in J} c_t,$$

for some set of indices less than i. Notice that $\partial(\eta)$ is homologous to zero at level $g(j-1)$. So we can add this equation to (6) and pass to homology to obtain:

$$[\partial(\eta)] + [a(\sigma_j)] = [a(\sigma_j)] = \sum_{k \in I} h_k + \sum_{t \in J} h_t.$$

But all the indices on the right hand side are less than i, and so this contradicts the definition of $y(a(\sigma_j))$. Therefore, h_i dies at level $g(j)$.