FAITHFUL SETS OF TOPOLOGICAL DESCRIPTORS AND THE ALGEBRAIC K-THEORY OF MULTI-PARAMETER ZIG-ZAG GRID PERSISTENCE MODULES

by

Anna Katherine Schenfisch

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2023

ACKNOWLEDGEMENTS

My work is supported by the NSF GRFP Grant No. DGE-1649608 as well as the NIH/NSF DMS Grant No. 1664858.

I first thank the many wonderful faculty and staff who have supported me during my time here. In particular, I thank Dr. Brittany Fasy, my advisor, for all her help and guidance. She has strengthened in me a love of grammar, the joy of small details, and an appreciation for the ways that geometry, topology, and computer science can overlap. I also thank Dr. Ryan Grady for his many semesters of mentoring and support; I always leave our meetings reenergized, with some new mathematical or philosophical perspective to consider. Thanks also to the other members of my committee, Dr. Tomas Gideon, Dr. David Millman, and Dr. Dominique Zosso. Finally, I thank Katie Sutich, Jane Crawford, and Emily Keller; their "good mornings" are often are my first interactions each day, and their work to keep the department running smoothly is invaluable.

I am deeply grateful for all the friends I have made during my time in Bozeman. Thanks to Robin, Marziah, Kristen, Jan, Jackie, the Heppner/Adames and Narotzkys, Dan, Ben, Kenny, Luciano, Chris, Catherine, William, Jordan, Brad, Lucy, Sam, Tori, Scotty, Hannah, and all the other wonderful humans who enrich my life. Friends need not be human; I thank Barbara, Pieper, Hans, Alecto, and the other animals with whom I was fortunate to spend time. They remind me the importance of taking walks and naps, and occasionally help me solve combinatorics problems. Finally, I thank my family for all their love. I am especially grateful to my brother Mark, who has been my friend and advocate since day zero.

TABLE OF CONTENTS

1. PRELIMINARIES 1

1.1 Persistence Modules 1
1.5 Simplicial Complexes and Filtrations 3
1.9 Topological Descriptors 7
1.16 Faithful Sets and Topological Transforms 13

2. MOTIVATING QUESTIONS AND MAIN RESULTS 16

3. FAITHFUL DISCRETIZATIONS OF TOPOLOGICAL TRANSFORMS 26

3.1 A Faithful Discretization of the Augmented Persistent Homology Transform 26
3.1.1 Background Definitions 26
3.2.1 Tools for Building a Faithful Discretization 27
3.8.1 The Main Result: A Faithful Discretization of the APHT 32
3.12.1 Reconstruction Algorithm for Simplicial Complexes in $\mathbb{R}^d$ 35
3.12.2 Vertex Reconstruction 35
3.13.1 Higher-Dimensional Simplex Reconstruction 36
3.13.1.1 Computing k-Indegree 36
3.16.0.1 Simplex Certificate 41
3.17.1 Proof of Theorem 3.12 and Its Corollaries 43
3.19.1 Explicit Construction of a Faithful Discretization 45
3.19.2 Auxiliary Constructions 45
3.23.1 Building an Explicit Set 55
3.23.1.1 Directions for Vertices 56
3.25.1 Directions for Higher-Dimensional Simplices 59
3.27.1 Example of Building a Faithful Set 61
3.27.2 Tilt Examples 62
3.27.3 Vertex Isolating Directions 63
3.27.4 Perpendicular Directions 64
3.27.5 Simplex-Isolating Directions 67
3.28 Discretization is Faithful for Other Topological Transforms 69
3.28.1 Faithful Discretization of the ABCT/Other Augmented Dimension-Returning Transforms 69
3.29.1 Faithful Discretization of the AECC 70
3.35 Stability 76
3.41 Edge Reconstruction and Representation Using a Radial Binary Search 81
3.41.1 Tools 81
3.42.1 Oracle Framework 84

TABLE OF CONTENTS – CONTINUED

3.43.1 Constructions and Data Structures 84
3.46.1 The Reconstruction 88
3.46.2 Fast Edge Reconstruction 88
3.49.1 Putting it Together: Full Reconstruction 100
3.53 Discussion 102

4. ORDERING CLASSES OF TOPOLOGICAL DESCRIPTORS 104

4.1 Introduction 104
4.2 Preliminary Considerations 105
4.3.1 Zoo of Descriptor Examples 106
4.3.2 Cardinality Conventions 107
4.5 Relating Descriptors 109
4.14 The Topological Descriptors 114
4.15 Partial Order 116
4.22 Bounds on Faithful Sets 121
4.22.1 Non-Augmented Descriptor Bounds 121
4.27.1 Augmented Descriptor Bounds 128
4.33 Discussion 134

5. ZIG-ZAG PERSISTENCE MODULES: COSHEAVES AND K-THEORY 136

5.1 Introduction 136
5.2 Why K-theory? 138
5.2.1 Flavors and History of K-theory 139
5.2.2 K-theory and persistence 141
5.2.3 What we do 142
5.3 Constructible coSheaves 144
5.3.1 Stratified/Constructible Basics 144
5.16.1 Operations on coSheaves 150
5.17.1 Entrance Paths and Their Representations 151
5.22 Waldhausen's K-theory 153
5.28 Persistence Modules, Persistence Cosheaves, and Filtrations 157
5.28.1 Persistent Definitions 158
5.32.1 Filtered Spaces and Cosheaves 164
5.32.1.1 Monotone and Index Filtrations 164
5.32.1.2 From Index Persistence Modules to Monotone Persistence Modules 165
5.34.0.1 One-Parameter Augmented Descriptors via Index Filtrations 167
5.36.0.1 Multi-Parameter Augmented Descriptors 173

TABLE OF CONTENTS – CONTINUED

5.37 An Equivalence Result 175
5.42.1 Proof of Theorem 5.42 177
5.42.2 Psuedo and lax (co)limits 178
5.45.1 The Grothendieck Construction 179
5.47.1 Proving the Theorem 180
5.51 K-theory of Multi-Parameter Modules 182
5.52.1 K-Theory of Vect-Valued coSheaves 182
5.61.1 Pointed Set Valued coSheaves 193
5.65 Euler Characteristic Surfaces and Virtual Diagrams 197
5.65.1 Euler Characteristic Surfaces and K_0 197
5.67.1 Virtual Diagrams 199
5.70 Discussion 203

REFERENCES CITED 205

LIST OF TABLES

Table Page

3.1 Directions that are vertex- and simplex-isolating for the simplicial complex of Figure 3.5, computed as described in Construction 1. 68

3.2 Attributes of the edge arc object. 86

LIST OF FIGURES

Figure Page

1.1 The six descriptors of interest generated by the direction indicated by the arrow for the simplicial complex at the top right of the figure. Both $\hat{\chi}$ and $\hat{\beta}$ are shown as two plots, the projections onto the first and second coordinates. 12

3.1 Wedge centered at a one-simplex 31

3.2 Vertex-Isolating Directions 36

3.3 Computing k-indegree 37

3.4 Geometric interpretation of Algorithm 3.6 51

3.5 The simplicial complex K used as an example in this section. We first find vertex-isolating directions: the basis vectors, along with a fourth direction that orders the filtration grid of K_0 with respect to the basis vectors uniquely. Next, for σ_1 and σ_2, we find simplex-isolating directions; directions that are K_0-perpendicular to σ_1 and σ_2, as well as directions that tilt these perpendicular directions so that proper faces of σ_1 (and σ_2) are "popped" above σ_1 (respectively, σ_2). 62

3.6 A two-simplex in $\mathbb{R}^3$ (left) stratifies $\mathbb{S}^2$ where each stratum is a region containing all directions that define the same partial order on vertices. Notice that the set of directions perpendicular to any pair of vertices forms a great circle and the two directions perpendicular to $[v_1, v_2, v_3]$ correspond to the two three-way intersections of these great circles. 78

3.7 An edge arc EA centered at vertex $EA.v = v$. Other attributes of the edge arc include its start and stop angles, $EA.\alpha_1 = 1.75$ radians and $EA.\alpha_2 = \pi \approx 3.14$ radians, the array of vertices $EA.verts = \{v_1, v_2\}$, and the count of edges $EA.count = 1$. Here, we also see that 1-$\textsc{Indeg}(v, e_1) = 1$ and 1-$\textsc{Indeg}(v, e_2) = 2$. 87

LIST OF FIGURES – CONTINUED

Figure Page

3.8 The splitting of edge arc EA into EA_ℓ and EA_r, as in Algorithm 3.12. The large gray region is the region containing all edges of *bigedges*. That is, all edges whose angle with the positive e_1-axis is at least $EA.\alpha_1$. On Line 5 of the algorithm, we compute the number of edges in EA_ℓ by first computing the indegree of $EA.v$ in direction s from the diagram in direction s, then we subtract the number of edges in *bigedges* that are below the height of $EA.v$ in direction s (i.e., below the blue line). By the pigeonhole principal, we find $EA_r.count = EA.count - EA_\ell.count$. 90

3.9 One step of Algorithm 3.13 95

4.1 The two simplicial complexes considered in the proof of Lemma 4.20 119

4.2 Given $S = \{s, s'\}$, the envelope for σ is shown in grey (and includes the area covered by σ). 123

4.3 With only two directions in our direction set on the left side of the figure, we see that the envelope P_σ for the maximal edge σ is still three-dimensional (here, P_σ is the unbounded region above σ between the two planes). This means that we could construct an adversarial two-simplex (shown in pink) contained entirely in P_σ that would not be detected by either non-augmented descriptors in the set. On the right, the addition of a third direction reduces P_σ to a linear subspace.....124

4.4 Nested triangles as discussed in Lemma 4.29 129

4.5 The two simplicial complexes considered in the proof of Lemma 4.30.....130

4.6 The regions of observability described in the proof of Lemma 4.30. K is shown as solid black edges and K' as dashed edges. For any lower-star filtration in a direction contained in R_i, the event at vertex v_i differs when considering K or K', thus, such directions are able to distinguish K from K'. Note that any direction outside the regions of observability (i.e., the non-shaded portions of the circle) is not able to distinguish K from K'. 133

LIST OF FIGURES – CONTINUED

Figure Page

4.7 An example of Construction 3 for $m = 4$. The regions of observability are shown below each clothespin. By construction, each of these four double wedges define disjoint regions of $\mathbb{S}^2$. 134

5.1 A stratified circle, $S^1 \to [1]$ as in Example 5, where $v \mapsto 0$ and $\alpha \mapsto 1$. 145

5.2 The poset $I = [2]^2$ defines a natural cubulation of $\mathbb{R}^2$, which then leads to a natural stratification on $\mathbb{R}^2$, denoted $(\mathbb{R}^2; I)$. Each strata of $(\mathbb{R}^2; I)$ is a cube of the associated cubulation.....161

5.3 Examples of open intervals occurring in Construction 4 164

5.4 An example of the map C. The relevant interval for the point, e.g., m_2 is $[2, 7)$, since the image of each simplex added in that interval under the filter function f is m_2. Then $[2, 6)$ is mapped to m_2 and $[6, 7)$ is mapped to $[m_2, m_3)$. 167

5.5 A bifiltration (left) illustrating one nuance when defining an augmented filtration cosheaf over stratified $\mathbb{R}^2$ (right). In final event of the filtration (at $(2, 2)$), we see the instantaneous birth and death of a cycle coming from the horizontal inclusion map, but no such instantaneous event coming from the vertical map. There is therefore no natural assignment of a single vector space to the zero-strata at $(2, 2)$ that captures this behavior. 173

5.6 The unfolding of a three-dimensional cube C with height two in each parameter (an instance of the cube $X' \setminus \mathfrak{c}$ discussed in the induction on height in the proof of Theorem 5.59). The thick blue submodule U is a closed connected collection of codimension-two faces of the cube along which we can unfold (right). Since the unfolding (left) is a net, connected components of both the $C \setminus U$ and U cannot be more than two-parameter modules. 191

LIST OF FIGURES – CONTINUED

Figure Page

5.7 The result of applying δ to the cosheaf shown on the top of the figure is the persistence diagram shown on the bottom. In the middle, we have drawn the associated barcode. In the spirit of [14], we have shown all bars as closed intervals to emphasize that they do not necessarily arise from a monotone filtration. Note the presence of length-zero barcodes and on-diagonal points, corresponding to indecomposable elements supported at a single point. 201

LIST OF ALGORITHMS

Algorithm Page

3.1 `ComputeIndegree`$(\sigma, s, k, \hat{\rho}(K,S), T = \{\})$ 38

3.2 `Certificate`$(V, \hat{\rho}(K,S))$ 42

3.3 `ReconstructComplex`$(\hat{\rho}(K,S))$ 44

3.4 `Perp`(P) 46

3.5 `PlaneFill`(P, s) 49

3.6 `Tilt`(s, s', P) 51

3.7 `TiltToPop`(P, V, W, s) 53

3.8 `PointIso`(P) 57

3.9 `EvenIndegree`$(\sigma, s, \widehat{\chi}(K,S), T = \{\})$ 72

3.10 `AECCCertificate`$(V, K_{|V|-2}, \widehat{\chi}(K,S))$ 74

3.11 `AECCReconstructComplex`$(\widehat{\chi}(K,S))$ 77

3.12 `SplitArc`$(EA, bigedges, \theta)$ 89

3.13 `UpEdges`$(v, V_v, in_v, \theta, \hat{\rho}(e_2))$ 94

3.14 `FindEdges`(V) 98

ABSTRACT

Given a geometric simplicial complex, the uncountable set of (augmented) persistence diagrams corresponding to lower-star filtrations taken with respect to all possible directions uniquely correspond to the simplicial complex, i.e., the set is *faithful*. While this hints towards interesting applications in shape comparison, the set of all possible directions is uncountably infinite, and so has no hope of computability. In practice, one might use a finite approximation, but faithfulness of this approximation is not guaranteed.

Motivated by the need for both computability and provable faithfulness, we provide an explicit description of a finite faithful set of augmented persistence diagrams. We then show this construction applies to augmented Euler characteristic curves and augmented Betti curves, and is stable under particular perturbations. In the specific case where the underlying complex is a graph, we provide an improved construction that utilizes a radial binary search. We then shift focus to comparing the cardinalities of minimal faithful sets of descriptors as a way to define and order equivalence classes of topological descriptor types. Focusing on six topological descriptor types commonly used in practice, we give a partial order on their corresponding equivalence classes, as well as give bounds on the sizes of minimum faithful sets for each descriptor type.

Next, we broaden our view to zig-zag grid persistence modules, functors whose domain categories are posets with grid-like structure. We begin by explicitly defining such persistence modules in terms of constructible cosheaves over stratified Euclidean space, including a careful treatment of augmented persistence modules, which are analogous to the aforementioned augmented descriptors who played a central role in discussion of faithful sets. Exodromy gives us an equivalence between persistence modules as a functor category and as constructible cosheaves; we furthermore show the equivalence of these categories with a category of constructible functors out of $\mathbb{R}^d$ with a fixed stratification and localized at weak equivalences, essentially "standardizing" modules so that the category has a clear monoid structure. We compute the algebraic K-theory of zig-zag grid persistence modules, using a double inductive argument to show the K-theory is additive over strata. Finally, we identify connections to related topics, such as the virtual diagrams of Bubenik and Elchensen, as well as Euler characteristic and Betti curves/surfaces/manifolds. We hope a study of K-groups will provide interesting insights into the nature of persistence modules, and we indicate ways in which the zeroth and first K-groups may be interpreted.

CHAPTER ONE

PRELIMINARIES

Although each chapter utilizes its own specific terminology, all work presented here is in some way related to topological data analysis (TDA), and more specifically, persistent homology. This general toolbox contains concepts and definitions that are universially relevant to the dissertation, which we introduce here. We assume the reader is familiar with foundational topics in topology, such as homology. For further information, see, e.g., [25,38].

1.1 Persistence Modules

At their core, all subsequent chapters stem from the following definition.

Definition 1.2 (Persistence Module)**.** *Given a poset category $\mathcal{P}$, a* persistence module over $\mathcal{P}$ *is a functor $P : \mathcal{P} \to \Omega$.*

In practice, the target category Ω is often $\mathsf{Vect}_{\mathbb{F}}$ or $\mathsf{grVect}_{\mathbb{F}}$, the categories of finite-dimensional vector spaces over a field $\mathbb{F}$ or graded finite-dimensional vector spaces over $\mathbb{F}$, respectively. We may refer to the functor category of persistence modules over $\mathcal{P}$ as $\mathcal{P}$-*indexed persistence modules*. Commonly, we may see $\mathcal{P} = [\eta]$ for some $\eta \in \mathbb{N} \cup \{0\}$, which has objects $\{0, 1, \ldots, \eta\}$ and morphisms corresponding to the usual ordering on $\mathbb{N} \cup \{0\}$.

Commonly, and especially when invariants of modules such as persistence or rank invariant are being considered, we think of objects of $\mathcal{P}$ being a subset of the objects of some other ordered set, such as $\mathbb{R}^d$. This additional structure is a convenient way to define

the following specific type of persistence modules.

Definition 1.3 (Monotone Persistence Modules)**.** *Let $P : \mathcal{P} \to \Omega$ be a persistence module for $\mathcal{P}$, a poset whose objects are a subset of a poset $\mathcal{P}'$. If, for all $a, b \in \mathcal{P}$, $a \leq_{\mathcal{P}} b$ if and only if $a \leq_{\mathcal{P}'} b$, then we say that P is a* monotone poset with respect to $\mathcal{P}'$ *and the corresponding module is* monotone persistence module*. If P is a monotone persistence module and, specifically, we have $\mathcal{P}' = \mathbb{R}^d$ under the product order, we say P is a d-*parameter monotone persistence module*.*

Note that, for general values of d, d-parameter modules are often referred to collectively as *multi-parameter persistence modules*. Multi-parameter monotone persistence modules are a subset of a larger class of persistence modules, which we refer to as *zig-zag grid* modules, which have no restriction on the order on $\mathcal{P}$.

In order to formally define zig-zag grid modules, we note that for any locally finite poset, we may consider a representation of the poset via a directed graph that has vertices for each object, and a directed edge $q \to p$ whenever $q \leq p$. We simplify this representation using a *Hasse diagram*, which contains the minimum number of edges necessary to deduce all relations in the poset. We call the relations corresponding to the edges in the Hasse diagram *generating relations* of the poset. Hasse diagrams are typically drawn in $\mathbb{R}^2$ so that the direction of relation $q \leq p$ is represented with an edge such that the node representing q has a lower y-coordinate than the node representing p. When we discuss Hasse diagrams as being homeomorphic, we mean their corresponding undirected graphs are isomorphic.

Definition 1.4 (Multi-Parameter Zig-Zag Grid Persistence Modules)**.** *Let $P : \mathcal{P} \to \Omega$ be a persistence module with the objects of $\mathbb{P}$ a subset of $\mathbb{R}^d$. If the Hasse diagram of $\mathbb{P}$ is homeomorphic to the Hasse diagram of a product of d (non-trivial) one-parameter monotone posets with discrete sets of objects, we say that $P : \mathcal{P} \to \Omega$ is a d-*parameter zig-zag grid module*. Suppose $\mathcal{P}' \subset \mathbb{R}^d$ is a poset with the same objects as $\mathbb{P} \subset \mathbb{R}^d$ but with different orderings. Although $\mathcal{P}$ and $\mathcal{P}'$ may have homeomorphic Hasse diagrams, we emphasize that persistence modules over $\mathcal{P}$ and $\mathcal{P}'$ are distinct.*

Here, we caution the reader: historically, multi-parameter persistence modules have also been called multi-dimensional persistence modules. To avoid overuse of the word dimension, we elect to use the word "parameter" whenever appropriate.

Additionally, the term "persistence module" has been used in various instances in the literature to refer to only monotone persistence modules. We use the term persistence module to mean a general module of the form in Definition 1.2, and use the adjectives "monotone" or "zig-zag" and "zig-zag grid" wherever appropriate to avoid ambiguity.

1.5 Simplicial Complexes and Filtrations

In applications, persistence modules often arise from filtrations of topological spaces, in particular, filtrations of finite simplicial complexes. Filtering finite simplicial complexes is a classic source of one-parameter monotone persistence modules.

A *geometric k-simplex* σ is the convex hull of a set of $k+1$ affinely independent points in $\mathbb{R}^d$, denoted $\sigma = [v_0, v_1, \ldots, v_k]$. Each of these points is called a *vertex*, and we denote the vertex set of σ by $\text{verts}(\sigma)$; at times, we use σ in place of $\text{verts}(\sigma)$ in places where using $\text{verts}(\sigma)$ would make equations cumbersome to read. We write $\overline{\sigma}$ to denote the *closure* of σ,

which is σ taken together with all its faces, considered as a closed subspace of $\mathbb{R}^d$. We call k the *dimension* of σ, and define $\dim(\sigma) := k$. For another simplex τ, we say that τ is a *face* of σ and σ is a coface of τ if $\text{verts}(\tau) \subseteq \text{verts}(\sigma)$; we denote this relation by $\tau \preceq \sigma$. If $\tau \preceq \sigma$ but $\tau \neq \sigma$, then τ is called a proper face of σ, denoted by $\tau \prec \sigma$.

We topologize a simplicial complex K with the Alexandroff topology. We use the notation K_i to mean the set of i-simplices in the simplicial complex[1] and the number of simplices in K_i by n_i. We let $n = \sum_{i \in \mathbb{N}} n_i$ and $\kappa = \max n_i$. The *degree* of $v \in K_0$ is the number of one-simplices (edges) that are cofaces of v, and we denote this as $\deg(v)$. We denote the affine space spanned by V by $\text{aff}(V)$.

Unless explicitly stated otherwise, we only consider finite geometric simplicial complexes, and generally omit the words "finite" and "geometric" from their descriptions. Now that we have defined a simplicial complex, we define two particular types of filtration.

Definition 1.6 (Multi-Parameter Monotone Grid Filtration)**.** *Let K be a simplicial complex and $I = I_1 \times \ldots \times I_d$ where each $I_n \subset \mathbb{R}$ is finite. A d-*parameter filtration *of K is a sequence of subcomplexes $\{L_i\}_{i \in I}$ such that, for every $i < j$ (with respect to the product order on I) there is an inclusion of spaces $L_i \hookrightarrow L_j$ and so that $L_{\min\{i \in I\}} = \emptyset$ and $L_{\max\{i \in I\}} = K$.*

Note that the minimum (and maximum) of I is unique, and is equivalent to the point $(\min\{I_1\}, \min\{I_2\}, \ldots, \min\{I_d\})$ (and similarly for the maximum). Just as in the one-parameter case, all inclusion maps are in the direction of increasing index. If K has n non-empty simplices, the number of subcomplexes of the filtration is at most $(n+1)^d$.

[1] In Chapter 5, we use the notation $K_i(C)$ to mean the ith K-group of C. Since Chapter 5 never refers to the set of i-simplices in a complex, we hope this double use of notation is clear.

Applying the homology functor to a one-parameter filtration, we obtain the persistence module $P : [\eta] \to \mathsf{grVect}$, whose image may be specified as

$$H(L_0) \to H(L_1) \to H(L_2) \to \ldots \to H(L_\eta).$$

Here, we assume that homology is computed using field coefficients (e.g., $\mathbb{Z}_2$). Then, for each pair of complexes $L_i \subset L_j$ in a filtration, the homology groups $H_k(L_i)$ and $H_k(L_j)$ are vector spaces, and the maps $f_k^{i,j} \colon H_k(L_i) \to H_k(L_j)$ are linear maps. In a very similar manner, we may apply the homology functor to the d-parameter filtration of Definition 1.6 to obtain a d-parameter persistence module $P : I \to \mathsf{grVect}$.

In Chapters 3 and 4, we are especially interested in *lower-star filtrations*, which are one-parameter filtrations that naturally arise when considering how a simplicial complex is embedded/immersed in an ambient space.

Definition 1.7 (Lower-Star Filter and Filtration)**.** *For each vertex $v \in K_0$, the height of v in direction s is given by the dot product, $s \cdot v$. The* lower-star filter function in direction s*, denoted $h_s : K \to \mathbb{R}$, defines a "height" of each simplex in K where $h_s(\sigma) = \max\{s \cdot v \mid v \in \text{verts}(\sigma)\}$, i.e., $h_s(\sigma)$ is the height of the highest vertex in σ with respect to s. We omit the domain of h_s in notation when clear from context.*

Notice that, for $r, t \in \mathbb{R}$ such that $r \leq t$, we have $h_s^{-1}(-\infty, r] = h_s^{-1}(-\infty, t]$ if and only if no vertex has height in the interval $(r, t]$. If all vertices are at distinct heights with respect to s, then the total number of distinct subcomplexes (η in the above equations) is $\Theta(n_0)$ and

there exists an ordering of the vertices $\{v_0, v_1, \ldots, v_{n_0-1}\}$ such that our filtration is:

$$\emptyset \subset h_s^{-1}(-\infty, s \cdot v_0] \subset h_s^{-1}(-\infty, s \cdot v_1] \subset \ldots \subset h_s^{-1}(-\infty, s \cdot v_{n_0-1}].$$

Finally, we end by describing one more important type of filtration, the *index filtration.*

Definition 1.8 (One-Parameter Index Filter and Filtration)**.** *For simplicial complex $K \subset \mathbb{R}^d$ and filter function f, consider a total ordering of the simplices*

$$\sigma_1, \sigma_2, \ldots, \sigma_n \tag{1.1}$$

such that if either $f(\sigma_i) < f(\sigma_j)$ or $\sigma_i \prec \sigma_j$, then $i < j$. We call such a total ordering an index filter compatible with f*; note that multiple index filters may be compatible with f, and that all filters are related by a permutation of the simplices. Defining $L'_j := \{\sigma_i \mid i \leq j\}$, the* index filtration *is the nested sequence of subspaces of K:*

$$\emptyset = L'_0 \subset L'_1 \subset L'_2 \subset \ldots \subset L'_n = K.$$

If a d-parameter filtration is a one-parameter index filtration when restricted to one-parameter subfiltration, i.e., each inclusion map is the inclusion of a single simplex, we say it is a *d-parameter index filtration.* Just as in the one-parameter case, a d-parameter index filtration may be *compatible* with a d-parameter monotone filtration if we may arrive at the monotone filtration by composing strings of inclusions. Note that a d-parameter index filtration has at most $(n+1)^d$ subcomplexes.

Remark 1. *Index filtrations are themselves monotone. Index filtrations are compatible with themselves, but to no other index filtrations.*

1.9 Topological Descriptors

Any persistence module, whether thought of as arising from a filtration or thought of as an abstract functor, leads to *topological descriptors*. In some contexts, it is preferable to view topological descriptors as arising solely from persistence modules, in which case, topological descriptors need not have any connection to an underlying filtration or even simplicial complex. In other settings (indeed, most applications), it is preferable to view topological descriptors more concretely, as invariants of filtrations.

We defer choosing a viewpoint until necessary in Chapter 4 (Definition 4.3). For now, we identify six specific topological descriptors that are of particular interest.

Definition 1.10 (Persistence Diagram (ρ))**.** *Let $f : K \to \mathbb{R}$ be a filter function Letting $\beta_k^{i,j}$ denote the rank of $f_k^{i,j}$, $(a,b)^m$ denote m copies of the point (a,b), and $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, we define the* persistence diagram *as the following multiset:*

$$\rho_k(f) := \left\{ (m_i, m_j)^{\mu^{(i,j)}} \ \ s.t. \ (m_i, m_j) \in \overline{\mathbb{R}}^2 \ \ and \ \mu^{(i,j)} = \beta^{i,j-1} - \beta^{i,j} - \beta^{i-1,j-1} + \beta^{i-1,j} \right\}.$$

In the case of a k-dimensional persistence diagram, each $(m_i, m_j) \in \rho_k(f)$ represents a k-dimensional homological generator α that is born at i (that is, $[\alpha] \in H(L_i)$ but $[\alpha] \notin \mathrm{im}(f_k^{i-1,i})$) and dies going into j (that is, j is the smallest index such that there exists $[\alpha'] \neq [\alpha] \in H(L_{i-1})$ with $[\alpha] = [\alpha']$ in $H(L_j)$).

Since simplices can appear at the same parameter value in a general filtration, it is

possible that the birth and the death of a k-cycle happen simultaneously, in which case, that cycle is not represented in the persistence diagram. Similarly, if the introduction of a simplex in the filtration does not change the homology of the subcomplex, the simplex is not represented in the diagram. However, having every simplex "appear" in the persistence diagram is helpful, in addition to being natural. Thus, we introduce *augmented persistence diagrams* ($\hat{\rho}$'s), which contain this information.

Definition 1.11 (Augmented Persistence Diagram ($\hat{\rho}$))**.** *Given a filter $f\colon K \to \mathbb{R}$, let f' be a compatible index filtration. For $k \in \mathbb{N}$, the* k-dimensional augmented persistence diagram *is the following multiset:*

$$\hat{\rho}_k(f) := \{(f(\sigma_i), f(\sigma_j))\}_{(i,j)\in\rho_k(f')}\,, \tag{1.2}$$

The augmented persistence diagram ($\hat{\rho}$) *of f is the union of all k-dimensional augmented persistence diagrams: $\hat{\rho}(f) := \cup_{k\in\mathbb{Z}}\hat{\rho}_k(f)$.*

Note that while we make a distinction between ρ's and $\hat{\rho}$'s, the idea of augmented persistence diagrams is not new. Indeed, in many standard algorithms for computing persistence (see e.g., [25]), topological events with trivial lifespan are explicitly computed but then discarded from the output, since, in many data theoretic applications, these events are seen as noise. However, as we will see, events with trivial lifespan are vital in some applications, and have been used or discussed in the literature. For example, the definition of "persistence diagram" in McCleary and Patel [45] is the same as our definition of $\hat{\rho}$. In [46], Mémoli and Zhou define *barcodes with ephemeral points* equivalent to our notion

of $\hat{\rho}$ using filtered chain complexes. Using filtered chain complexes as a tool for obtaining an augmented structure is also explored in [15].

Recording invariants other than homology throughout a filtration leads to other topological descriptor types. For instance, recording Betti numbers gives us *Betti curves*. Just as with (augmented) persistence diagrams, there is a distinction between the augmented and non-augmented versions of Betti curves.

Definition 1.12 (Betti Curve (β) and Augmented Betti Curve ($\hat{\beta}$))**.** *Given a filter $f\colon K \to \mathbb{R}$, let f' be a compatible index filter. Let $m_1 < m_2 < \ldots < m_\eta$ be the ordered set of filter parameters of f that witness a change in homology and $\{L_i := f^{-1}(-\infty, m_i]\}_{i=1}^{\eta}$ be the filtration of K determined by f. The kth* Betti curve *(β_k) is a step function $\beta^{(f,k)} : \mathbb{R} \to \mathbb{Z}$ defined for $p \in [m_i, m_{i+1})$ by*

$$\beta_k^f(p) := \beta_k(L_{m_i}).$$

The collection of Betti curves for each dimension is known as the Betti curve (BC)*. We may omit the superscript f when it is clear from context.*

We call $\sigma \in K$ positive *(respectively,* negative*)* for *β_k if the inclusion of σ into the index filtration associated to f' increases (respectively, decreases) β_k. We will denote the positive (negative) i-simplices for β_k by L_i^+ (and L_i^-) Then the kth* augmented Betti curve *($\hat{\beta}_k$) is the function $\hat{\beta}_k^f : \mathbb{R} \to \mathbb{Z}^2$ defined by the following.*

$$\hat{\beta}_k^f(p) := \left(\left|\{\sigma \in L_k^+ \mid f(\sigma) \leq p\}\right|, \left|\{\sigma \in L_{k+1}^- \mid f(\sigma) \leq p\}\right| \right),$$

In other words, at every point of kth Betti curve, we know the number and dimension of k and $k+1$ simplices. The collection of augmented Betti curves for each dimension is known as the augmented Betti curve ($\hat{\beta}$). *Note then that, from the $\hat{\beta}$, we can see how many of each dimension of simplex was added in every step of the filtration. We may omit the superscript f when it is clear from context.*

We may also choose to record *Euler characteristic* throughout a filtration, which gives rise to *(augmented) Euler characteristic curves.*

Definition 1.13 (Euler Characteristic Curve (χ) and Augmented Euler Characteristic Curve ($\hat{\chi}$))**.** *Given a filter $f\colon K \to \mathbb{R}$, let f' be a compatible index filtration. Let $\{L_i := f^{-1}(-\infty, t_i]\}_{i=1}^{n}$ be the filtration of K determined by f. The* Euler characteristic curve *is a step function $\chi_f : \mathbb{R} \to \mathbb{Z}$ defined by the following equation:*

$$\chi_f(p) := \sum_{k=0}^{\infty} (-1)^k n_k^{(i)},$$

where $p \in [t_i, t_{i+1})$ and $n_k^{(i)}$ is the number of k-simplices in L_i.

Recall that in the index filtration associated to f', a single simplex is added at a time. Then from the curve $\chi_{f'}$, we can identify if a simplex added was even- or odd-dimensional, based on if inclusion of the simplex into the index filtration causes the Euler characteristic to increase or decrease. Then for the compatible monotone filtration f, the augmented Euler

characteristic curve ($\hat{\chi}$) *is the function* $\widehat{\chi}_f : \mathbb{R} \to \mathbb{Z}^2$ *defined by*

$$\widehat{\chi}_f(p) = \left(\sum_{k=0}^{\infty} n_{2k}^{(i)}, \sum_{k=0}^{\infty} n_{2k+1}^{(i)} \right)$$

In other words, $\hat{\chi}$ *represents* χ *as a parameterized count of even- and odd-dimensional simplices. We may omit the subscript* f *when it is clear from context.*

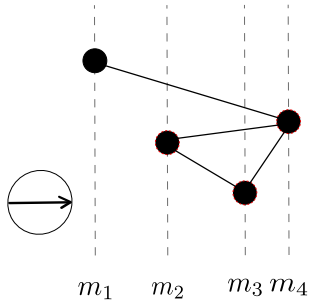
See Figure 1.1 for an example of each of the six types of topological descriptors defined above. Note that each definition of augmented descriptors above depended on a choice of compatible index filtration, although there are generally many compatible index filtrations for a particular monotone filtration. We therefore end with a justification that the resulting descriptor is independent of this choice, and first prove this claim for the case of $\hat{\rho}$'s.

Lemma 1.14 (Augmented Persistence Diagrams are Well-Defined)**.** *Let* K *be a simplicial complex, and let* $f\colon K \to \mathbb{R}$ *be a monotonic function. Then, the augmented persistence diagram* $\hat{\rho}_f$ *is well-defined.*

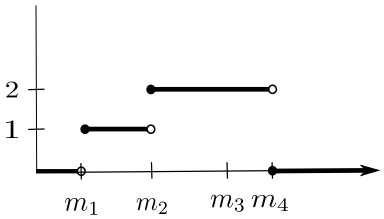
Proof. For $t \in \mathbb{N}$, let $f_t\colon K \to \mathbb{N}$ be all filter functions compatible with f ordered such that either and f_{t+1} are either same function ($f_t = f_{t+1}$), or they agree except for one transposition for each. As in Equation (1.2), define the following multisets:

$$\mathcal{A}_t := \{(f(\sigma_i), f(\sigma_j))\}_{(i,j)\in\rho_k(f_t)}$$

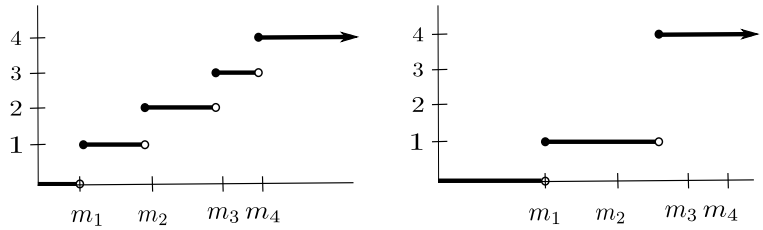
We prove by induction that $\mathcal{A}_t = \mathcal{A}_1$ for each $t \in \mathbb{N}$. The base case is simple: $\mathcal{A}_1 = \mathcal{A}_1$. For the inductive assumption, let $t \geq 1$ and assume that $\mathcal{A}_t = \mathcal{A}_1$. If $f_t = f_{t+1}$, we are



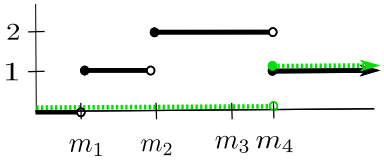
(a) K



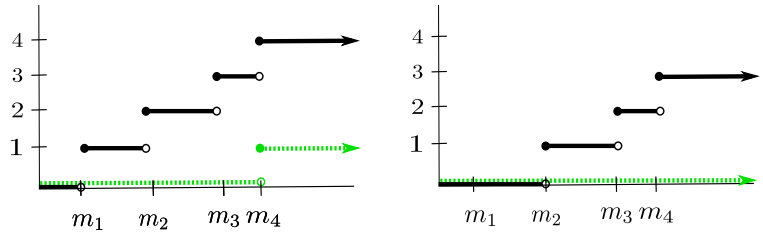
(b) ECC



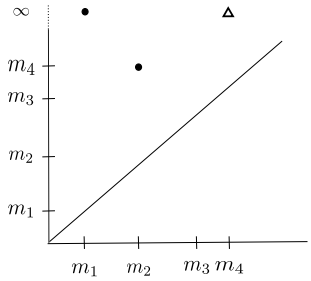
(c) $AECC$ (left: even-dimensional, right: odd-dimensional)



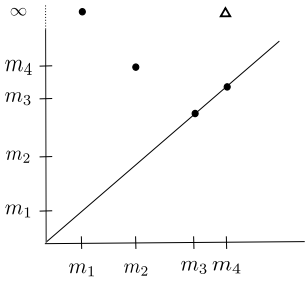
(d) BC



(e) ABC (left: positive for β_k, right: negative for β_k. Green: β_1. Black: β_0.)



(f) PD



(g) APD

Figure 1.1: The six descriptors of interest generated by the direction indicated by the arrow for the simplicial complex at the top right of the figure. Both $\hat{\chi}$ and $\hat{\beta}$ are shown as two plots, the projections onto the first and second coordinates.

done. Otherwise, f_t and f_{t+1} differ by one transposition. Thus, there exist $\sigma_t, \sigma_t' \in K$, where $f_t(\sigma_t) = f_{t+1}(\sigma_t) - 1$ and $f_t(\sigma_t') = f_t(\sigma_t') + 1$, and $f_t(\sigma) = f_{t+1}(\sigma)$ for all other $\sigma \in K$. Since f_t is compatible with f, we know that $f(\sigma) \leq f(\sigma')$. Since f_{t+1} is also compatible with f, we know that $f(\sigma') \leq f(\sigma)$. Hence, $f(\sigma) = f(\sigma')$.

Note that the persistence diagrams $\rho_k(f_t)$ and $\rho_k(f_{t+1})$ induce matchings on K, i.e., the birth-death pair $(i, j]$ induces the match (σ_i, σ_j). By [17, p. 122], there are three cases where the matchings can change: (1) two birth events swap; (2) two death events swap; (3) a birth event and a death event swap. In all three cases, the pairing swaps correspond to σ and σ'. Since $f(\sigma) = f(\sigma')$, we have $\mathcal{A}_{t+1} = \mathcal{A}_t$. □

Since Betti number and Euler characteristic can be determined from knowing all persistence pairs, we have the following corollary.

Corollary 1.15. *The augmented Euler characteristic curve and the augmented Betti curves are well-defined.*

1.16 Faithful Sets and Topological Transforms

In Chapters 3 and 4, we are concerned with sets of topological descriptors generated from lower-star filtrations with respect to different directions over a particular simplicial complex. A desirable property for such sets of descriptors to have is *faithfulness*.

Definition 1.17 (Faithful Set)**.** *Let K be a geometric simplicial complex. Let $\mathrm{D}(K, S) = \{(\mathrm{D}(K, s), s)\}_{s \in S}$, be an indexed set of topological descriptors of type D. We say that $\mathrm{D}(K, S)$* is faithful *if, for any K' we have $\mathrm{D}(K', S) = \mathrm{D}(K, S)$ if and only if $K' = K$.*

We often also refer to a set of descriptors as faithful for K_0, which is similarly defined. That is, the set $\mathrm{D}(K, S)$ *is* faithful for K_0 *if, for any simplicial complex K', we have* $\mathrm{D}(K', S) = \mathrm{D}(K, S)$ *if and only if* $K'_0 = K_0$.

There are many ways to see if a set is faithful; the following lemma provides one way.

Lemma 1.18 (Sufficient and Neccessary Condition for Faithfulness)**.** *Let K be a simplicial complex,* D *a topological descriptor type, and P a set of parameters indexing the set* $\mathrm{D}(K, P)$. *Then the set* $\mathrm{D}(K, P)$ *is faithful if and only if*

$$\bigcap_{p \in P} \left\{ K' \in \mathbb{R}^d \mid \mathrm{D}(K', p) = \mathrm{D}(K, p) \right\} = \{K\}.$$

Proof. First, suppose that $\mathrm{D}(K, P)$ is faithful. It is immediate that the intersection contains at least K. Suppose, towards a contradiction, that the intersection contains some $K' \neq K$. Then, for all $p \in P$, we have $\mathrm{D}(K', p) = \mathrm{D}(K, p)$, meaning that $\mathrm{D}(K', P) = \mathrm{D}(K, P)$, contradicting the faithfulness of $\mathrm{D}(K, P)$.

Next, suppose that the intersection only contains K. This means there is no $K' \neq K$ such that $\mathrm{D}(K', p) = \mathrm{D}(K, p)$ for all $p \in P$, i.e., there is no $K' \neq K$ such that $\mathrm{D}(K', P) = \mathrm{D}(K, P)$, so we see that $\mathrm{D}(K, P)$ is faithful. $\square$

Finally, when considering a set of descriptors corresponding to lower-star filtrations such that the set is parameterized by the entire sphere of directions, we refer to the set as a *directional topological transform.*

Definition 1.19 (Directional Topological Transform)**.** *Let* $K \subset \mathbb{R}^d$ *be a simplicial complex and* D *be a type of topological descriptor corresponding to lower-star filtrations. Then* $\mathrm{D}(K, \mathbb{S}^{d-1})$ *is the* directional topological transform associated to D and K.

In particular, we are interested in the *(augmented) persistent homology transform ((A)PHT)* where $\mathrm{D} = \rho$ or $\hat{\rho}$, the *(augmented) Betti curve transform ((A)BCT)* where $\mathrm{D} = \beta$ or $\hat{\beta}$, and the *(augmented) Euler characteristic curve transform ((A)ECCT)* where $\mathrm{D} = \chi$ or $\hat{\chi}$. We may use the shorthand notation $APHT(K)$ to mean $\hat{\rho}(K, \mathbb{S}^{d-1})$, and similarly for the other transforms of interest.

We define directional topological transforms in terms of lower-star filtrations; note that other authors may use, e.g., height filtrations instead [3,20,34]. This distinction is nontrivial; for instance, in theorem 1.2 of [3], the authors discuss an instance when height filtration-based PHT can be determined solely from homology in degree zero. Notice that this will never be true for the lower-star filtration-based PHT; consider the complex consisting of just two-simplex versus the complex consisting of the boundary of a two-simplex, which have equivalent degree zero information in all directions. For our current purposes, the distinction between height filtration-based transforms and lower-star filtration-based transforms is not so important, as we take information in all degrees; note that a lower-star filtration of a simplicial complex is simply a (more computable) coarsening of a height filtration.

CHAPTER TWO

MOTIVATING QUESTIONS AND MAIN RESULTS

In this chapter, we summarize the motivating questions and main results of each chapter. Through doing so, we aim to provide an overview of the content as well as highlight connections between chapters. We end each section with a grateful acknowledgement of collaborators, reviewers, and other contributors who made this work possible.

Chapter 3 Our study of topological transforms began when reading and discussing [61], where the PHT and ECCT were first introduced and explored. In [61], faithfulness of the entire transforms are discussed (i.e., faithfulness of the uncountably infinite set parameterized by the entire sphere of directions), which is infeasible to use in computational applications. Applications require a finite set of directions, which in [61], were chosen somewhat arbitrarily, and were not proven to be a faithful set.

Motivated by these observations and with the goal of connecting applications to provably correct direction-choosing techniques, we asked the following question.

Question 2.1. *Given a finite geometric simplicial complex $K \in \mathbb{R}^d$, can we find an explicit faithful set of $\hat{\rho}$'s, $\hat{\beta}$'s, or $\hat{\chi}$'s that is finite? That is, can we faithfully discretize the APHT, ABCT, or AECCT?*

We answer this question in the affirmative, first for the APHT, using directions that satisfy certain properties in relation to the underlying simplicial complex.

Theorem 3.12 (Sufficient Conditions for Representation)**.** *Let K be a simplical complex GP-immersed in $\mathbb{R}^d$ such that $\dim(K) = \kappa < d$, and let $S \subset \mathbb{S}^{d-1}$ such that (K, S) is vertex- and simplex-isolating. Then, $\hat{\rho}(K, S)$ is a faithful discretization of $APHT(K)$.*

Furthermore, we provide algorithms to find explicit directions satisfying these conditions, given an input simplicial complex (summarized in Construction 1).

Surprisingly, our results relied only on knowing the dimension of events throughout a filtration, and did not require any birth-death pairing information available in $\hat{\rho}$'s. Thus, we have an immediate corollary that vertex- and simplex-isolating directions correspond to faithful discretizations of not only the ABCT, but for any topological transform of augmented dimension-returning transforms (Corollary 3.29).

Since we cannot tell the dimension of an event (i.e., what dimension(s) of homology changed) just from a single $\hat{\chi}$, showing that a set of vertex- and simplex-isolating directions also corresponds to a faithful discretization of the AECCT is less straightforward. In Section 3.29.1, we show how the algorithms for $\hat{\rho}$'s and $\hat{\beta}$'s can be adapted for $\hat{\chi}$'s, and are able to conclude that such a set faithfully discretizes the AECCT (Theorem 3.34).

For the setting of general simplicial complexes, our goal is to provide answers to Question 2.1 using the tool of reconstruction, not to necessarily reconstruct complexes in algorithmically efficient ways. Having answers to Question 2.1, we ask a related question.

Question 2.2. *Can we find a more efficient algorithm for simplicial complex reconstruction?*

We observe that edges adjacent to a common vertex can be radially ordered through projection onto $\mathbb{R}^2$. This allows us to perform a radial binary search for edges at each vertex. Incorporating the radial binary search into a sweepline algorithm that finds all edges above each vertex, we have the following result.

Theorem 3.50 (Graph Reconstruction)**.** *Let G be an unknown graph immersed in $\mathbb{R}^d$. If we use a vertex reconstruction method that requires $\Theta(\mathbb{V}_{time})$ time and $\Theta(\mathbb{V}_{dgm})$ diagrams, we can reconstruct G using $\Theta(\mathbb{V}_{dgm} + d + m \log n)$ diagrams in $\Theta(\mathbb{V}_{time} + n^2 + m \log n(\log n + d + \Pi))$ time.*

We are able to use this search method on one-skeletons of general simplicial complexes (Corollary 3.51). Since higher-dimensional simplices do not have a well-defined notion of radial ordering, we are not able to directly extend the radial binary search technique. Finding an answer to Question 2.2 for general simplicial complexes remains as ongoing work.

Chapter 3 is an extension of [29] (Section 3.8.1) and [30] (Section 3.41), which are collaborations with Brittany Fasy, David Millman, Samuel Micka, and Lucia Williams. This work began in [8], which was joint work with the aforementioned collaborators, along with Robin Belton, Rostik Mertz, Daniel Salinas, and Jordan Schupach. Additionally, we would like to thank the CompTaG club at Montana State University and the many reviewers for their thoughtful feedback on this work.

Chapter 4 Since Euler characteristic is a much coarser invariant than Betti numbers, which is coarser than homology, it makes intuitive sense that more Euler characteristic curves might be necessary to form a faithful set with minimum cardinality than, e.g., persistence diagrams.

We also observe a stark difference in the ability of non-augmented Euler characteristic curves to represent a shape when compared to augmented persistence diagrams; they feel vastly *weaker* than augmented persistence diagrams. These sorts of observations lead us to the following motivating question.

Question 2.3. *What is a reasonable definition of an ordering relation for general topological descriptors types?*

We define our ordering relations on equivalence classes of topological descriptor types. We define what we mean when we say two descriptor types are in the same equivalence class (Definition 4.6), when one equivalence class is weaker than another (Definition 4.8), and show that these notions are well defined (Lemma 4.7 and Lemma 4.9, respectively).

Although the relations apply to general classes of topological descriptors, we are particularly interested in topological descriptors that are commonly used in practice. Thus, we focus on six common types of descriptors.

Question 2.4. *What relations are there among elements of* $\{[\chi], [\beta], [\rho], [\hat{\chi}], [\hat{\beta}], [\hat{\rho}]\}$*? What (if anything) can we say about minimal faithful sets of these types of descriptors?*

Through a combination of reduction arguments and constructions, we find the following poset, which represents progress towards the first half of Question 2.4.

Theorem 4.21 (Partial Ordering)**.** *The following diagram is a correct partial order on strength classes of topological descriptors.*

$$\begin{tikzcd}
{[\hat{\chi}]} \arrow[r, "\leq"] & {[\hat{\beta}]} \arrow[r, "<"] & {[\hat{\rho}]} \\
{[\chi]} \arrow[u, "<"] \arrow[r, "\leq"] & {[\beta]} \arrow[u, "<"] \arrow[r, "\leq"] & {[\rho]} \arrow[u, "<"]
\end{tikzcd}$$

In Section 4.22, we make progress towards the second half of Question 2.4. In particular, we identify specific simplicial complexes for which the cardinality of minimum faithful sets is large. For non-augmented descriptors, the bound is not surprising; we know that if a set of descriptors is faithful for a complex, each maximal k-simplex requires $d-k+1$ perpendicular directions in the corresponding set of directions (Corollary 4.25). The main challenge of the following result is keeping track of groups of complexes that are all perpendicular to a common hyperplane, thus allowing a single direction to "count" as a perpendicular direction to multiple simplices at once.

The construction of a simplicial complex in $\mathbb{R}^d$ for $d > 2$ consisting of $n_1 > d-1$ maximal edges arranged so that no j lie on the same $(j+1)$-plane gives the following result.

Theorem 4.27 (Lower Bound for Worst-Case Non-Augmented Descriptor Complexity)**.** *Let* $\mathrm{D} \in \{\chi, \beta, \rho\}$*. Then the worst-case cardinality of a minimal descriptor set is* $\Omega(d+n_1)$*.*

We have a similar result for augmented descriptors, again through a construction of a family of simplicial complexes comprised entirely of maximal edges, this time in $\mathbb{R}^2$.

Theorem 4.32 (Lower Bound for Worst-Case Augmented Descriptor Complexity)**.** *Let* $\hat{\mathrm{D}} \in \{\hat{\chi}, \hat{\beta}, \hat{\rho}\}$*. Then the worst-case cardinality of a minimal descriptor set is* $\Omega(n_0)$*.*

Theorem 4.32 is surprising. Since augmented descriptors have an event at every vertex, one might expect the minimum cardinality of a faithful set of augmented descriptors to be independent of the size of the complex (and in many cases, this is true). However, we were able to prove our constructed family of simplicial complexes requires any faithful set

to include descriptors corresponding to a very specific set of directions, whose size depends on the number of vertices of the complex.

Chapter 4 is an ongoing collaboration with Brittany Fasy, David Millman, and Samuel Micka. We began asking the types of questions that led to this work in [27], which was joint work with the aforementioned collaborators as well as Lucia Williams. We additionally thank Ryan Grady for many helpful discussions.

Chapter 5 We have so far focused on one-parameter monotone filtrations, which correspond to one-parameter monotone persistence modules. In Chapter 5, we broaden our view to multi-parameter zig-zag modules.

In order to use tools from category theory, we would like to view persistence modules from a number of perspectives, not just as functors from posets to abelian categories. In particular, we are interested in viewing persistence modules as constructible cosheaves over their parameter space, stratified by event times and relations between event times.

Considering persistence modules in sheaf/cosheaf theoretic terms is not novel. In fact, the topological transforms of Chapter 3 are a type of persistence module themselves, and it is increasingly popular to view topological transforms as sheaves/cosheaves over the sphere of directions, stratified by regions of directions that see the height of vertices of a simplicial complex in the same order [3, 31, 39].

However, to our knowledge, a precise definition of persistence module cosheaves, specifically persistence modules whose parameter space is Euclidean (rather than a sphere), is not available in the literature. We are also interested in creating a definition that allows for augmented structure. Thus, a first question is as follows.

Question 2.5. *How can we explicitly define multi-parameter persistence modules as constructible cosheaves over $\mathbb{R}^d$? Can we make our definition in a way that allows for a notion of augmented persistence modules?*

We first describe a cosheaf structure for multi-parameter modules in Construction 4, which we call the *propagated persistence cosheaf*, since the costalks over zero-strata propagate to costalks over neighboring points in the directions of the underlying poset. This reflects the notion that features in persistent homology appear and then remain until some later event (where "later" is dictated by the poset structure, which may not be a linear order).

To extend this cosheaf construction to augmented persistence modules, we first focus on the one-parameter case. In Definition 5.33, we carefully define a map of stratified spaces, C, that sends the stratified space corresponding to an index filtration to the stratified space corresponding to a monotone filtration, in such a way that we can relate their corresponding persistence cosheaves via pushforward.

Then, in Definition 5.36, we define an augmented persistence cosheaf, using information about instantaneous events available to us from the map C, so that instantaneous events are recorded in open sets containing zero-strata, but not in open sets that don't contain zero-strata (this reflects the notion that instantaneous events appear and disappear at a single parameter value, but do not impact parameter values between events). We also note that the information about instantaneous events can also be defined via boundary and cycle information of the monotone module alone (Lemma 5.35).

We then broaden our attention to defining augmented persistence cosheaves on multi-parameter modules. We note that cosheaves valued in Vect may not be sufficient for a direct adaptation of our one-parameter cosheaf, since a zero-strata in the multi-parameter setting may correspond to an instantaneous event in one parameter but not another. We therefore propose an extension to grVect that keeps track of each parameter in its own level.

Now that we have a definition of persistence modules as constructible cosheaves over a stratified parameter space, we would like to make the connection to other notions of persistence modules explicit.

Theorem 5.42. *The category of (marked) multi-parameter zig-zag modules localized at weak equivalences is equivalent to the category of constructible cosheaves on $\mathbb{R}^d$ stratified by the natural numbers. That is, we have an equivalence of categories*

$$\mathsf{ZZmod}[\mathcal{W}^{-1}] \cong \mathsf{Fun}(\mathsf{Ent}_{\Delta}(\mathbb{R}^d;\mathbb{N}^d), \mathsf{Vect}) \simeq \mathsf{cShv}^{\mathsf{Vect}}_{\mathsf{cbl}}((\mathbb{R}^{\mathsf{d}};\mathbb{N}^{\mathsf{d}})).$$

Persistence modules viewed as a category are a good candidate for K-theory computations, in particular since they are an example of a *Waldhausen category*, whose additional structure lets us use tools like Waldhausen additivity. This leads us to the following question.

Question 2.6. *What is the algebraic K-theory of the category of multi-parameter (zig-zag) persistence modules?*

We first approach Question 2.6 for the restricted setting of one-parameter zig-zag persistence modules. Using Waldhausen additivity, we find that the K-theory of one-

parameter persistence modules is additive over strata; that is, we can express the K-theory of the whole persistence module as the K-theory of persistence modules valued over each zero-strata (corresponding to events in the filtration) and valued over each one-strata (corresponding to parameter values between events in the filtration). We show this result through induction over the number of zero-strata (Theorem 5.56).

We then move on to answering Question 2.6 in the general case, for multi-parameter modules. Using Theorem 5.56 as our base case, we induct on the number of parameters as well as the *height* of the parameter space, which is a measure of the maximum number of zero-strata in any single parameter direction (induction on height is analogous to induction on number of zero-strata in the one-parameter case).

Each part of the induction is a geometric argument. First, we show that a parameter space with height $h+1$ can be "sliced" into pieces, all of which have height no more than h. Then, given a $(d+1)$-parameter space with height two (i.e., a cube), we use a result in computational geometry to show the space can be broken apart into lower-parameter spaces. In both the induction on height and on parameters, we again use Waldhausen additivity to "glue" the pieces of the division back together. Finally, we have the following main result.

Theorem 5.59 (K-Theory of Multi-Parameter Zig-Zag Modules)**.** *Let X be a cubical grid d-manifold with a finite number of strata. There is an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \bigvee_{x_0 \in X_0} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_0)) \vee \bigvee_{x_1 \in X_1} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_1)) \vee \ldots \vee \bigvee_{x_d \in X_d} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_d))$$

where X_i is the set of i-strata of X.

Chapter 5 is an extension of [37] and is the result of collaboration with Ryan Grady. We thank Peter Bubenik for several discussions related to zig-zag persistence and other mathematical topics in TDA. Finally, we thank the anonymous referee for feedback and suggestions, which greatly enhanced the readability of [37], and therefore also Chapter 5.

CHAPTER THREE

FAITHFUL DISCRETIZATIONS OF TOPOLOGICAL TRANSFORMS

3.1 A Faithful Discretization of the Augmented Persistent Homology Transform

3.1.1 Background Definitions

Here, we give an overview of relevant background information and definitions, following the notation established in [7, 8]. More foundational definitions are given in Chapter 1. In this section, we take a generous assumption of general position.

Assumption 1. (General Position) *In what follows, we say a point set $V \subset \mathbb{R}^d$ is in* general position *if, for all k with $1 \leq k \leq d+1$, every subset of V of size k is affinely independent.*

Note that Assumption 3 is almost always met. A *GP-immersed simplicial complex K* is a finite set of geometric simplices such that the vertex set of K is in general position. Note that this is a stronger condition than the usual notions of immersions. For instance, the intersection of an edge's endpoint lying in the interior of another edge is not allowed, since this would cause three vertices to be colinear. However, two edges could intersect transversely in $\mathbb{R}^2$ and would still satisfy Assumption 3.

In this chapter, we are specifically interested in persistence diagrams for *lower-star filtrations* of a simplicial complex K that is GP-immersed in $\mathbb{R}^d$ with respect to a direction $s \in \mathbb{S}^{d-1}$. As discussed in Chapter 1, a filtration gives rise to an *(augmented) persistence diagram*, notated $\hat{\rho}$. Since we only use $\hat{\rho}$'s in this section, we use "diagram" as

shorthand for "directional $\hat{\rho}$." Also recall that the set of diagrams parameterized by the entire sphere of directions is called the *(augmented) persistent homology transform ((A)PHT)*.

The foundational paper by Turner et al. [61] showed that the PHT and ECCT (*Euler characteristic curve transform*) are *faithful*. That is, the data of $PHT(K)$ or $ECC(K)$ is sufficient to know which simplicial complex K generated the sets of descriptors. Topological transforms are parameterized by the sphere of directions; we aim to find a finite and discrete subset of the sphere of directions that retains the property of having topological descriptors that are faithful for a particular simplicial complex. In this section, we focus on faithfully discretizing the APHT; a formal definition of this notion is given as follows.

Definition 3.2 (Faithful Discretization of $APHT(K)$)**.** *Given a simplicial complex K GP-immersed in $\mathbb{R}^d$ and a finite set $S \subset \mathbb{S}^{d-1}$, we say that $\hat{\rho}(K,S) := \{(s, \hat{\rho}(h_s))\}_{s \in S} \subset APHT(K)$ is a* faithful discretization of $APHT(K)$ *if, for any other simplicial complex $K' \neq K$ GP-immersed in $\mathbb{R}^d$, there exists an $s_0 \in S$ such that $\hat{\rho}(h_{s_0}) \neq \hat{\rho}(h'_{s_0})$, where $h'_{s_0} \colon K' \to \mathbb{R}$ is the lower-star filter function of K' with respect to direction s_0. That is, no other simplicial complex GP-immersed in $\mathbb{R}^d$ has the same parameterized set of directional diagrams.*

3.2.1 Tools for Building a Faithful Discretization

We now give a lemma that relates simplices to points in a $\hat{\rho}$, discuss the general position assumption, and define a tool used in our proofs of faithfulness called *filtration hyperplanes*.

Lemma 3.3 (Simplex Count)**.** *Let K be a simplicial complex. Let $k \in \mathbb{N}$ and $c \in \mathbb{R}$. Let $f \colon K \to \mathbb{R}$ be a monotonic function. Then, the k-dimensional simplices of K with a*

function value of c are in one-to-one correspondence with the points in the following multiset:

$$\{(a,b) \in \hat{\rho}_k(f) \ \ s.t. \ \ a = c\} \cup \{(a,b) \in \hat{\rho}_{k-1}(f) \ \ s.t. \ \ b = c\}\,. \tag{3.1}$$

Proof. Let $f'\colon K \to \mathbb{N}$ be an index filter function compatible with f, and let $\{L'_i\}_{i\in\mathbb{N}}$ be the corresponding index filtration, where $L'_i := f'^{-1}(-\infty, i]$. Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be the ordering of simplices in K such that $f'(\sigma_i) = i$. Consider the sets

$$C_B := \{(i,j) \in \rho_k(f') \text{ s.t. } f(\sigma_i) = c\} \quad C_D := \{(i,j) \in \rho_{k-1}(f') \text{ s.t. } f(\sigma_j) = c\}$$

and let $C = C_B \cup C_D$.

We start by defining a bijection $\phi\colon f^{-1}(c) \to C$. Let $\sigma_i \in K_k$ such that $f(\sigma_i) = c$. Since adding a single k-simplex either increases β_k by one or decreases β_{k-1} (see, e.g., [25, pp. 120–121]), either $\beta_k(L'_i) = \beta_k(L'_{i-1}) + 1$ or $\beta_{k-1}(L'_i) = \beta_{k-1}(L'_{i-1}) - 1$, but not both.

Case 1: (β_k increases). There exists a unique birth at index i in $\rho_k(f')$. Thus, let $j \in \bar{\mathbb{R}}$ be such that $(i,j) \in \rho_k(f')$. Then, $(f'(\sigma_i), f'(\sigma_j)) = (i,j) \in C_B$. Define $\phi(\sigma_i) = (i,j)$.

Case 2: (β_{k-1} decreases). There is a unique death at index i in $\rho_k(f')$. Thus, let $j \in \bar{\mathbb{R}}$ be such that $(j,i) \in \rho_k(f')$. Then, $(f'(\sigma_j), f'(\sigma_i)) = (j,i) \in C_D$. Define $\phi(\sigma_i) = ((f'(\sigma_j), f'(\sigma_i)) = (j,i)$.

In other words, each σ_i is mapped to the persistence pair containing $f'(\sigma_i)$. If $\phi(\sigma) = \phi(\tau) = (i,j)$, then, by construction, both σ and τ are the same type of event. WLOG, assume that

they are birth events. Then, $f'(\sigma) = f'(\tau)$, which, by injectivity of f' means that $\sigma = \tau$. Thus, we have shown that ϕ is an injection.

We show that ϕ is a surjection by contradiction. Suppose there exists $(i, j) \in C$ such that there does not exist $\sigma \in f^{-1}(c)$ with $\phi(\sigma) = (i, j)$. Since $(i, j) \in C$, then either $f(\sigma_i) = c$ or $f(\sigma_j) = c$ (or both). WLOG, suppose $f(\sigma_i) = c$. Then, $\sigma_i \in f^{-1}(c)$ and, by construction, we have $\phi(\sigma_i) = (i, j) \in C_B$.

By Equation (1.2) and since $\hat{\rho}(f)$ is well-defined (see Lemma 4.9), we know that the map $\psi\colon \rho_k(f') \to \hat{\rho}_k(f)$ defined by the assignment $\psi(i, j) = (f(\sigma_i), f(\sigma_j))$ is a bijection. Finally, since $C \subset \rho_k(f')$, and since both functions are bijections, the composition $\psi \circ \phi$ is exactly the one-to-one correspondence that we sought. $\square$

Next, we define a structure that helps build a geometric intuition for several of the proofs that follow.

Definition 3.4 (Filtration Hyperplane)**.** *Let $s \in \mathbb{S}^{d-1}$ be a unit vector, and let $c \in \mathbb{R}$. Let $H(s, c)$ be the $(d-1)$-dimensional hyperplane that passes through the point $cs \in \mathbb{R}^d$ and is perpendicular to s. We define the closed half-spaces above and below this hyperplane with respect to direction s by $H^{\uparrow}(s, c)$ and $H^{\downarrow}(s, c)$, respectively.*

Let V be a finite set of vertices in $\mathbb{R}^d$ and let $h_s\colon V \to \mathbb{R}$ be the lower-star filter function for direction s. The filtration hyperplanes of V *are the set of hyperplanes*

$$\mathbb{H}(s, V) := \{H(s, h_s(v))\}_{v \in V}.$$

All hyperplanes in $\mathbb{H}(s, V)$ are parallel to each other and perpendicular to the direction s. Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Since the births in $\hat{\rho}_0(h_s)$ are in one-to-one correspondence with the vertices of K by Lemma 3.3, there is a filtration hyperplane at every height at which a vertex lies in direction s. In directions where vertices are at the same height, the filtration hyperplanes are not unique.

By observing intersections of a sufficient number of linearly independent filtration hyperplanes, we can form a grid of points of intersections, which will be a crucial tool in our vertex-reconstruction arguments.

Definition 3.5 (Filtration Grid)**.** *For $n \geq d$, let $s_1, s_2, \ldots s_n \in \mathbb{S}^{d-1}$ be linearly independent and let $P \in \mathbb{R}^d$ be a pointset. We define the* filtration grid of P with respect to $\{s_1, s_2, \ldots, s_n\}$ *to be the grid of points, A, arising from choosing one hyperplane in each set $\mathbb{H}(s_i, K_0)$ for each $1 \leq i \leq n$. That is, the filtration grid is the collection of n-way intersections of filtration hyperplanes. Note that $P \subseteq A$ and $|A| \leq |P|^d$.*

See Figure 3.2 for a figure illustrating filtration hyperplanes and a filtration grid. Finally, we define a tool that is useful in stating conditions and results succinctly.

Definition 3.6 (Wedge)**.** *Let $P \subset \mathbb{R}^d$ be a finite set of points and aff$(P) \neq \mathbb{R}^d$. Let s_1, s_2 be two directions in $\mathbb{S}^{d-1}$ that are orthogonal to aff(P) and $s_1 \neq s_2$. Let h_1, h_2 be the height of P with respect to s_1 and s_2, respectively. The* wedge *between s_1 and s_2 pivoted about P is the closure of the symmetric difference between $H^{\downarrow}(s_1, h_1)$ and $H^{\downarrow}(s_2, h_2)$, which we denote as $H^{\downarrow}(s_1, h_1) \triangle H^{\downarrow}(s_2, h_2)$.*

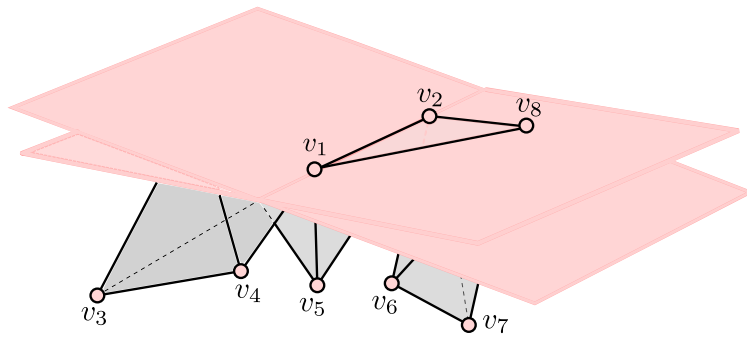


Figure 3.1: A wedge (shaded in pink) pivoting about the one-simplex $[v_1, v_2]$. To test if $[v_1, v_2, v_8] \in K$, we notice that the two-indegrees for $[v_1, v_2]$ in the two directions defining the wedge do not differ, indicating that v_1, v_2, and v_8 do not define a two-simplex in K by Theorem 3.17.

Figure 3.1 illustrates a wedge pivoted around the vertex set $\{v_1, v_2\}$. The next definition describes specific directions that are perpendicular to a simplex.

Definition 3.7 (P-Perpendicular)**.** *Let $V \subset P \subset \mathbb{R}^d$ such that P is in general position. We say that a direction $s \in \mathbb{S}^{d-1}$ is* P-perpendicular *to V if for all $u \in P$ and $v \in V$, $s \cdot u = s \cdot v$ if and only if $u \in V$. In other words, s is perpendicular to aff(V) and no other vertex in P is at the same height as the vertices in V.*

Next, we define when a direction is a slight tilt of an initial direction that is P-perpendicular to some V, so that a subset W (and no other points) pops above $V \setminus W$.

Definition 3.8 ((P, V, W, s)-Perturbation)**.** *Let $W \subset V \subset P \subset \mathbb{R}^d$ such that P is in general position. Let $s \in \mathbb{S}^{d-1}$ be P-perpendicular to V. Then, we say a direction $s_* \in \mathbb{S}^{d-1}$ is a* (P, V, W, s)-perturbation *if the following statements hold:*

(i) The direction s_* is P-perpendicular to aff$(V \setminus W)$.

(ii) The points in W are above $V \setminus W$ with respect to s_*.

(iii) For all $p \in P \setminus V$, p is strictly above (below) the height of $V \setminus W$ with respect to s_* if and only if it is strictly above (below, respectively) V with respect to s.

If V defines the vertices of a simplex σ, then we may use σ in place of verts(σ) in the definition and construction above. Likewise for W.

3.8.1 The Main Result: A Faithful Discretization of the APHT

In this section, we give sufficient properties for a set of diagrams to faithfully discretize the APHT of a given simplicial complex, the proof of which is given in the remainder of the section. To give a concrete construction, we also provide an explicit set of directions that satisfies these properties.

Definition 3.9 (Vertex-Isolating)**.** *We say that the pair (K, S) is* vertex-isolating *if $\hat{\rho}(K, S)$ has d linearly-independent directions, denoted $\{s_1, s_2, \ldots, s_d\}$, and one additional direction that uniquely orders the the filtration grid of K_0 with respect to $\{s_1, s_2, \ldots, s_d\}$. We may also say that S is* vertex-isolating with respect to K.

Definition 3.10 (Simplex-Isolating)**.** *Let $\sigma \in K \setminus K_0$. We say that the pair (K, S) is* σ-isolating *if S includes a direction s_σ such that:*

(a) s_σ is K_0-perpendicular to σ, and

(b) for each $\emptyset \neq W \subsetneq V =$ verts(σ), S includes a direction that is a (K_0, V, W, s_σ)-perturbation.

If the pair (K, S) is σ-isolating for every simplex $\sigma \in K$, we say (K, S) is simplex-isolating.

Note that (K, S) may only be simplex-isolating if $\dim(K) < d$; that is, K must have positive codimension with its ambient space. Figure 3.1 illustrates directions that are $[v_1, v_2, v_8]$-isolating, where $V = \{v_1, v_2, v_8\}$ and $W = \{v_8\}$. The two filtration hyperplanes through $V \setminus W$ with respect to the indicated directions (a direction s_V that is K_0-perpendicular to V and a direction that is a (K, V, W, s_V)-perturbation) form a "wedge" containing the vertices of W.

Next, we make the observation that if there are directions are σ-isolating for all maximal simplices of a complex, then the set of directions is simplex-isolating. Namely, there are directions that are σ-isolating for *all* simplices of a complex, regardless of maximality.

Lemma 3.11 (Recursive Nature of σ-Isolating Directions)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$, and let S be a set of directions so (K, S) is σ-isolating for every maximal simplex σ. Then, (K, S) is τ-isolating for every simplex $\tau \in K$.*

Proof. Let $\tau \in K$. Then, either τ is a maximal simplex or τ is a proper face of a maximal simplex. If τ is maximal, the claim follows immediately by assumption. Suppose, then, that τ is a proper face of a maximal simplex τ_m. Denote the vertex sets of τ and τ_m by T and T_m, respectively. Since S satisfies Statement (a) of Definition 3.10,there exists a direction $s_m \in S$ that is K_0-perpendicular to T_m. Then, since S also satisfies Statement (b), in particular when $V = T_m$ and $W = T_m \setminus T$, there is a direction $s_\tau \in S$ that is a $(K_0, T_m, T_m \setminus T, s_m)$-perturbation; this direction is K_0-perpendicular to $T_m \setminus (T_m \setminus T) = T$, as desired.

If τ is a vertex, the second part of the claim is vaccuously true. We therefore proceed assuming there exists some $\tau' \prec \tau$ and let $T' = \text{verts}(\tau')$. Since $T_m \setminus (T \setminus T') \subset T_m$, and

since S satisfies Statement (b) of Definition 3.10, there exists a direction $s_p \in S$ that is a $(K_0, T_m, T_m \setminus (T \setminus T'), s_m)$-perturbation.

We show that s_p satisfies the three properties of Definition 3.8. By definition, s_p is K_0-perpendicular to $T_m \setminus (T_m \setminus (T \setminus T') = T \setminus T'$. Hence, the direction s_p satisfies Statement (i) of Definition 3.8. Also by definition, the points of $T_m \setminus (T \setminus T')$ are above $T \setminus T'$ with respect to s_p. Since $T' \subseteq T_m \setminus (T \setminus T')$, we conclude that the points in T' are above $T \setminus T'$ with respect to s_p. Hence, s_p satisfies Statement (ii) of Definition 3.8. Finally, let $p \in K_0 \setminus T$. If $p \in K_0 \setminus T_m$, then p is strictly above (below) the height of $\tau \setminus \tau'$ with respect to s_p if and only if it is strictly above (below) the height of $\tau \setminus \tau'$ with repsect to s_m by defintion. Furthermore, this means p is above (below, respectively) $\tau \setminus \tau'$ if and only if it is above (below) with respect to s_τ. If, instead, $p \in K_0 \setminus (T_m \setminus T)$, then $p \in \text{verts}(\sigma) \setminus (T \setminus T')$, so p is necessarily above $\tau \setminus \tau'$ with respect to s_τ and s_p. Hence, s_p satisfies Statement (iii) of Definition 3.8. Thus, the direction s_p is a $(K_0, \tau, \tau', s_\tau)$-perturbation. □

The remainder of this section focuses on proving the following theorem:

Theorem 3.12 (Sufficient Conditions for Representation)**.** *Let K be a simplical complex GP-immersed in $\mathbb{R}^d$ such that $\dim(K) = \kappa < d$, and let $S \subset \mathbb{S}^{d-1}$ such that (K, S) is vertex- and simplex-isolating. Then, $\hat{\rho}(K, S)$ is a faithful discretization of $APHT(K)$.*

In Section 3.23.1, we present algorithms that compute an explicit set of directions that is vertex- and simplex-isolating with respect to a simplicial complex K. That is, we arrive at an explicit faithful discretization of the APHT. However, before computing explicit directions,

we first prove Theorem 3.12 for general sets of directions satisfying the conditions of being vertex- and simplex-isolating for a given simplicial complex.

3.12.1 Reconstruction Algorithm for Simplicial Complexes in $\mathbb{R}^d$

In the following section, we prove Theorem 3.12 (Sufficient Conditions for Representation) by reconstructing a GP-immersed simplicial complex. Our method first finds all zero-simplices, then all one-simplices, and so on. In what follows, let K be a simplicial complex GP-immersed in $\mathbb{R}^d$ and let S be a set of directions that is vertex- and simplex-isolating for K. We use $\hat{\rho}(K, S)$ to reconstruct K.

3.12.2 Vertex Reconstruction

Our method for finding zero-simplices is a straightforward generalization of the method of [8]. We briefly state the result here.

Lemma 3.13 (Vertex Reconstruction, [8, Theorem 9])**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Then, given a set of directions S that is vertex-isolating, $\hat{\rho}(K, S)$ can reconstruct K_0.*

We refer the reader to [8] for full details, but provide a proof sketch here. Since S is vertex-isolating, it contains a set of linearly-independent directions $\{s_1, s_2 \ldots, s_d\}$. Let A denote the filtration grid of K_0 with respect to $\{s_1, s_2, \ldots, s_d\}$. Since S is vertex-isolating, it also contains some direction s_{d+1} that uniquely orders points of A, leading to exactly n_0 pairwise intersections with $\mathbb{H}(s_{d+1}, K_0)$ and A. It is then a simple matter of checking for the locations of these intersections, for example, with a brute force algorithm. See Figure 3.2 for an example of a vertex-isolating set and how such a set may be used for vertex reconstruction.

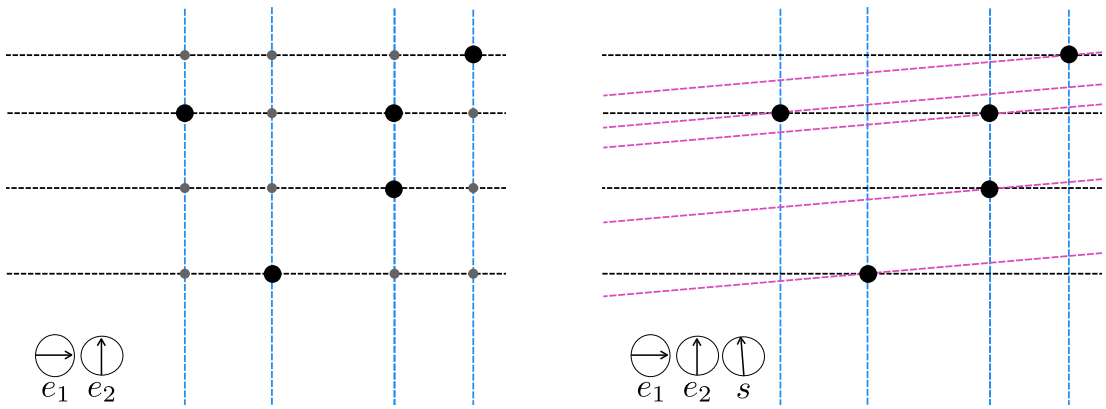


Figure 3.2: The vertex set P (large black points) defines the filtration grid of P with respect to $\{e_1, e_2\}$, denoted A (grey and black dots in the left figure). The direction s, indicated on the right, uniquely orders the points of A. Thus, since e_1 and e_2 are linearly independent and since the direction s uniquely orders the points of the A, the set $\{e_1, e_2, s\}$ is vertex-isolating for the given vertex set. To locate the vertices of the set, we simply need to identify all intersections of $\mathbb{H}(s, P)$ (diagonal pink dashed lines) with A.

3.13.1 Higher-Dimensional Simplex Reconstruction

Assuming that K_0 is known, we reconstruct the simplices of the simplicial complex K.

3.13.1.1 Computing k-Indegree

The key to determining whether a simplex exists is the *k-indegree* of a simplex, the count of k-dimensional cofaces of a simplex σ at the same height as σ in a particular direction.

Definition 3.14 (k-Indegree for Simplex)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$ and $\sigma \in K$. Let $s \in \mathbb{S}^{d-1}$ be a direction K_0-perpendicular to aff(σ). Then, the k-*indegree of σ in direction s *is the number of k-dimensional cofaces of σ that have the same height as σ in direction s.*

To compute the k-indegree of a simplex $\sigma \in K$, in Algorithm 3.1, we combine information from wedges constructed from σ-isolating directions. Simplex-isolating directions are

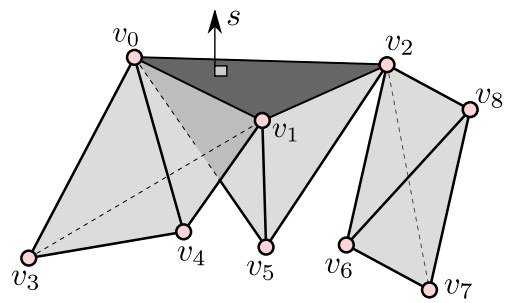


Figure 3.3: Computing the three-indegree for a two-simplex (triangle) in $\mathbb{R}^4$. The simplex σ is shown in dark gray. The direction $s \in \mathbb{S}^2$ is orthogonal to $\text{aff}(\sigma)$ such that all other vertices shown are below σ (note that $s = s_\sigma$ from Definition 3.10(a)). Although the three-indegree of σ is one, $\hat{\rho}(s)$ sees three tetrahedron at the same height as σ. Recursively using the three-indegree of all faces of σ in tilted directions (that is, directions given in Definition 3.10(b)), the three-indegree of σ can be defined by subtracting the indegrees of faces of σ in specific directions $(3 - 1 - 1 = 1)$ given in Equation (3.6).

directions that differ by small enough values so that they define a wedge of regions between filtration hyperplanes that only contain the vertices of a single potential k-simplex and no other vertices (see Figure 3.1 for an illustration of this idea).

Although a direction s being K_0-perpendicular to $\text{aff}(\sigma)$ means all zero-simplices of σ are at the same height in direction s, not all k-simplices at this height contribute to the k-indegree of σ; see Figure 3.3. This is why diagrams from Step 2a. of Theorem 3.27 are not sufficient alone. Thus, we have Algorithm 3.1, which finds the k-indegree of a given simplex in a particular direction. To prove correctness of Algorithm 3.1, we use the following lemma.

Lemma 3.15 (Indegree Contributors)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $k \in \mathbb{R}$ and $\tau, \sigma, \sigma' \in K$ such that $\tau \prec \sigma$ and $k = \dim(\sigma) = \dim(\sigma')$. Let s be K_0-perpendicular to σ and let s_τ be a (K_0, σ, τ, s)-perturbation. Then, σ' contributes to the k-indegree of τ in direction s_τ if and only if σ' is at the same height as σ in direction s_τ and $\tau = \sigma \cap \sigma'$.*

Algorithm 3.1: `ComputeIndegree`$(\sigma, s, k, \hat{\rho}(K,S), T = \{\})$

Input: $\sigma \in K$, $s \in S$ s.t. s is K_0-perpendicular to σ, $k \in \mathbb{N}$ s.t. $k > \dim(\sigma)$, $\hat{\rho}(K,S)$, where S satisfies the assumptions of Theorem 3.12, and a table T for memoization.

Output: the k-indegree for σ

1: $c \leftarrow$ height of σ in direction s

2: `numDeaths` $\leftarrow$ number of $(k-1)$-dimensional deaths in $\hat{\rho}(h_s)$ at height c

3: `numBirths` $\leftarrow$ number of k-dimensional births in $\hat{\rho}(h_s)$ at height c

4: `doubleCounts` $\leftarrow 0$

5: **for** $\tau \prec \sigma$ in non-descending order by dimension **do**

6: $W \leftarrow \mathrm{verts}(\sigma) \setminus \mathrm{verts}(\tau)$

7: $s_\tau \leftarrow$ a $(K_0, \mathrm{verts}(\sigma), W, s)$-perturbation

8: **if** $T[\tau]$ was not computed yet **then**

9: $T[\tau] \leftarrow$ `ComputeIndegree`$(\tau, s_\tau, k, \hat{\rho}(K,S), T)$

10: **end if**

11: `doubleCounts` $\leftarrow$ `doubleCounts` $+ T[\tau]$

12: **end for**

13: **return** `numDeaths` $+$ `numBirths` $-$ `doubleCounts`

Proof. Suppose that $f\colon K \to \mathbb{R}$ and $f_\tau\colon K \to \mathbb{R}$ are the filter functions for the directions s and s_τ, respectively.

($\Rightarrow$) Suppose that σ' contributes to the k-indegree of τ in direction s_τ. Then, by the definition of k-indegree, $\tau \prec \sigma'$ and σ' is at the same height as τ, which is the same height as σ. Since $\tau \prec \sigma$ by assumption, we have $\tau \preceq \sigma \cap \sigma'$. We must now show that $\sigma \cap \sigma' \preceq \tau$.

By contradiction, suppose that $\sigma \cap \sigma' \not\preceq \tau$. Then, there exists a vertex $v \in \sigma \cap \sigma'$ such that $v \notin \tau$. Thus, $v \in \sigma \setminus \text{verts}(\tau)$. Since s_τ is a (K_0, σ, τ, s)-perturbation, we know that $s_\tau \cdot v > s_\tau \cdot w$ for all $w \in \text{verts}(\tau)$, a contradiction to the claim that σ' contributes to the k-indegree of τ in direction s_τ. Therefore, $\sigma \cap \sigma' \preceq \tau$ as required.

($\Leftarrow$) Suppose that σ' is at the same height as σ in direction s_τ and that $\tau = \sigma \cap \sigma'$. Denote the dimension of τ by j. Since σ' is a k-simplex, τ is a j-simplex, and $\tau \prec \sigma'$, we can write $\tau = [v_0, v_1, \ldots, v_j]$ and $\sigma' = [v_0, v_1, \ldots, v_k]$ where $v_i \in K_0$. Then,

$$f_\tau(\sigma') = \max_{i=0}^{k} f_\tau(v_i) = \max\left(\max_{i=0}^{j} f_\tau(v_i), \max_{i=j+1}^{k} f_\tau(v_i)\right) = \max\left(f_\tau(\tau), \max_{i=j+1}^{k} f_\tau(v_i)\right). \quad (3.2)$$

Since σ' is at the same height as σ in direction s_τ and $\tau \prec \sigma$, σ' is also at the same height as τ in direction s, meaning that $f(v_i) \leq f(\tau)$ for all $0 \leq i \leq$ k. Since v_i is not in τ for $i > j$, it must also be the case that $f(v_i) < f(\tau)$ for $i > j$. Since s_τ is a (K_0, σ, τ, s)-perturbation, by Statement (iii) of Definition 3.8, any vertex below τ in direction s is also below τ in direction s_τ. Thus, $f_\tau(v_i) < f_\tau(\tau)$ for all $j < i \leq k$ and

$$\left(\max_{i=j+1}^{k} f_\tau(v_i)\right) < f_\tau(\tau).$$

Then, by Equation (3.2), $f_\tau(\sigma') = f_\tau(\tau)$. This taken together with $\tau \prec \sigma'$ shows that σ' contributes to the k-indegree of τ in direction s_τ. $\square$

Although Lemma 3.15 allows us to state formally which cofaces contribute to a simplex's k-indegree, the computation of k-indegree requires more than a single diagram (see

Figure 3.3). Thus, we use an inclusion-exclusion style argument to compute the k-indegree in Algorithm 3.1, using T, a table that keeps track of any contributers to k-indegree that would be double counted. We prove correctness of Algorithm 3.1 in the following theorem.

Theorem 3.16 (Computing k-Indegree)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $\sigma \in K$ and $s \in \mathbb{S}^{d-1}$ such that s is K_0-perpendicular to σ. Then, for $k > \dim(\sigma)$,* `ComputeIndegree`$(\sigma, s, k, \hat{\rho}(K, S))$ *returns the k-indegree of σ in direction s.*

Proof. We prove the claim inductively on $j = \dim(\sigma)$. For the base case ($j = 0$), let $k > j$ and consider the zero-simplex $[v]$. Let $h_s \colon K \to \mathbb{R}$ be the filter function for direction s. We note that this base case is a generalization of [8, Lemma 11]. However, unlike in [8, Lemma 11], we are only making an argument for the k-indegree at a single vertex and not all vertices. As such, we can relax the requirement that no two vertices in K_0 have the same height in direction s and just require that no vertices in $K_0 \setminus \{v\}$ have the same height in direction s as v. Thus, we have that k-indegree of σ is equal to the number of k-simplices that have height $h_s(v)$, which, by Lemma 3.3, is

$$|f^{-1}(f(v))| = |\{(a, b) \in \hat{\rho}_{k-1}(h_s) \text{ s.t. } b = f(v)\}| + |\{(a, b) \in \hat{\rho}_k(h_s) \text{ s.t. } a = f(v)\}|. \quad (3.3)$$

In Algorithm 3.1, notice that if σ is a single vertex, we do not enter the loop that starts on Line 4. Thus, the return value is exactly the number given in Equation (3.3).

For the inductive assumption, let $j \geq 0$. We assume that Algorithm 3.1 returns the k'-indegree of τ in direction s, for all $\tau \in K_j$ and all $k' > j$. For the inductive step, let $\sigma \in K_{j+1}$. Let $k > j + 1$. Now, we compute the k-indegree of σ in direction s. Using Lemma 3.3, we

know that the number of k-simplices with height $h_s(\sigma)$ in direction s is:

$$\delta := |\{(a,b) \in \hat{\rho}_{k-1}(h_s) \text{ s.t. } b = f(\sigma)\}| + |\{(a,b) \in \hat{\rho}_k(h_s) \text{ s.t. } a = f(\sigma)\}|. \tag{3.4}$$

Let F_σ denote this set of simplices, let $\sigma' \in F_\sigma$, and let $\tau \prec \sigma$. Suppose that s_τ is a $(K_0, \text{verts}(\sigma), \text{verts}(\sigma \setminus \tau), s)$-perturbation. By Lemma 3.15, the k-simplex σ' contributes to the k-indegree of τ in direction s_τ if and only if $\tau = \sigma \cap \sigma'$.

Then, we can isolate each face of $\tau \prec \sigma$ and compute the k-indegree of τ using Equation (3.4) and add or subtract it from the k-indegree of σ, alternating by dimension of τ. This ensures that no coface of $\tau \prec \sigma$ adds to the k-indegree of σ. Formally, this is seen in the equation for the k-indegree of σ

$$\delta - \sum_{\tau \prec \sigma} \delta_\tau, \tag{3.5}$$

where δ_τ is the k-indegree of τ in the corresponding tilted directions. In Algorithm 3.1, `numDeaths+numBirths` is equal to δ, and the values δ_τ are computed in Line 8 of Algorithm 3.1. Thus, the return value matches Equation (3.6). □

3.16.0.1 Simplex Certificate

Using the k-indegree and the diagrams guaranteed by Lemma 3.11, we are able to isolate potential k-simplices between two hyperplanes centered at a simplex. In Algorithm 3.2, we certify if a set of $(k+1)$-vertices forms a k-simplex by first checking if there are appropriate directions in our set to create a wedge centered at a subset of k vertices. If not, we know

there is no simplex there. If there are directions in the set that form the appropriate wedge, we then use the difference in the indegree between the two filtration hyperplanes defining the wedge to check for the presence of a k-simplex.

Algorithm 3.2: `Certificate`$(V, \hat{\rho}(K, S))$

Input: $V \subset K_0$, a set of vertices and $\hat{\rho}(K, S)$, where S is simplex-isolating (Definition 3.10).

Output: TRUE if V defines a $(|V| - 1)$-simplex of K, FALSE otherwise

1: $k \leftarrow |V| - 1$

2: $w \leftarrow$ a vertex of V

3: **if** there is a direction s_U that is K_0-perpendicular to V (as in Definition 3.10(a)) and if s_L is a $(K_0, V, \{w\}, s)$-perturbation (as in Definition 3.10(b)) **then**

4: $x \leftarrow |$`ComputeIndegree`$(V, s_U, k, \hat{\rho}(K, S)) -$ `ComputeIndegree`$(V, s_L, k, \hat{\rho}(K, S))|$

▷ Algorithm 3.1

5: **return** $x == 1$

6: **else**

7: **return** FALSE

8: **end if**

Theorem 3.17 (Correctness of Simplex Certificate)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $V \subseteq K_0$. Then, Algorithm* 3.2 *returns* TRUE *iff V defines a simplex.*

Proof. We must show that `Certificate`$(V, \hat{\rho}(K, S))$ returns TRUE iff V defines a simplex. By Lemma 3.11, every there are directions as described in Line 3 for every set of vertices V and vertex $w \in V$ where V forms a simplex of K. Thus, if no such directions exist for V in

our set of directions, we immediately know V does not form a simplex, and return FALSE on Line 7. Suppose, then, the directions of Line 3 exist for V and $w \in V$, so we compute the difference of indegrees on Line 4. By construction, all vertices of $V \setminus \{w\}$ are at the same height with respect to s_U and s_L; we denote these heights by c_U and c_L, respectively.

Since s_U is perpendicular to V, we have $c_U = s_U \cdot w$. Since s_L is a (K, V, w, s_U)-perturbation, we see that $c_L > s_L \cdot w$ (by Statement (ii) of Definition 3.8). Thus, we know $w \in H^{\downarrow}(s_U, c)$ and $w \notin H^{\downarrow}(s_L, c')$, which means that $w \in \mathbb{W} := H^{\downarrow}(s_U, c) \triangle H^{\downarrow}(s_L, c')$.

Next, we show that only the vertices of V lie in $\mathbb{W}$. Assume, for contradiction, that there exists a vertex $w_* \in \mathbb{W} \cap K_0 \setminus V$. Then, since w_* can't be at height c_U (since s_U is assumed to be K_0-perpendicular to aff(V)), we have $s_U \cdot w_* < c_U$. Then, by Statement (iii) of Definition 3.8, $s_L \cdot w_* < c_L$, which implies that w_* is not in $\mathbb{W}$, a contradiction. Thus, $\mathbb{W}$ only contains the vertices of V as required.

We now show that the k-indegree can be used to determine if V defines a simplex in K. Since $w \notin H^{\downarrow}(s_U, c_U)$ and $w \in H^{\downarrow}(s_L, c_L)$, every simplex that contributes to the k-indegree of $V \setminus \{w\}$ in direction s_U also contributes to the k-indegree of $V \setminus \{w\}$ in direction s_L. The only thing contributing to the k-indegree of $V \setminus \{w\}$ in direction s_L but not in direction s_U is the potential simplex defined by V. Thus, Algorithm 3.2 returns TRUE iff $V \in K_k$. $\square$

3.17.1 Proof of Theorem 3.12 and Its Corollaries

Using the results from the previous subsection, we arrive at Algorithm 3.3 that fully reconstructs a GP-immersed simplicial complex.

Algorithm 3.3: `ReconstructComplex`($\hat{\rho}(K, S)$)

Input: $\hat{\rho}(K, S)$, where S is vertex- and simplex-isolating (Definitions 3.9 and 3.10).

Output: simplicial complex K.

1: $K_0 \leftarrow$ vertices of K, as found using the methods of [8, Theorem 9]

2: **for** $V \subseteq K_0$ with $1 < |V| \leq d$ and in non-decreasing size of V **do**

3: $k \leftarrow |V| - 1$

4: **if** $V \setminus \{v_i\} \in K_{k-1}$ for all $v_i \in V$ **then**

5: **if** `Certificate`(V) == TRUE **then**

6: Add V to K_k

7: **end if**

8: **end if**

9: **end for**

10: **return** $K_0 \cup K_1 \cup \cdots \cup K_\kappa$

Theorem 3.18 (Simplicial Complex Reconstruction)**.** *Let K be a κ-dimensional simplicial complex GP-immersed in $\mathbb{R}^d$, such that $\kappa \leq d - 1$. Let $S \subset \mathbb{S}^{d-1}$ satisfy the assumptions of Theorem 3.12 (that is, (K, S) is vertex- and simplex-isolating). Then, Algorithm 3.3 reconstructs K.*

Proof. We begin by reconstructing K_0 on Line 1 using the methods of [8, Theorem 9] and the vertex-isolating directions. Algorithm 3.3 then iterates over all subsets of vertices $V \subset K_0$. Suppose that σ is the simplex defined by the vertex set V. We do not yet know if $\sigma \in K$. Since sets are included in non-decreasing size, K_{k-1} is finalized by the time V

is considered. The condition that $\partial(\sigma) \subset K_{k-1}$ is checked on Line 4. For such a vertex set, V, the algorithm certifies if V forms a $(|V|-1)$-simplex in K. By Theorem 3.17, `Certificate`$(V, \hat{\rho}(K,S))$ (Algorithm 3.2) returns TRUE iff $\sigma \in K$. Since we iterate over all subsets of K_0, Algorithm 3.3 eventually finds all simplices, and therefore reconstructs the simplicial complex K. □

This theorem concludes the proof of Theorem 3.12.[1] Note that if a set of directions S is vertex- and simplex-isolating for a simplicial complex K, it is also vertex- simplex-isolating for any subcomplex of K. An immediate corollary is as follows:

Corollary 3.19 (Subcomplexes are Represented)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $S \subset \mathbb{S}^{d-1}$ satisfy the assumptions of Theorem 3.12. Let K' be any subcomplex of K, and let $h'_s : K' \to \mathbb{R}$ be the lower-star filter function of K' in direction s. Then $\hat{\rho}(K', S)$ is a faithful discretization of $APHT(K')$.*

3.19.1 Explicit Construction of a Faithful Discretization

Here, we provide algorithms to explicitly build a set of directions $S \subseteq \mathbb{S}^{d-1}$, given a simplicial complex K (GP-immersed in $\mathbb{R}^d$), so (K, S) is vertex- and simplex-isolating.

3.19.2 Auxiliary Constructions

This section describes the methods that are used to compute the directions in Theorem 3.27: how to use Gram-Schmidt to find a direction orthogonal to a given plane, how to find a set of points that span a given affine subspace, how to "tilt" one direction

[1] If the complex contains codimension-zero simplices (that is, if $\kappa = d$), we can still compute a representative set. The details of this case, including modified reconstruction algorithm, can be found in [29], version 3.

towards another direction while controlling the order of a given set of points, and how to use tilt in order to "pop" a subset of vertices off of a given hyperplane.

Computing a Perpendicular Direction We often compute directions orthogonal to the affine plane spanned by some set of points. To do so, we use Gram-Schmidt orthogonalization to first find $n-1$ vectors in the space spanned by the points, and then perform Gram-Schmidt orthogonalization on the standard basis vectors until we find one that is not in the space spanned by the points. The result is the orthogonal vector that we desire. While intuitive, the explicit formulation is not easily accessible from literature, so we derive the formulation here.

Algorithm 3.4: `Perp`(P)

Input: $P = \{p_0, p_1, \ldots, p_k\} \subset \mathbb{R}^d$, a set of affinely independent points such that $k \leq d$.

Output: $s \in \mathbb{S}^{d-1}$, a direction that is orthogonal to aff(P).

1: $A \leftarrow$ the $d \times k$ matrix containing $p_i - p_0$ in column i for for $1 \leq i \leq k-1$

2: $\{b_i\}_{i=1}^{k} \leftarrow$ basis vectors for the space spanned by A ▷ QR-decomposition

3: $i \leftarrow 0$

4: $s \leftarrow 0$

5: **while** $s = 0$ **do**

6: $s \leftarrow$ the output of Gram-Schmidt on $\{b_i\}_{i=1}^{d-k}$ with e_i

7: $i \leftarrow i + 1$

8: **end while**

9: **return** s

Lemma 3.20 (Computing a Perpendicular Direction)**.** *Let $P \subset \mathbb{R}^d$ be a point set in general position with $|P| \leq d$. Then, the output of Algorithm 3.4 is a direction $s \in \mathbb{S}^{d-1}$ orthogonal to aff(P) and runs in $\Theta(d^3)$ time.*

Proof. Denote the points of P by $p_0, p_1, \ldots, p_k$. In Line 1, we define A as the $d \times k$ matrix consisting of $p_i - p_0$ in column i. By Theorem 7.1 in [59], A has a reduced QR-factorization in which the columns of Q are an orthonormal basis for the space spanned by A, or aff(P). We find Q using k iterations of Gram-Schmidt on Line 2.

Then, in the loop at Lines 5–8, we iteratively test the result of performing Gram-Schmidt on the columns of Q and a basis vector. The result is zero if the basis vector is contained in aff(P). As soon as the result is nonzero, (which has to happen since aff(P) has positive codimenison, so can't be spanned by all e_is), we have found a vector orthogonal to aff(P), and the value returned on Line 9 is as desired.

Building the matrix A and normalizing s takes constant time. Finding the reduced QR-factorization takes k iteration of Gram-Schmidt, (Theorem 8.1, [59]) and finding the final vector that is orthogonal to the space spanned by P requires at most d iterations of Gram-Schmidt. Since each iteration of Gram-Schmidt takes $O(dk)$ time, we see that the total runtime is $\Theta(d^3)$. □

The ability to find a direction perpendicular to some set of points using Algorithm 3.4 is be used in Algorithm 3.7. Before introducing Algorithm 3.7, we first provide two additional algorithms that are be utilized in Algorithm 3.7.

Plane Filling Given a point set $P \subset \mathbb{R}^d$ of k affinely independent points and a direction s, Algorithm 3.5 finds a complementary set of points P' such that $\text{aff}(P' \cup P)$ has only two perpendicular directions, s and s'. In other words, we find enough points in $\mathbb{R}^d$ so that they, along with the original point set P, 'fill' the plane. We do so by first considering a matrix A that describes the equation of the $(d-1)$-plane orthogonal to s containing the points of P. Recall that the left null space of A is the space of all vectors n such that $n^T A = 0$. Thus, for such a vector n, the points $n - p_0$ are also in the plane described by A. We are able to find $d - |P|$ of these vectors and corresponding points by computing a basis of the left null space, since this space is $d - |P|$ dimensional.

Lemma 3.21 (Plane Filling)**.** *Algorithm 3.5 is correct and runs in $\Theta(d^3)$ time.*

Proof. We first prove the runtime of this algorithm by analyzing what is done in each line of the algorithm. First, we initialize P' to the empty set. Then, we construct a matrix A (Line 2) containing $p_i - p_0$ in column i for the first $d - k$ columns, and the vector s in the last column, which takes $\Theta(nd)$ time. As the points in P are affinely independent and s is orthogonal to $\text{aff}(P)$, the dimension of the column space of the matrix A (Line 2) is k, which means that there are $d - k$ vectors in its nullspace. We find the nullspace by first finding a basis for the space spanned by A in Line 3, via a full QR-decomposition [59], and then by using Gram-Schmidt in Lines 5–13, taking $\Theta(d^3)$ time. Line 7 always has at least $d - k$ nonzero outputs, and we stop once we have $d-k$ directions perpendicular to $\text{aff}(P)$. We then iteratively add points to P' in Line 9. Computing the full QR-decomposition and performing Gram-Schmidt dominates the algorithm; hence, the running time for Algorithm 3.5 is $\Theta(d^3)$.

Algorithm 3.5: `PlaneFill`(P, s)

Input: $P = \{p_0, p_1, \ldots, p_k\} \subset \mathbb{R}^d$, a set of affinely independent points such that $k \leq d$; $s \in \mathbb{S}^{d-1}$, a direction that is orthogonal to aff(P)

Output: a set $P' \subset \mathbb{R}^d$ of $d - |P|$ points at the same height as P with respect to direction s such that $\dim(\text{aff}(P' \cup P)) = d - 1$.

```
P' ← ∅
A ← the d × k matrix containing p_i − p_0 in column i for for 1 ≤ i ≤ k − 1 and s in the kth column
{b_i}_{i=1}^{k+1} ← basis vectors for the space spanned by A        ▷ QR-decomposition
j ← 0
for i from 1 to d do
while j < d − k do
s_j ← the output of Gram-Schmidt on {b_i}_{i=1}^{k+1} with e_i
if s_j ≠ 0 then
P' ← P' ∪ {s_j + p_0}
j ← j + 1
end if
end while
end for
return P'
```

Finally, for correctness, we prove that the output of Algorithm 3.5 has the desired properties. By construction, we have $|P'| = d - |P|$, i.e., $\dim(\text{aff}(P' \cup P)) = d - 1$. To show each point in P' is at the same height as P, it suffices to show that $s \cdot p_0 = s \cdot p'_i$ for all $p'_i \in V'$. Since each s_j considered in Line 9 is in the basis for the left null space of A and since s is a column of A, all such vectors s_j are orthogonal to s. Then, indeed, for all $p'_i \in V'$, we have $s \cdot p'_i = s \cdot (s_j + p_0) = s \cdot p_0$. □

Tilting In the Algorithm 3.6, we find a direction that is a slight tilt of one input direction towards another, so that no vertex orders change. First, we explain the geometric intuition of the algorithm. Let S be the set of n line segments in $\mathbb{R}^2$:

$$S := \left\{ \overline{(0, s \cdot p), (1, s' \cdot p)} \right\}_{p \in P}.$$

Each line segment in S represents a linear interpolation between $(0, s \cdot p)$ and $(1, s' \cdot p)$, which correspond to the heights in directions s and s' of each point in P. Moreover, we can parameterize each line segment in S as $(1-t)s \cdot p + ts' \cdot p$ for $t \in [0, 1]$ (see the grey lines in Figure 3.4). Then, vertical cross sections record vertex heights with respect to some direction that is an interpolation of s and s'. We want to identify a particular $t_* > 0$ such that the ordering of the heights of points in direction $s_t = (1 - t_*)s + t_* s'$ is consistent with the ordering of the points in direction s. Notice that no swapping of point order can occur before the intersection of black lines in Figure 3.4; thus, we choose the direction that is halfway to this point of the interpolation (corresponding to the dashed vertical line in Figure 3.4) as the output of Algorithm 3.6.

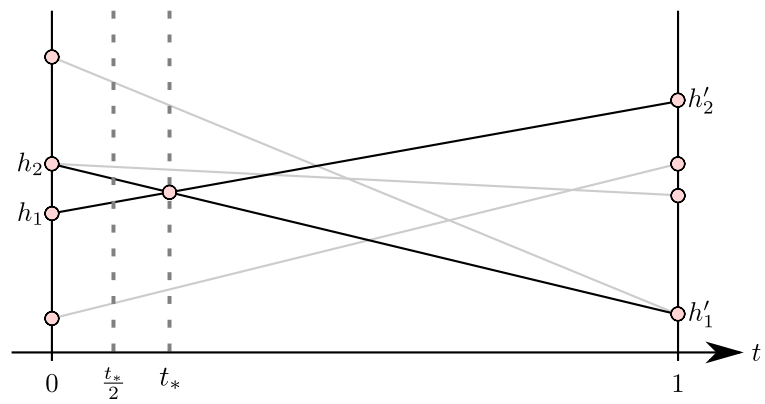


Figure 3.4: The solid grey lines in the figure above indicate the changing heights of points as we swing direction s towards s'. Although we do not explicitly compute the grey lines, we know by simple geometry that no intersection of grey lines (and, in particular, no swapping of point orders) occurs before the t-value t_*, which corresponds to the intersection of the closest pairwise heights of points on the left and the extremal heights of points on the right, as indicated by the black lines. Since there are no crossings of line segments before $\frac{1}{2}t_*$, there is therefore no change in the order of points with respect to direction $s_t = (1 - \frac{1}{2}t_*)s + \frac{1}{2}t_*s'$.

Algorithm 3.6: `Tilt`(s, s', P)

Input: $P \subset \mathbb{R}$ finite; $s, s' \in \mathbb{S}^{d-1}$.

Output: $s_t \in \mathbb{S}^{d-1}$, a direction satisfying Statements (i)-(iii) of Lemma 3.22.

1: $h_1, h_2 \leftarrow$ heights of points that are closest with respect to s such that $h_1 \neq h_2$

2: $h_1', h_2' \leftarrow$ minimum and maximum heights of points with respect to s'

3: $t_* \leftarrow$ the solution to $(1-t)h_1 + th_2' = (1-t)h_2 + th_1'$

4: **return** $(1 - \frac{1}{2}t_*)s + \frac{1}{2}t_*s'$

Lemma 3.22 (Tilt)**.** *Let $P \subset \mathbb{R}^d$ be a finite set. Let $s, s' \in \mathbb{S}^{d-1}$. Then, using Algorithm 3.6, we can compute a third direction $s_t =$ `Tilt`(s, s', P) in $\Theta(|P| \log |P| + d + \Pi)$ time such that the following properties holds for all points $p_1, p_2 \in P$:*

(i) If p_1 is strictly above (below) p_2 with respect to direction s, then p_1 is strictly above (below, respectively) p_2 with respect to direction s_t.

(ii) If p_1 and p_2 are at the same height with respect to direction s and p_1 is strictly above (below) p_2 with respect to direction s', then p_1 is strictly above (respectively, below) p_2 with respect to direction s_t.

(iii) If p_1 is is at the same height as p_2 with respect to both directions s and s', then p_1 and p_2 are at the same height with respect to direction s_t.

Proof. First let h_1, h_2 be the heights of points in P with the closest unequal heights in direction s, and let h'_1, h'_2 be the heights of points in P with the extreme heights in direction s', as in Lines 1 and 2. Consider the lines between segments connecting h_1 to h'_2 and h_2 to h'_1; they intersect at some point i that is at least as close to the left as the leftmost non-zero intersection of all linear interpolations of point heights with respect to directions s and s' (see Figure 3.4). Let t_* denote the first coordinate of i as on Line 3. Since $\frac{1}{2}t_* < t_*$, segments in the interval $(0, \frac{1}{2}t_*]$ of linear interpolations of point heights must not have any crossings, and so the ordering of points that have unique heights with respect to s is maintained, i.e., we see that s_t satisfies Statement (i). Furthermore, if points have the same height with respect to direction s, they correspond to an intersection at zero of heights in the linear interpolation of point heights, meaning that s_t orders points equivalently to how s' orders the points, satisfying Statement (ii) and Statement (iii).

To find h_1, h_2, h'_1, h_2, we can sort the heights of points in directions s and s' in $\Theta(|P| \log |P|)$ time. Finding i from the resulting two segments takes constant time, and returning $s_t = (1 + \frac{1}{2}t_*)s + \frac{1}{2}t_* s'$ takes $\Theta(d)$ time. Thus, the total runtime of Algorithm 3.6 is $\Theta(|P| \log |P| + d + \Pi)$. $\square$

Tilting to Pop Given a point set $P \subset \mathbb{R}^d$ in general position, two sets $W \subseteq V \subseteq P$, and a direction s that is P-perpendicular to V, Algorithm 3.7 calculates a direction that is close to s that "pops" all of the vertices in W either entirely above or below the vertices in $V \setminus W$.

Algorithm 3.7: `TiltToPop`(P, V, W, s)

Input: P, V, W point sets in $\mathbb{R}^d$ such that P is in general position, $W \subseteq V \subseteq P$, and $|V| \leq d$; s, a direction P-perpendicular to V.

Output: $s_* \in \mathbb{S}^{d-1}$, a direction that is a (P, V, W, s)-perturbation.

1: $V' \leftarrow$ `PlaneFill`(V, s) ▷ Algorithm 3.5

2: $s' \leftarrow$ `Perp`$(\text{aff}(V' \cup (V \setminus W) \cup \{W - s\}))$ ▷ Algorithm 3.4

3: **if** W is below $V \setminus W$ with respect to s' **then**

4: $s' \leftarrow -s'$

5: **end if**

6: **return** $s_* =$ `Tilt`$(s, s', P \cup V')$ ▷ Algorithm 3.6

The algorithm begins in Line 1 by finding a set of points $V' \subset \mathbb{R}^d$ such that $\text{aff}(V' \cup V)$ is a $(d-1)$-dimensional subspace of $\mathbb{R}^d$ in $\Theta(d^3)$ time by Lemma 3.21 The additional points help us control which way to tilt s. In particular, the direction s' (computed on Line 2 in $\Theta(d^3)$ by Lemma 3.20) is perpendicular to $\text{aff}(V' \cup (V \setminus W))$, and is the direction towards which we can tilt in order to "pop" W off of the plane orthogonal to s at height $V \setminus W$. Since there are two choices for s' in Line 2, the if statement in Lines 3 and 4 ensures that the direction is the one such that W is above $V \setminus W$, and is completed in $\Theta(d)$ time. Finally, we return s_* on Line 6 using Algorithm 3.6, taking $\Theta(|P| \log |P| + d + \Pi)$ by Lemma 3.22.

Lemma 3.23 (Tilting to Pop)**.** *Let P be a finite point set in $\mathbb{R}^d$ in general poistion. Let $W \subseteq V \subseteq P$ and let $s \in \mathbb{S}^{d-1}$ be perpendicular to aff(V). Then Algorithm 3.7 finds* `TiltToPop`(P, V, W, s) *in $\Theta(|P| \log|P| + d^3)$ time, resulting in a (P, V, W, s)-perturbation.*

Proof. The runtime was justified in the paragraph above detailing the algorithm. Recall what it means for the returned direction to be a (P, V, W, s)-perturbation:

(i) The points in W are above $V \setminus W$ with respect to direction returned.

(ii) For all $p \in P \setminus V$, p is strictly above (below) the height of $V \setminus W$ with respect to the direction returned iff it is strictly above (below, respectively) V with respect to s.

(iii) The direction returned is P-perpendicular to $V \setminus W$.

Before proving Statements (i)-(iii), we establish three properties of s' returned on Line 2. First, since aff$(V' \cup (V \setminus W))$ has codimension one and the points are in general position, s' is automatically P-perpendicular to this space.

Next, we show that all vertices of W have the same height with respect to s'. Let $w_1, w_2 \in W$. Since $(w_1 - s), (w_2 - s) \in$ aff$(V' \cup (V \setminus W) \cup \{W - s\})$ we have $s' \cdot (w_1 - s) = s' \cdot (w_2 - s)$. Then by adding $s' \cdot s$ to both sides, we obtain $s' \cdot w_1 = s' \cdot w_2$, i.e., all points of W are at the same height with respect to s'.

Finally, we show $W \notin$ aff$(V' \cup (V \setminus W) \cup \{W - s\})$. Suppose not. Then, since both W and $W - s$ are in aff$(V' \cup (V \setminus W) \cup \{W - s\})$, the vector s is parallel to aff$(V' \cup (V \setminus W) \cup \{W - s\})$. Furthermore, since both planes contain the same linearly independent set of $(d-1)$ points, namely $V' \cup V$, we have aff$(V' \cup (V \setminus W) \cup \{W - s\}) =$ aff$(V' \cup V)$. But then s is parallel to

$\text{aff}(V' \cup V)$, contradicting the property that s is P-perpendicular to $\text{aff}(V' \cup V)$. Thus, we see that s' orders all vertices of W on the same side of $V \setminus W$; on Line 4, we ensure that all vertices of W are specifically above $V \setminus W$ with respect to s', therefore, by Lemma 3.22((ii)), all points of W are above $V \setminus W$ with respect to the direction s_* returned on Line 6 and we have shown s_* satisfies Statement (i).

We are now ready to show s_* satisfies Statements (ii) and (iii). Since both s and s' are perpendicular to $\text{aff}(V \setminus W)$, the output $\texttt{Tilt}(s, s', P \cup V') = s_*$ is perpendicular to $\text{aff}(V \setminus W)$ by Lemma 3.22((iii)). To show that no other vertex of P is at the same height of vertices in $V \setminus W$ with respect to s_*, let $p \in P \setminus (V \setminus W)$. If $p \in P \setminus V$, since p is strictly above or below $V \setminus W$ with respect to direction s, then by Lemma 3.22((i)), p is strictly above or below $V \setminus W$ with respect to s_*, showing Statement (ii). If $p \in W$, then it is at the same height as $V \setminus W$ in direction s and above $V \setminus W$ in direction s', thus, by Lemma 3.22((ii)), p is above $V \setminus W$ in direction s_* and we have shown that s_* is P-perpendicular to $V \setminus W$, so Statement (iii) of the current lemma is satisfied, concluding our proof. □

3.23.1 Building an Explicit Set

Next, we use the algorithms of Section 3.19.2 to construct directions in Theorem 3.27 for a given simplicial complex.

Construction 1 (Constructing a Faithful Set)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$ such that $\dim(K) < d$, and let S be the following set of directions constructed iteratively as follows:*

Step 1 Initially, let S be the standard basis vectors $(e_1, e_2, \ldots, e_d)$, plus the additional direction

that is the output of $\texttt{PointIso}(K_0)$.

Step 2 For every maximal $\sigma \in K$,

a. *Let* $s = \texttt{Perp}(\sigma)$ *and let* H *be the set of all vertices with the same height as* σ *in direction* s, *and let* $W = \texttt{PlaneFill}(H, s) \cup (H \setminus \sigma)$ *and* $V = W \cup \sigma$. *Then, add the direction* $s_\sigma := \texttt{TiltToPop}(K_0 \cup W, V, W, s)$ *to* S.

b. *For each* $\tau \in K$ *such that* $\tau \prec \sigma$, *add the direction* $\texttt{TiltToPop}(K_0, \sigma, \tau, s_\sigma)$ *to* S.

The remainder of this section shows that Construction 1 forms a faithful discretization.

3.23.1.1 Directions for Vertices

Constructing a set of directions that faithfully represents a vertex set has been explored in previous work. By [8, Lemma 7], it suffices to construct a set of d linearly independent directions, plus one additional direction so that there are exactly n_0 intersections of size $d+1$ among all associated filtration hyperplanes. However, the construction of this final direction given in [8, Lemma 8] requires stricter general position assumptions, namely, that no two vertices share any e_i-coordinate for $1 \leq i \leq d$. Algorithm 3.8 produces such a $(d+1)$st direction when our pointset satisfies only the mild conditions of of Assumption 3.

The next lemma proves correctness of Algorithm 3.8.

Lemma 3.24 (Correctness of Algorithm 3.8)**.** *Let* P *be a finite point set in* $\mathbb{R}^d$ *in general position. Then,* $\texttt{PointIso}(P)$ *returns a direction that uniquely orders the filtration grid of* P *with respect to* $\{e_1, e_2, \ldots, e_d\}$ *and runs in* $\Theta(d^2|P|^d \log |P|)$ *time.*

Algorithm 3.8: `PointIso`(P)

Input: $P \subset \mathbb{R}^d$, a point set in general position

Output: $s \in \mathbb{S}^{d-1}$, a direction that uniquely orders the filtration grid of P with respect to $\{e_1, e_2, \ldots, e_d\}$

1: $A \leftarrow$ the filtration grid of P with respect to $\{e_1, e_2, \ldots, e_d\}$

2: $s \leftarrow e_1$

3: **for** i from 2 to d **do**

4: $s \leftarrow$ `Tilt`(s, e_i, A)

5: **end for**

6: **return** s

Proof. First, we analyze runtime. On Line 4, we call Algorithm 3.6, which by Lemma 3.22 takes $\Theta(|A| \log |A| + d)$ time. Since this is called $d-1$ times during the loop on Lines 3–5 and since the filtration grid has cardinality $|A| = \Theta(|P|^d)$, the total runtime of Algorithm 3.8 is $\Theta(d(|P|^d \log |P|^d + d)) = \Theta(d^2 |P|^d \log |P|)$.

Next, we show Algorithm 3.8 is correct. Let π_i be the standard projection map onto the $(e_1, e_2, \ldots e_i)$-plane. As on Line 1, let A be the filtration grid of P with respect to $\{e_1, e_2, \ldots, e_d\}$ and note that A is a grid of at most $|P|^d$ points. Let j be the number of times the loop on Lines 3–5 has been completed.

We use the loop invariant that s totally orders the unique points of the image $\pi_{j+1}(A)$. We initialize $s = e_1$ on Line 2. Since e_1 totally orders the points of $\pi_1(A)$, the loop invariant is satisfied before entering the loop.

Next, we show the loop invariant is maintained. Let s^i denote the ith value of s in

Algorithm 3.8 (so s^1 is the initial direction defined in Line 2, s^2 is the direction updated by tilting towards e_2 the first time we encounter Line 3, etc. Note that this means $i = j + 1$.) Suppose that the loop invariant is true going into the for loop of Lines 3–5. Recall by Lemma 3.22 that $\texttt{Tilt}(s^{i-1}, e_i, A)$ produces a direction s^i so that, for all $a_1, a_2 \in A$,

(i) If a_1 is strictly above (below) a_2 with respect to direction s^{i-1}, then a_1 is strictly above (below, respectively) a_2 with respect to direction s^i.

(ii) If a_1 and a_2 are at the same height with respect to direction s^{i-1} and a_1 is strictly above (below) a_2 with respect to direction e_i, then a_1 is strictly above (respectively, below) a_2 with respect to direction s^i.

(iii) If a_1 is is at the same height as a_2 with respect to both directions s^{i-1} and e_i, then a_1 and a_2 are at the same height with respect to direction s^i.

Since s^{i-1} provided a total order of $\pi_{i-1}(A)$ by assumption and given the statements above, we conclude that s^i totally orders $\pi_i(A)$. Suppose that after the loop terminates, $s^d = \texttt{PointIso}(P)$, totally orders the points of $\pi_d(A)$. Then, since $\pi_d(A) = A$, the final direction totally orders the points of A. Finally, by our initial runtime analysis, the loop terminates and thus, we see that Algorithm 3.8 is correct. □

Using the previous lemma, we are now able to construct a set of directions that is vertex-isolating with respect to a simplicial complex. The output of Algorithm 3.8 is vertex-isolating by definition, so we immediately have the following corollary.

Corollary 3.25 (Constrution of Step 1 Directions)**.** *Let $K \subset \mathbb{R}^d$ be a GP-immersed simplicial complex. Then the basis directions $e_1, e_2, \ldots, e_d$, along with the direction that is the output of* $\texttt{PointIso}(K_0)$*, form a set that is vertex-isolating.*

3.25.1 Directions for Higher-Dimensional Simplices

Next, we show how auxiliary constructions of Section 3.19.2 can be used to construct sets of directions that faithfully represent all higher-dimensional simplices. If a simplex σ is less than $(d-1)$-dimensional, the direction returned by $\texttt{Perp}(\sigma)$ is not guaranteed to be K_0-perpendicular to σ. That is, other vertices may have the same height as σ with respect to this direction. Thus, to ensure we have a direction that places σ at a unique height, we "pop" off any of these extra vertices using `TiltToPop`, returning a tilted direction that is guaranteed to be K_0-perpendicular to σ.

Lemma 3.26 (Construction of Step 2a. Directions)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let σ be a maximal simplex of K. Let $s = \texttt{Perp}(\sigma)$, H denote the set of all vertices with the same height as σ in direction s, $W = \texttt{PlaneFill}(H, s) \cup (H \setminus \sigma)$, and $V = W \cup \sigma$. Then $\texttt{TiltToPop}(K_0 \cup W, V, W, s)$ is K_0-perpendicular to σ. Furthermore, this direction can be computed in $O((n_0 + d)\log(n_0 + d) + d^3)$.*

Proof. We first assert that the inputs of $\texttt{TiltToPop}(K_0 \cup W, V, W, s)$ are valid; note that, by construction, V contains $d-1$ points in general position, meaning that no other points have the same height as V with respect to s. Thus, s is $(K_0 \cup W)$-perpendicular to V. Also by construction, we have $W \subseteq V \subseteq K_0 \cup W$. Then, by Lemma 3.23, the direction returned by $\texttt{TiltToPop}(K_0, V, W, s)$ is a (K_0, V, W, s)-perturbation. In particular, by Statement (i) of

Definition 3.8, this means the direction is K_0-perpendicular to $V \setminus W = \sigma$, as was desired.

Finding the set H takes $\Theta(n_0)$ time. We compute $s = \texttt{Perp}(\sigma)$ in $\Theta(d^3)$ time by Lemma 3.20. We compute $W = \texttt{PlaneFill}(H, s) \cup (H \setminus \sigma)$, in $\Theta(d^3)$ time by Lemma 3.21. Finally, we compute $\texttt{TiltToPop}(K_0 \cup W, V, W, s)$. $\texttt{TiltToPop}$ has runtime $\Theta((n_0 + |W|)\log(n_0 + |W|) + d^3)$ by Lemma 3.23. Since W has size $O(d)$, all these operations in total take $O((n_0 + d)\log(n_0 + d) + d^3)$ time. □

Finally, we show that Construction 1 indeed forms a faithful discretization of the APHT and analyze its size and time complexity.[2]

Theorem 3.27 (Explicit Faithful Discretization)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$ such that $\dim(K) < d$, and let S be the set of directions from Construction 1. Then, $\hat{\rho}(K, S)$ is a faithful discretization of size $O(n2^\kappa + d)$, and S can be computed in $O(\log(n_0 + d)(d^2 n_0^d + nd + nn_0 2^\kappa) + n2^\kappa d^3)$ time.*

Proof. By Lemma 3.24, the directions added to S in Step 1 are vertex-isolating (Definition 3.9). By Lemma 3.26, the directions added to S in Step 2a. are K_0-perpendicular to every maximal simplex, satisfying the first condition of being simplex-isolating (Definition 3.10(a)). By Lemma 3.23, the directions added to S in Step 2b. are (P, V, W, s)-perturbations, satisfying the second condition of being simplex-isolating (Definition 3.10(b)). Thus, the directions of Step 2 are simplex-isolating. Since (K, S) is both vertex- and simplex-isolating, by Theorem 3.12, (K, S) is a faithful discretization.

[2]Notice that we have a set of requirements that leads to a set of directions. It is possible that the set of directions has a smaller cardinality than the set of requirements, namely, if a direction satisfies two or more requirements. However, since we are considering an upper bound, we count directions assuming each direction satisfies just one requirement.

Now, we analyze size and time bounds. By Lemma 3.24, the $\Theta(d)$ directions added to S in Step 1 can be computed in time $\Theta(d^2 n_0^d \log n_0)$.

Next, we give bounds for Step 2a.. By Lemma 3.26, for each maximal $\sigma \in K$, the direction s_σ in Step 2a. can be computed in time $O((n_0+d)\log(n_0+d)+d^3)$. Since the total number of maximal simplices is $O(n)$, the total time required to compute the directions in Step 2a. is $O\left(n((n_0+d)\log(n_0+d)+d^3)\right)$.

Given s_σ and $\tau \prec \sigma$, by Lemma 3.23, a single direction in Step 2b. can be computed in time $\Theta(n_0 \log n_0 + d^3)$. Since for every maximal i-simplex of K, we compute one direction for each of its proper faces, each i-simplex adds a total of $2^{i+1}-2$ directions in Step 2b.. Letting $\kappa = \dim K$, the number of directions in Step 2b. is $\Theta(n2^\kappa)$. Hence, the total time required to compute the directions in Step 2b. is $O(n2^\kappa(n_0 \log n_0 + d^3))$.

Thus, combining the results above, we find that the set S has $O(n2^\kappa + d)$ directions and can be computed in time $O(\log(n_0+d)(d^2 n_0^d + nd + nn_0 2^\kappa) + n2^\kappa d^3)$. □

3.27.1 Example of Building a Faithful Set

Here, we walk through a small example of building a faithful discretization of the APHT, following Construction 1. Suppose we are given the simplicial complex K in $\mathbb{R}^3$, shown in Figure 3.5. Specifically, K has: five vertices, $v_0 = (1,0,0)$, $v_1 = (0,1,0)$, $v_2 = (0,0,1)$, $v_3 = (0,2,0)$, and $v_4 = (0,2,2)$; four edges, $[v_0, v_1]$, $[v_1, v_2]$, $[v_0, v_2]$, and $[v_0, v_3]$; and the single two-simplex $[v_0, v_1, v_2]$. There are three maximal simplices, v_4, $\sigma_1 = [v_0, v_3]$, and $\sigma_2 = [v_0, v_1, v_2]$.

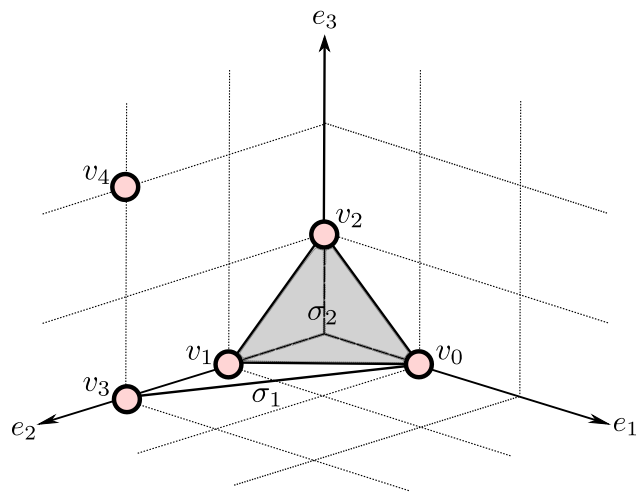


Figure 3.5: The simplicial complex K used as an example in this section. We first find vertex-isolating directions: the basis vectors, along with a fourth direction that orders the filtration grid of K_0 with respect to the basis vectors uniquely. Next, for σ_1 and σ_2, we find simplex-isolating directions; directions that are K_0-perpendicular to σ_1 and σ_2, as well as directions that tilt these perpendicular directions so that proper faces of σ_1 (and σ_2) are "popped" above σ_1 (respectively, σ_2).

3.27.2 Tilt Examples

Since many algorithms used in Construction 1 call Algorithm 3.6 (`Tilt`), we walk through two examples in this section, using inputs that are relevant for later subsections.

We first describe the details of the call $\texttt{Tilt}(s, s', P)$ where $s = (1, 0, 0)$, $s' = (0, 1, 0)$ and $P = \{0, 1, 2\}^3 \subset \mathbb{R}^3$. We begin by finding the heights of points in P with respect to the direction s that are closest but distinct (Line 1 of Algorithm 3.6). There are only three distinct heights of points of P with respect to the direction s, namely 0, 1, and 2. Choosing the first pair that satisfies our desired property, we have $h_1 = 0$ and $h_2 = 1$. Next, on Line 2, we find heights of points of P in the direction s' that are extremal, which are $h'_1 = 0$ and $h'_2 = 2$. Then, on Line 3, we compute the solution to $(1-t)h_1 + th'_2 = (1-t)h_2 + th'_1$, which

is $t_* = 1/3$. Finally, we return the tilted direction

$$s_t = \left(1 - \frac{1}{2}t_*\right) e_1 + \frac{1}{2}t_* e_2 = \left(\frac{5}{6}, \frac{1}{6}, 0\right) \approx (0.8333, 0.1667, 0).$$

Next, we detail a more involved instance of `Tilt`, using $s = (2\sqrt{5}/5, \sqrt{5}/5, 0)$, $s' = (0.8, 0.4, 0.4472)$, and $P = K_0$. First, on Line 1, we identify two heights of vertices with respect to s that are closest. The direction s orders vertices in an evenly spaced way (v_0, v_3, and v_4 all have height $2\sqrt{5}/5$, v_1 has height $\sqrt{5}/5$, and v_2 has height 0). Going through the sorted list, we first encounter $h_1 = 0$ and $h_2 = \sqrt{5}/5$. On Line 2, we find the minimum and maximum heights with respect to s'. These are $h'_1 = 0.4$ (the height of v_1 with respect to s') and $h'_2 \approx 1.6944$ (the height of v_4 with respect to s'). Finally, on Line 3 of Algorithm 3.6, we find the solution to $(1-t)h_1 + th'_2 = (1-t)h_2 + th'_1$, which is $t_* \approx 0.2568$ This gives us the final output of `Tilt`(s, s', K_0), the direction

$$s_t = \left(1 - \frac{1}{2}t_*\right) s + \frac{1}{2}t_* s' \approx (0.8823, 0.4412, 0.0574).$$

3.27.3 Vertex Isolating Directions

First, we find the directions of Step 1 of Construction 1. Specifically, we find directions that are vertex-isolating with respect to K. By Lemma 3.25, the three standard basis directions, e_1, e_2, e_3, as well as a fourth direction as computed in `PointIso`(K_0) (Algorithm 3.8) are vertex-isolating with respect to K. Thus, we walk through the computation of Algorithm 3.8. First, on Line 1, we begin by defining A, the filtration

grid of K_0 with respect to $\{e_1, e_2, \ldots, e_d\}$. In our example, A is a set of 27 gridpoints, $\{0,1,2\}^3 \subset \mathbb{R}^3$. In Line 2, we initialize $s = (1,0,0)$, and enter the loop on Lines 3–5. The first iteration calls $\texttt{Tilt}((1,0,0),(0,1,0),A)$ (Algorithm 3.6), tilting e_1 towards e_2. As detailed in Section 3.27.2, this gives us $s^2 = (1-\frac{1}{2}t_*)e_1 + \frac{1}{2}t_* e_2 \approx (0.8333, 0.1667, 0)$. We are then ready for the next and final iteration of the loop in Algorithm 3.8, finding the fourth direction in our set of vertex-isolating directions, which is

$$s = \texttt{Tilt}(s^2, (0,0,1), A) \approx (0.8013, 0.1603, 0.0385).$$

3.27.4 Perpendicular Directions

Next, we compute directions as described in Step 2a. of Construction 1 (equivalently, directions satisfying Statement (a) of Definition 3.10). In the language of Definition 3.10, for σ_i $(i = 1,2)$ we need a direction s_{σ_i} that is K_0-perpendicular to σ_i. We begin with the two-simplex, σ_2, the more straightforward computation. First, we compute $\texttt{Perp}(\sigma_2)$ (Algorithm 3.4). In Line 1 of Algorithm 3.4, we find a matrix A with $v_i - v_0$ in the ith column, which, for the two-simplex σ_2 is

$$A = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Next, in Line 2, we use a QR-decomposition to find the basis vectors for the space spanned by A, which are, $b_1 = (-\sqrt{2}/2, \sqrt{2}/2, 0)$ and $b_2 = (-\sqrt{6}/6, -\sqrt{6}/6, \sqrt{6}/3)$. Then, in Lines 5–

8, we perform Gram-Schmidt on b_1 and b_2 with each basis vector until the result is non-zero. The loop terminates after one iteration and we find

$$s_{\sigma_2} = \texttt{Perp}(\sigma_2) = (\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3) \approx (0.5774, 0.5774, 0.5774).$$

The direction s_{σ_2} is already K_0-perpendicular to σ_2 since σ_2 has codimension zero; we confirm this by completing the procedure suggested in Step 2a. of Construction 1. Let H be the set of all vertices with the same height as σ_2 in direction s_{σ_2}, meaning that $H = \emptyset$. Then, $W = \texttt{PlaneFill}(H, s) \cup (H \setminus \sigma_2) = \emptyset$ and $V = W \cup \sigma_2 = \sigma_2$, so $\texttt{TiltToPop}(K_0, V, W, s_{\sigma_2}) = \texttt{TiltToPop}(K_0, \sigma_2, \emptyset, s_{\sigma_2}) = s_{\sigma_2}$ (this is simply highlighting the fact that, since the output of $\texttt{Perp}(\sigma_2)$ is already K_0-perpendicular to σ_2, we do not need to adapt it in any way).

Next, we shift focus to σ_1. To construct the direction described by Step 2a. of Construction 1 for the edge σ_1, we again begin by finding $\texttt{Perp}(\sigma_1)$. This begins with the column matrix containing $v_3 - v_0$,

$$A_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

Then, we use QR-decomposition to find the basis for the space spanned by A_1, which consists of the single vector $b = (-\sqrt{5}/5, 2\sqrt{5}/5, 0)$. Performing Gram-Schmidt on b and e_1 yields the direction $s = \texttt{Perp}(\sigma_1) = (2\sqrt{5}/5, \sqrt{5}/5, 0)$. However, note that although s is perpendicular to σ_1, it is not K_0-perpendicular, since the vertex v_4 is at the same

height as σ_1 with respect to s. Thus, we correct s by following the procedure suggested in Construction 1, Step 1. First, let H_1 be the set of all vertices with the same height as σ_1 in direction s, i.e., $H_1 = \{v_0, v_3, v_4\}$. Since $\text{aff}(H_1)$ has codimension one with $\mathbb{R}^3$, $\texttt{PlaneFill}(H_1, s) = \emptyset$, so that

$$\begin{aligned} W_1 &= \texttt{PlaneFill}(H_1, s) \cup (H_1 \setminus \sigma_1) \\ &= \emptyset \cup (\{v_0, v_3, v_4\} \setminus \{v_0, v_3\}) \\ &= \{v_4\}. \end{aligned}$$

Finally, setting $V_1 = W \cup \sigma_1 = \{v_0, v_3, v_4\}$, we use $\texttt{TiltToPop}(K_0 \cup W_1, V_1, W_1, s)$ (Algorithm 3.7) to find a direction K_0-perpendicular to σ_1. In Line 1, we find $V' = \texttt{PlaneFill}(V_1, s)$ but since $\text{aff}(V_1)$ has codimension one with $\mathbb{R}^3$, we simply have $V' = \emptyset$. Next, in Line 2, we compute

$$\begin{aligned} s' &= \texttt{Perp}(\text{aff}(V' \cup (V_1 \setminus W_1) \cup \{W_1 - s\})) \\ &= \texttt{Perp}(\text{aff}(v_0, v_3, v_4 - s)) \\ &= \texttt{Perp}(\text{aff}(v_0, v_3, (-2\sqrt{5}/5, 2 - \sqrt{5}/5, -2))) \\ &\approx (0.8, 0.4, 0.4472). \end{aligned}$$

Although for this example, s' happens to be already K_0 perpendicular to σ_1, this is not generally the case. Thus, we continue on Line 6 and compute $s_{\sigma_1} = \texttt{Tilt}(s, s', K_0 \cup V') = \texttt{Tilt}(s, s', K_0) \approx (0.8823, 0.4412, 0.0574)$ using Algorithm 3.6 (see Section 3.27.2 for details),

which is therefore also the output of $\texttt{TiltToPop}(K_0 \cup W, V, W, s)$, and a direction guaranteed to be K_0-perpendicular to σ_1.

3.27.5 Simplex-Isolating Directions

Next, we find the directions of Step 2b. in Construction 1 (directions satisfying Statement (b) of Definition 3.10). That is, for every nonempty proper face of σ_1 and σ_2, we need a direction that pops the face above the remaining vertices of the maximal simplex. We walk through the procedure for a single face of σ_2; the remaining computations are similar and are summarized in Table 3.1.

Consider the vertex $[v_0] \prec \sigma_2$. By Theorem 3.27 Step 2, we can use the output of $\texttt{TiltToPop}(K_0, \sigma_2, v_0, s_{\sigma_2})$ (Algorithm 3.7) to find a direction that pops v_0 above $\sigma_2 \setminus v_0 = [v_1, v_2]$. First, in Line 1, we find $V' = \texttt{PlaneFill}(\sigma_2, s_{\sigma_2})$ using Algorithm 3.5. Since σ_2 has codimension one in $\mathbb{R}^3$, this is simply $V' = \emptyset$. Next, on Line 2 of Algorithm 3.7 we compute

$$\begin{aligned} s' &= \texttt{Perp}(\text{aff}(V' \cup (\sigma_2 \setminus v_0) \cup \{v_2 - s_{\sigma_2}\})) \\ &= \texttt{Perp}(\text{aff}(\{v_1, v_2\} \cup \{v_0 - s_{\sigma_2}\})) \\ &\approx (0.9636, 0.1890, 0.1890). \end{aligned}$$

Since v_0 is above $[v_1, v_2]$ with respect to s', we do not enter the IF loop on Line 3. Finally, we use Algorithm 3.6 and compute $s = \texttt{Tilt}(s_{\sigma_2}, s', K_0) \approx (0.6598, 0.4944, 0.4944)$, which is therefore also the output to $\texttt{TiltToPop}(K_0, \sigma_2, v_0, s_{\sigma_2})$.

Table 3.1: Directions that are vertex- and simplex-isolating for the simplicial complex of Figure 3.5, computed as described in Construction 1.

Property	Description	Computed Directions (rounded to four decimal places)
Definition 3.9	vertex-isolating	(1,0,0)
		(0,1,0)
		(0,0,1)
		(0.8013, 0.1603, 0.0385)
Statement (a)	K_0-perpendicular to σ_2	$s_{\sigma_2} = (0.5774, 0.5774, 0.5774)$
of Definition 3.10	K_0-perpendicular to σ_1	$s_{\sigma_1} = (0.8823, 0.4412, 0.0574)$
Statement (b)	$(K_0, \sigma_2, [v_0], s_{\sigma_2})$-perturbation	(0.6598, 0.4944, 0.4944)
of Definition 3.10	$(K_0, \sigma_2, [v_1], s_{\sigma_2})$-perturbation	(0.5357, 0.6187, 0.5357)
	$(K_0, \sigma_2, [v_2], s_{\sigma_2})$-perturbation	(0.5357, 0.5357, 0.6187)
	$(K_0, \sigma_2, [v_0, v_1], s_{\sigma_2})$-perturbation	(0.5953, 0.5953, 0.4866)
	$(K_0, \sigma_2, [v_0, v_2], s_{\sigma_2})$-perturbation	(0.5959, 0.4835, 0.5959)
	$(K_0, \sigma_2, [v_1, v_2], s_{\sigma_2})$-perturbation	(0.5235, 0.5880, 0.5880)
	$(K_0, \sigma_1, [v_0], s_{\sigma_2})$-perturbation	(0.5561, 0.5560, 0.5866)
	$(K_0, \sigma_1, [v_3], s_{\sigma_2})$-perturbation	(0.5649, 0.5720, 0.5844)

Summary of Computations In Table 3.1, we see the complete list of computed directions, including those that were discussed in detail previously in this section.

3.28 Discretization is Faithful for Other Topological Transforms

We have so far focused on a faithful discretization of the APHT. However, there are many other types of topological descriptors that could be used to summarize a shape. The ECCT was introduced at the same time as the PHT in [61], and was also shown to be faithful (see also [20, 34]). Since then, many papers in applications have sampled the (A)ECCT [2, 24, 43] or (A)BCT [55, 57]. In this section, we show that a direction set that is vertex- and simplex-isolating for a simplicial complex K also gives us faithful discretizations of the ABCT and other augmented dimension-returning transforms, as well as the AECCT.

3.28.1 Faithful Discretization of the ABCT/Other Augmented Dimension-Returning Transforms

Once the vertex set is known, the reconstruction algorithm uses limited information from the persistence diagrams: namely, the height and dimension of each simplex in a given direction.[3] Thus, our results can be extended to any topological descriptor that contains this information, which we call *augmented dimension-returning transforms*, including the ABCT. We summarize this observation in the following corollary.

Corollary 3.29 (A Faithful Discretization of the ABCT and Other Augmented Dimension-Returning Transforms)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $S \subset \mathbb{S}^{d-1}$ be a direction set that satisfies the assumptions of Theorem 3.12. Then, the set S*

[3] In fact, if the vertex locations are known, then we are only using the total order of the vertices (with respect to the directions in S). Specifically, for each direction s and each vertex $v \in K_0$, we define an equivalence relation on K such that the equivalence class for v is: $[v]_s := \{\sigma \in K \text{ s.t. } h_s(\sigma) = h_s(v)\}$. The set of equivalence classes is totally orded by the height in direction s.

parameterizes a faithful discretization of the ABCT and other augmented dimension-returning transforms for K.

3.29.1 Faithful Discretization of the AECC

In practice, Augmented Euler Characteristic Curves ($\hat{\chi}$'s) are often preferred to $\hat{\rho}$'s due to the existence of faster algorithms for computing the curves. For instance, [52] gives an algorithm to compute χ in linear time. This faster computation time makes the (A)ECCT a good canidate for processing large amounts of data. For instance, in [2], CT scans of barley seeds and spikes are considered in discrete "slices" with respect to 158 directions, creating a sampling of the associated Euler characteristic curves and therefore also sampling the ECCT. Through this sampling, they are able to use machine learningt o distinguish between 28 barley phenotypes, agreeing with a biologically based classification.

However, unlike $\hat{\rho}$'s or $\hat{\beta}$'s, $\hat{\chi}$'s are not dimension-returning. In particular, this means that we cannot directly compute k-indegree as in Algorithm 3.1 when using $\hat{\chi}$'s. However, as this section shows, we can adapt the notion of k-indegree to a notion of even- and odd-indegree. We will then show how to use even- and odd-indegree to reconstruct all higher-dimensional simplices, therefore showing that a set of directions that is vertex- and simplex-isolating gives us a faithful discretezation of the AECCT. That is, we will show that the set of directions described in Section 3.8.1 that formed a faithful discretization of the APHT also form a faithful discretization of the AECCT.

The remainder of this section serves to prove this claim (Theorem 3.34). Our main tool is the following adaptation of the definition of k-indegree to even/odd-indegree, where,

rather than counting simplices of a particular dimension, we count even-dimensional (or odd-dimensional) simplices.

Definition 3.30 (Even/Odd-Indegree for Simplex)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$ and $\sigma \in K$. Let $s \in \mathbb{S}^{d-1}$ be a direction K_0-perpendicular to aff(σ). Then, the* even-indegree of σ in direction s *(respectively,* odd-indegree *of σ in direction s) is the number of even-dimensional (respectively, odd-dimensional) cofaces of σ that have the same height as σ with respect to the direction s.*

Just as we did with the notion of k-indegree, we will need to use an inclusion-exclusion type calculation to find the even/odd-indegree of a simplex. The following lemma helps us prove the correctness of this calculation.

Lemma 3.31 (Even/Odd-Indegree Contributers)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $\tau, \sigma, \sigma' \in K$ such that $\tau \prec \sigma$ and $\dim(\sigma')$ is even (respectively, odd). Let s be K_0-perpendicular to σ and let s_τ be a (K_0, σ, τ, s)-perturbation. Then, σ' contributes to the even-indegree (respectively, odd-indegree) of τ in direction s_τ if and only if σ' is at the same height as σ in direction s_τ and $\tau = \sigma \cap \sigma'$.*

We omit a proof, as it is identical to the proof of Lemma 3.15. We now present Algorithm 3.9, which computes even-indegree. The case for computing odd-indegree is nearly identical. We assert the correctness Algorithm 3.9 in the following theorem.

Theorem 3.32 (Computing Even/Odd-Indegree)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $\sigma \in K$ and $s \in \mathbb{S}^{d-1}$ such that s is K_0-perpendicular to σ. Then, for $k > \dim(\sigma)$,* `EvenIndegree`$(\sigma, s, \widehat{\chi}(K, S))$ *returns the even-indegree of σ in direction s.*

Algorithm 3.9: `EvenIndegree`$(\sigma, s, \widehat{\chi}(K,S), T = \{\})$

Input: $\sigma \in K$, $s \in S$ s.t. s is K_0-perpendicular to σ, $\widehat{\chi}(K,S)$, where S satisfies the assumptions of Theorem 3.12, and a table T for memoization.

Output: the even-indegree for σ

1: $c \leftarrow$ height of σ in direction s

2: `allEven` $\leftarrow$ number of even-dimensional events in $\widehat{\chi}(h_s)$ at height c

3: `doubleCounts` $\leftarrow 0$

4: **for** $\tau \prec \sigma$ in non-descending order by dimension **do**

5: $W \leftarrow \text{verts}(\sigma) \setminus \text{verts}(\tau)$

6: $s_\tau \leftarrow$ a $(K_0, \text{verts}(\sigma), W, s)$-perturbation

7: **if** $T[\tau]$ was not computed yet **then**

8: $T[\tau] \leftarrow$ `EvenIndegree`$(\tau, s_\tau, \widehat{\chi}(K,S), T)$

9: **end if**

10: `doubleCounts` $\leftarrow$ `doubleCounts` $+ T[\tau]$

11: **end for**

12: **return** `allEven` $-$ `doubleCounts`

Proof. The proof is similar to the proof of Theorem 3.16. We prove the claim inductively on $j = \dim(\sigma)$ and let $h_s : K \to \mathbb{R}$ denote the filter function for direction s. For the base case ($j = 0$), consider $\sigma = [v]$, a single zero-simplex. Since s is K_0-perpendicular to σ, the even-indegree of v is exactly equal to the number of even-simplices that have height $s \cdot v$, which can immediately be read off of $\hat{\chi}(s)$. In Algorithm 3.9, notice that since the vertex v does not have any proper faces, we do not enter the loop that starts on Line 4. Thus, the

return value is exactly the number of even simplices with height $s \cdot v$.

For the inductive assumption, let $j \geq 0$. We assume that Algorithm 3.9 returns the even-indegree of τ in direction s, for all $\tau \in K_j$. For the inductive step, let $\sigma \in K_{j+1}$. Now, we compute the even-indegree of σ in direction s. Let F_σ denote the set of even-dimensional simplices with height $h_s(\sigma)$ in direction s, whose cardinality is reported in $\hat{\chi}(s)$. Let $\sigma' \in F_\sigma$, and let $\tau \prec \sigma$. Suppose that s_τ is a $(K_0, \mathrm{verts}(\sigma), \mathrm{verts}(\sigma \setminus \tau), s)$-perturbation. By Lemma 3.31, σ' contributes to the even-indegree of τ in direction s_τ if and only if $\tau = \sigma \cap \sigma'$.

Then, we can isolate each face of $\tau \prec \sigma$, identify the even-indegree of τ by the inductive assumption, and add or subtract it from the even-indegree of σ, alternating by dimension of τ. This ensures that no coface of $\tau \prec \sigma$ adds to the even-indegree of σ. Formally, this is seen in the equation for the even-indegree of σ

$$\delta - \sum_{\tau \prec \sigma} \delta_\tau, \tag{3.6}$$

where δ_τ is the even-indegree of τ in the corresponding tilted directions. In Algorithm 3.9, `allEven` is equal to δ, and the values δ_τ are computed in Line 8. Thus, the return value of Algorithm 3.9 matches the value of Equation (3.6). □

Theorem 3.33 (Correctness of AECC Simplex Certificate)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$. Let $V \subseteq K_0$. Then, Algorithm* 3.10 *returns* TRUE *iff the vertex set V defines a $(|V| - 1)$-simplex of K.*

Algorithm 3.10: $\texttt{AECCCertificate}(V, K_{|V|-2}, \widehat{\chi}(K,S))$

Input: $V \subset K_0$, a vertex set, $K_{|V|-2}$, the $(|V|-2)$-skeleton of K and $\widehat{\chi}(K,S)$, where S is simplex-isolating (Definition 3.10).

Output: TRUE if V defines a $(|V|-1)$-simplex of K, FALSE otherwise

1: $w \leftarrow$ a vertex of V

2: **if** S has a direction s_U that is K_0-perpendicular to V and a direction s_L that is a $(K_0, V, \{w\}, s)$-perturbation **then**

3: **if** $|V|-1$ is even **then**

4: $\texttt{knownEven}_\texttt{U}, \texttt{knownEven}_\texttt{L} \leftarrow$ the number of even simplices of $K_{|V|-2}$ that contribute to the even-indegree of V in directions s_U and s_L, respectively

5: $x \leftarrow |(\texttt{EvenIndegree}(V, s_U, \widehat{\chi}(K,S)) - \texttt{knownEven}_\texttt{U}) - (\texttt{EvenIndegree}(\texttt{V}, \texttt{s}_\texttt{L}, \widehat{\chi}(\texttt{K},\texttt{S})) - \texttt{knownEven}_\texttt{L})|$ ▷ Algorithm 3.9

6: **else**

7: $\texttt{knownOdd}_\texttt{U}, \texttt{knownOdd}_\texttt{L} \leftarrow$ the number of even simplices of $K_{|V|-2}$ that contribute to the odd-indegree of V in directions s_U and s_L, respectively

8: $x \leftarrow |(\texttt{OddIndegree}(V, s_U, \widehat{\chi}(K,S)) - \texttt{knownOdd}_\texttt{U}) - (\texttt{OddIndegree}(\texttt{V}, \texttt{s}_\texttt{L}, \widehat{\chi}(\texttt{K},\texttt{S})) - \texttt{knownOdd}_\texttt{L})|$ ▷ Odd Algorithm 3.9

9: **end if**

10: **return** TRUE if $x \geq 1$

11: **else**

12: **return** FALSE

13: **end if**

Proof. We must show that `AECCCertificate`$(V, K_k, \widehat{\chi}(K, S))$ returns TRUE iff V defines a simplex. By Lemma 3.11, there are directions as described in Line 2 for every set of vertices V and vertex $w \in V$ where V forms a simplex of K. Thus, if no such directions exist for V in our set of directions, we immediately know V does not form a simplex, and return FALSE on Line 12. Suppose, then, that the directions described in Line 2 exist for V and $w \in V$. Using identical reasoning as in the proof of Theorem 3.17, we know that V and only V lies in the wedge $\mathbb{W} := H^{\downarrow}(s_U, c) \triangle H^{\downarrow}(s_L, c')$.

We then show that the even/odd-indegree can be used to determine if V defines a simplex in K. First, consider the case that $|V| - 1$ is even. Since we have $w \in H^{\downarrow}(s_U, c_U)$ and $w \notin H^{\downarrow}(s_L, c_L)$, every simplex that contributes to the even-indegree of $V \setminus \{w\}$ in direction s_U also contributes to the even-indegree of $V \setminus \{w\}$ in direction s_L. In addition, the only potential simplices contributing to the even-indegree of $V \setminus \{w\}$ in direction s_L that do not contribute to the even-indegree in direction s_U are the even-simplices defined by V or that have V as a face. To avoid lower-dimensional simplices contributing to the total count on Line 5, we subtract $\texttt{knownEven}_{\texttt{U}}$ and $\texttt{knownEven}_{\texttt{L}}$ from the even-indegrees of V in directions s_U and s_L, respectively. Thus, if V defines a maximal $(|V| - 1)$-simplex, the value returned on Line 5 is one. If V defines a $(|V| - 1)$-simplex that is the face of a higher dimensional even simplex, then the value returned on Line 5 is greater than one. If V does not define a simplex of K, this value will be zero. The case when $|V| - 1$ is odd is nearly identical. Thus, Algorithm 3.10 returns TRUE iff V defines a $(|V| - 1)$-simplex of K. $\square$

The last step of the full process of reconstruction is to note that vertex-reconstruction only relied on the presence of events, not their dimension. Therefore, vertex-isolating directions can be used with any augmented descriptor to reconstruct K_0. Finally, we can state the following main result.

Theorem 3.34 (Sufficient Conditions for Representation)**.** *Let K be a simplical complex GP-immersed in $\mathbb{R}^d$ such that $\dim(K) = \kappa < d$, and let $S \subset \mathbb{S}^{d-1}$ such that (K, S) is vertex- and simplex-isolating. Then, $\hat{\chi}(K, S)$ is a faithful discretization of $AECCT(K)$.*

The theorem is proven by Algorithm 3.11. We omit a proof of correctness for Algorithm 3.11, as it is nearly identical to Theorem 3.18.

3.35 Stability

In this section, we make observations about the stability of our discretization. We are interested specifically in the stability of finite discretizations of the APHT. All proofs are given with reference to the APHT; stability results for the ABCT and other augmented dimension-returning transforms, as well for the AECCT are given as corollaries.

Many important observations related to stability can be defined in terms of a stratifcation of the sphere induced by topological transforms, as has been noted in related work, e.g., [20,42]. Additionally, the strata have nice combinatorial properties that allow for a sheaf- or cosheaf-theoretic interpretation. The sheaf/cosheaf viewpoint has been championed by many, e.g., [3, 20, 31] (see also Chapter 5). The stability of the entire PHT (using height filtrations) is stated in [3, Theorem 4.2] in terms of interleaving distance between sheaves. Roughly, the stratification is defined by dividing the sphere of directions into regions (strata)

Algorithm 3.11: `AECCReconstructComplex`($\widehat{\chi}(K, S)$)

Input: $\widehat{\chi}(K, S)$, where S is vertex- and simplex-isolating (Definitions 3.9 and 3.10).

Output: simplicial complex K.

1: $K_0 \leftarrow$ vertices of K, as found using the methods of [8, Theorem 9]

2: **for** $V \subseteq K_0$ with $1 < |V| \leq d$ and in non-decreasing size of V **do**

3: $k \leftarrow |V| - 1$

4: **if** $V \setminus \{v_i\} \in K_{k-1}$ for all $v_i \in V$ **then**

5: **if** `AECCCertificate`(V) == TRUE **then**

6: Add V to K_k

7: **end if**

8: **end if**

9: **end for**

10: **return** $K_0 \cup K_1 \cup \cdots \cup K_\kappa$

where all directions within a particular stratum induce the same partial order on vertices (ordered by their height with respect to that direction) (see Figure 3.6). Further details on stratified spaces are given in Chapter 5. We begin by defining a strata-preserving map, which allows for a clean way to describe important relationships between sets of directions.

Definition 3.36 (Strata-Preserving Map)**.** *Given a set $S \subseteq \mathbb{S}^{d-1}$ and a simplicial complex $K \subset \mathbb{R}^d$, we say that the map $f : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ is* strata-preserving *with respect to K if, for every $s \in S$, the partial order ordering of vertices induced by their heights with respect to direction s and $f(s)$ is the same.*

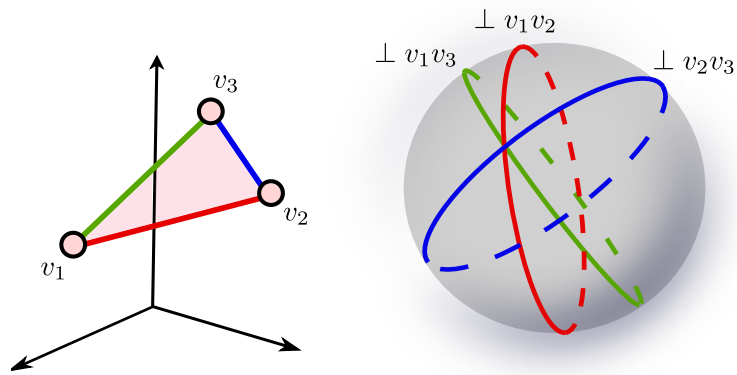


Figure 3.6: A two-simplex in $\mathbb{R}^3$ (left) stratifies $\mathbb{S}^2$ where each stratum is a region containing all directions that define the same partial order on vertices. Notice that the set of directions perpendicular to any pair of vertices forms a great circle and the two directions perpendicular to $[v_1, v_2, v_3]$ correspond to the two three-way intersections of these great circles.

Observing that Theorem 3.12 gives us flexibility for choosing which directions parameterize a faithful discretization, the first stability result shows if we perturb those directions in a strata-preserving way, we wind up with another faithful discretization.

Theorem 3.37 (Discretization is Robust to Perturbing Parameterization)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$, and let $S \subset \mathbb{S}^{d-1}$ be such that (K, S) is vertex- and simplex-isolating. If a function $f\colon \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ satisfies the following conditions:*

(i) there are directions $s_1, s_2, \ldots s_d, s_{d+1} \in S$ so that $f(s_1), f(s_2), \ldots, f(s_d)$ are linearly independent and so that $f(s_{d+1})$ uniquely orders the filtration grid of K_0 with respect to $\{f(s_1), f(s_2), \ldots, f(s_d)\}$; and

(ii) the map f is strata-preserving,

then $\hat{\rho}(K, f(S))$ is a faithful discretization of $APHT(K)$.

Proof. The directions ensured by Statement (i) mean that $f(S)$ is vertex-isolating. We must therefore show that $f(S)$ is simplex-isolating. Let σ be a maximal simplex. Since (K, S) is σ-isolating, S contains a direction s_σ that is K_0-perpendicular to σ. Since f is strata-preserving by Statement (ii), it preserves vertex order, so $f(s_\sigma)$ is also K_0-perpendicular to σ. Moreover, if s' is a (P, V, W, s_σ)-perturbation, then $f(s')$ is a $(P, V, W, f(s_\sigma))$-perturbation. Thus, we see that $(K, f(S))$ is σ-isolating.

Since $(K, f(S))$ is vertex- and simplex-isolating, we have met the assumptions of Theorem 3.12; hence, $\hat{\rho}(K, f(S))$ is a faithful discretization of $APHT(K)$. □

The results of Section 3.28 give us the following corollary.

Corollary 3.38. *Given a map $f : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$ as in Theorem 3.37, if $\hat{\mathrm{D}}$ is any augmented dimension-returning transform, then $\hat{\mathrm{D}}(K, f(S))$ is a faithful discretization of the corresponding topological transform. Notably, $\hat{\beta}(K, f(S))$ is a faithful discretization of $ABCT(K)$. Furthermore, $\hat{\chi}(K, f(S))$ is a faithful discretization of $AECCT(K)$.*

Since vertex heights change continuously as a direction changes continuously (continuous with respect to the standard topologies on $\mathbb{R}$ and $\mathbb{S}^{d-1}$), the next stability result shows that we can allow small edits to the embedding.

Theorem 3.39 (Parameterization is Robust to Vertex Perturbations)**.** *Let K be a simplicial complex GP-immersed in $\mathbb{R}^d$, and let S be as in Construction 1. Let L be a complex that arises by perturbing the vertices of K in such a way so that, for every $s \in S$, the order of vertex heights in L_0 with respect to s is the same order of vertex heights in K_0 with respect*

to s. Then, there exists a set $S' \subset \mathbb{S}^{d-1}$ that differs from S by at most a single direction such that $\hat{\rho}(L, S')$ is a faithful discretization of $APHT(L)$.

Proof. Let $g\colon K \to L$ be the simplicial map determined by the perturbation described in the theorem. Note that this map is a homeomorphism (and hence a bijection).

The original direction set S is simplex-isolating for L. By construction, (K, S) is simplex-isolating. Now, let σ be a maximal simplex. Since (K, S) is σ-isolating, there is a direction $s_\sigma \in S$ that is K_0-perpendicular to σ. Since K and L have the same vertex order with respect to direction s_σ, we know that s_σ is L_0-perpendicular to $g(\sigma)$. Moreover, if s' is a (P, V, W, s_σ)-perturbation, then s' is also a $(g(P), g(V), g(W), s_\sigma)$-perturbation. Hence, $(g(K), S) = (L, S)$ is $g(\sigma)$-isolating. Since g is a bijection, (L, S) is simplex-isolating.

It is not guaranteed that the direction set S is vertex-isolating for L. Suppose that $s_1, s_2, \ldots, s_{d+1} \in S$ are vertex-isolating directions for K, so that $s_1, s_2, \ldots s_d$ are linearly independent, and s_{d+1} orders the filtration grid of K with respect to $\{s_1, s_2, \ldots, s_d\}$ uniquely. Notice that the filtration grid of L with respect to $\{s_1, s_2, \ldots, s_d\}$ is distinct from the filtration grid of K; it is no longer guaranteed that s_{d+1} will order this new filtration grid uniquely. Thus, we may need to replace s_{d+1} with a new direction s'_{d+1} that *does* order the filtration grid of L with respect to $\{s_1, s_2, \ldots, s_d\}$ uniquely.

Thus, denoting the new direction set we obtained by modifying S with this single replacement S', we find that $\hat{\rho}(L, S')$ faithfully discretizes $APHT(L)$, as desired. □

Corollary 3.40. *Given a map $g : K \to L$ as in Theorem 3.39, if $\hat{\mathrm{D}}$ is any augmented dimension-returning transform, then there exists a set $S' \subset \mathbb{S}^{d-1}$ that differs from S by at*

most a single direction such that $\hat{\mathrm{D}}(L, S')$ *is a a faithful discretization of the corresponding topological transform. Notably,* $\hat{\beta}(L, S')$ *is a faithful discretization of* $ABCT(L)$*. Furthermore,* $\hat{\chi}(L, S')$ *is a faithful discretization of* $AECCT(L)$*.*

Note that in the previous theorem and corollary, the set of simplicial complexes L for which the original direction set S does *not* lead to a faithful discretization (that is, the change of the single direction is required) has measure zero with respect to Lebesgue measure.

3.41 Edge Reconstruction and Representation Using a Radial Binary Search

In this section, we improve upon the results of Section 3.8.1 as well as [8] and present an improved algorithm for graph—and, more generally one-skeleton—reconstruction. The improvement comes in reconstructing the edges, where we use a radial binary (multi-)search. The binary search employed takes advantage of the fact that the edges can be ordered radially with respect to a reference plane, a feature unique to graphs. The methods described here can apply to any augmented dimension-returning transform, but for clarity, we focus on $\hat{\rho}$'s.

3.41.1 Tools

We begin with an overview of the definitions and tools specific to this section and method. We write (V, E) for a graph and its vertex and edge sets, and use $n = |V|$ and $m = |E|$. In this section, we require slightly stricter general postition assumptions than those that were used in previous sections.

Assumption 2. *Let* $V \subset \mathbb{R}^d$ *be a finite set with* $d \geq 2$*. We say* V *is in* general position *if the following properties are satisfied:*

(i) Every set of $d + 1$ *points is affinely independent.*

(ii) No three points are colinear after orthogonal projection into the space $\pi(\mathbb{R}^d)$, where $\pi : \mathbb{R}^d \to \mathbb{R}^2$ is the orthogonal projection onto the plane spanned by the first two basis elements, e_1 and e_2.

(iii) Every point has a unique height with respect to the direction e_2.

We call a graph GP-immersed[4] *iff its vertex set is in general position in $\mathbb{R}^d$.*

If only the first two assumptions are met, we can deterministically find a basis that satisfies all three, at an added cost of $\Theta(|P| \log |P| + d + \Pi)$ time.

Lemma 3.42 (Creation of Orthonormal Basis)**.** *Given a point set $P \subset \mathbb{R}^d$ satisfying Assumption 2(i) and Assumption 2(ii), we can use two diagrams and $\Theta(|P| \log |P| + d + \Pi)$ time to create the orthonormal basis $\{b_1, b_2, e_3, e_4, \ldots, e_d\}$ so that all points of P have a unique height in direction b_2.*

Proof. Recall that Algorithm 3.6 (`Tilt`) takes diagrams from two linearly independent directions $s, s' \in \mathbb{S}^{d-1}$, the point set P, and returns a direction s_* in $\Theta(|P| \log |P| + d + \Pi)$ time[5] so that the following properties holds for all $p_1, p_2 \in P$: [6]

(i) If p_1 is strictly above (below) p_2 with respect to direction s, then p_1 is strictly above (below, respectively) p_2 with respect to direction s_*.

[4]Here, we use *immersed* rather than *embedded* in order to allow intersections of edges. Note, however, that this can only happen when $d = 2$.

[5]While Lemma 3.22 does not account for diagram computation time, there are two diagrams used in this process, hence our addition of $\Theta(\Pi)$ to the total runtime.

[6]A proof of correctness is given in Lemma 3.22.

(ii) If p_1 and p_2 are at the same height with respect to direction s and p_1 is strictly above (below) p_2 with respect to direction s', then p_1 is strictly above (respectively, below) p_2 with respect to direction s_*.

(iii) If p_1 is is at the same height as p_2 with respect to both directions s and s', then p_1 and p_2 are at the same height with respect to direction s_*.

We start with the standard basis for $\mathbb{R}^d$, $\{e_1, e_2, \ldots, e_d\}$, and we replace the first two basis elements as follows. Let b_2 be the direction obtained from $\texttt{Tilt}(e_1, e_2, P)$.

By Assumption 2(ii), no three points of P are colinear when projected onto the first two coordinates. In particular, this means no two points share the same heights in both the e_1 and e_2 directions. Then, by Statements (i)-(ii) above, the direction b_2 must order all vertices of P uniquely. Using only the first two coordinates of b_2 and e_1, we then perform Gram Schmidt orthanormalization to compute the first two coordinates of b_1. More precisely, letting $b_i^{(j)}$ denote the jth coordinate of b_i, we compute

$$\begin{pmatrix} b_1^{(1)} \\ b_1^{(2)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} b_2^{(1)} & b_2^{(2)} \end{pmatrix}^T, \begin{pmatrix} 1 & 0 \end{pmatrix}^T \right\rangle}{\left\| \begin{pmatrix} b_2^{(1)} & b_2^{(2)} \end{pmatrix}^T \right\|^2} \begin{pmatrix} b_2^{(1)} \\ b_2^{(2)} \end{pmatrix} \tag{3.7}$$

We then set $b_1^{(j)} = 0$ for $2 < j \leq d$, so that $b_1 \in \text{span}\{e_1, e_2\}$, $b_2 \perp b_1$, and $||b_1|| = 1$. Only considering the first two coordinates of b_2 and e_1 means this process takes constant time. The remaining e_i for $2 \leq i \leq d$ can be used to fill the basis.

Finally, we have a basis satisfying all assumptions of Assumption 2, namely, the computed set $\{b_1, b_2, e_3, e_4, \ldots, e_d\}$. $\square$

Again, we emphasize that, while the preceeding lemma shows that only the first two general position assumptions are necessary, we proceed assuming all items of Assumption 2 are satisfied when using the standard basis, for the sake of simplicity.

3.42.1 Oracle Framework

For simplicity of explanation, we assume an oracle framework. That is, we assume that we have no knowledge of the shape itself, but we have access to an oracle from which we can query directional diagrams. We encourage the reader to consider how the algorithms of Section 3.23.1 could be used in place of the oracle.

Definition 3.43 (Oracle)**.** *For a graph (V, E) GP-immersed in $\mathbb{R}^d$ and a direction $s \in \mathbb{S}^{d-1}$, the operation* `Oracle`(s) *returns the diagram $\hat{\rho}(s)$. We define $\Theta(\Pi)$ to be the time complexity of this oracle query and note that the space complexity of $\hat{\rho}(s)$ is $\Theta(n + m)$. We assume that the data structure returned by the oracle allows queries for specific birth or death values in $\Theta(\log n)$ time (for example, the we could have two arrays of persistence points, one sorted by birth values and one sorted by death values).*

3.43.1 Constructions and Data Structures

In this subsection, we introduce the edge arc object and other definitions useful for computing properties of immersed graphs. Throughout this section, we project points in $\mathbb{R}^d$ to the (e_1, e_2)-plane. As a result, we use "above (below)" without stating with respect to which direction as shorthand for "above (below) with respect to the direction e_2." This direction is intentionally chosen (and used in our GP assumption), as it corresponds to our intuition of above (below) in the figures. Given a vector x, when we write $\measuredangle x$, we mean the

angle that $\pi(x)$ makes with the positive e_1 axis.

Given a direction s and a vertex in a graph immersed in $\mathbb{R}^d$, we classify each edge (v, v') as either an "incoming" edge, when v' is below v with respect to s, or an "outgoing" edge, when v' is above v with respect to s. Note that all incoming edges have the same height as the vertex with respect to the e_2 direction. The following lemma relates the number of edges at a given height to points in the diagram.

Lemma 3.44 (Edge Count)**.** *Let (V, E) be a graph and let $c \in \mathbb{R}$. Let $f\colon V \sqcup E \to \mathbb{R}$ be a filter function. Then, the edges in E with a function value of c are in one-to-one correspondence with the following multiset of points in $\hat{\rho}(f)$, the diagram corresponding to f:*

$$\big\{(b, d) \in \hat{\rho}_1(f) \ \ s.t. \ \ b = c\big\} \cup \big\{(b, d) \in \hat{\rho}_0(f) \ \ s.t. \ \ d = c\big\}. \tag{3.8}$$

In other words, each edge corresponds to either a birth of a one-dimensional homological feature or a death of a zero-dimensional feature in $\hat{\rho}(f)$. This is a generalization of Lemma 3.3 in Section 3.2.1.

If f is a lower-star filtration in direction $s \in \mathbb{S}^{d-1}$, we note that whenever a vertex v has a unique height with respect to a direction s, the cardinality of the multiset above is exactly the one-indegree of v with respect to s. Since we will frequently use one-indegree, we notate this by 1-$\textsc{Indeg}(v, s)$.

Lemma 3.45 (Indegree Computation)**.** *Let (V, E) be a graph GP-immersed in $\mathbb{R}^d$. Let $v \in V$ and let $s \in \mathbb{S}^{d-1}$ such that $s \cdot v \neq s \cdot v'$ for any $v' \neq v \in V$. Then,* 1-$\textsc{Indeg}(v, s)$ *can be computed via the oracle using one diagram and $\Theta(\log n + \Pi)$ time.*

Proof. Let $\hat{\rho} = \texttt{Oracle}(s)$. By the assumption on s, the height of v with respect to the direction s is unique. Hence, we know that any edge at height $c = s \cdot v$ must be incident to v. Thus, by the definition of indegree, an edge has the height c if and only if it contributes to the indegree of v in direction s. Using Lemma 3.44, we count these edges by counting one-dimensional births and zero-dimensional deaths at height c. Since $\hat{\rho}_0$ and $\hat{\rho}_1$ are sorted by both birth and death values (see Definition 3.43) and since $\hat{\rho}$ has $\Theta(n+m)$ points, searching for these events takes $\Theta(\log n + \log m)$. Adding $\Theta(\Pi)$ for the oracle query and recalling that $m = O(n)$, the total runtime is $\Theta(\log n + \Pi)$. □

We conclude this section by introducing a data structure, the *edge arc object*; see Table 3.2 for a summary of the attributes of an edge arc and Figure 3.7 for an example. An

Table 3.2: Attributes of the edge arc object.

EA	**Edge Arc**
v	Vertex around which the edge arc is centered
(α_1, α_2)	Start and stop angles of the arc, with respect to the e_1 direction
verts	Array of vertices in arc radially ordered clockwise in (e_1, e_2)-plane
count	Number of edges incident to v within the arc

edge arc represents the region in the (e_1, e_2)-plane centered at v that is swept out between the two angles α_1 and α_2 (the word 'arc' is referring to the arc of angles between α_1 and α_2, where the angle is measured with respect to the positive e_1 axis). We only consider edge arcs in the upper half-space, with respect to the e_2 direction, so the maximal edge arc is the upper half-plane and the start and stop angles always satisfy $0 \leq \alpha_1 \leq \alpha_2 \leq \pi$. An edge

arc stores an array of vertices sorted radially clockwise about $\pi(v)$ in the (e_1, e_2)-plane in decreasing angle with the e_1-direction. By construction, the first vertex in the array must be closest to α_2 and the last closest to α_1. The edge arc also stores the count of edges of E that have vertices from *verts* as endpoints. In implementation, the angles α_1 and α_2 do not need to be stored directly, but we include them in psuedocode and discussions for clarity.

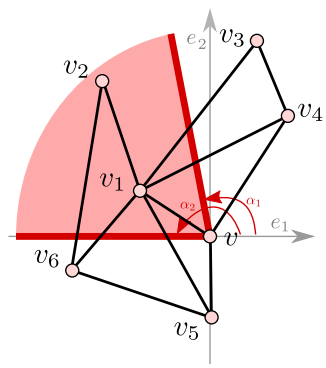


Figure 3.7: An edge arc EA centered at vertex $EA.v = v$. Other attributes of the edge arc include its start and stop angles, $EA.\alpha_1 = 1.75$ radians and $EA.\alpha_2 = \pi \approx 3.14$ radians, the array of vertices $EA.verts = \{v_1, v_2\}$, and the count of edges $EA.count = 1$. Here, we also see that 1-$\textsc{Indeg}(v, e_1) = 1$ and 1-$\textsc{Indeg}(v, e_2) = 2$.

Given some arc EA centered at vertex $v \in V$, we need to be able to compute $EA.count$, the number of edges contained EA that are adjacent to v. The following lemma provides such a computation. We omit a proof because it is a straightforward adaptation of Theorem 3.16 and [8, Lemma 13].

Lemma 3.46 (Arc Count)**.** *Let (V, E) be a graph GP-immersed in $\mathbb{R}^d$. Let EA be an edge arc object, and let $v = EA.v$. Let $s \in \mathbb{S}^{d-1}$ be the direction perpendicular to α_2 so that the arc is entirely below $s \cdot v$. Let E_* denote the edges with height $s \cdot v$ that are not contained*

in EA. If no vertex in V is at the same height as v in direction s, then

$$EA.count = \text{1-INDEG}(v, s) - |E_*|.$$

As an illustration, again, consider Figure 3.7. Consider the direction $s = e^{i(\alpha_2 - \pi/2)}$, which is perpendicular to $e^{i\alpha}$ by construction. In addition, the edge arc is below $s \cdot v$ (specifically, all vertices in $EA.verts$ are below $s \cdot v$). Then, 1-INDEG$(v, s) = 2$ and $E_* = \{[v, v_5]\}$. By Lemma 3.46, $EA.count = \text{1-INDEG}(v, s) - |E_*| = 2 - 1 = 1$. When we say that a list of vertices or edges is sorted clockwise around v, we mean that the list is sorted clockwise (cw) around $\pi(v)$ once projected into the (e_1, e_2)-plane with the largest angle first.

3.46.1 The Reconstruction

In this section, we provide an algorithm to reconstruct a graph (or the one-skeleton of a simplicial complex) using the oracle. We start with an algorithm to find edges, provided vertex locations are known. We end by describing the complete graph reconstruction method.

3.46.2 Fast Edge Reconstruction

In this subsection, we assume we have a graph (V, E), where the vertex set V is known, but E is unknown. Using the oracle and the known vertex locations, we provide a reconstruction algorithm to find all edges in E (Algorithm 3.14). This algorithm is a sweepline algorithm in direction e_2 that, for each vertex processed in the sweep, performs a radial binary multi-search (Algorithm 3.13). This search is enabled by Algorithm 3.12, which splits an edge arc object into two edge arcs, each containing half of the vertices.

In Algorithm 3.12, we find a direction s in the (e_1, e_2)-plane such that half of the vertices

Algorithm 3.12: `SplitArc`$(EA, bigedges, \theta)$

Input: EA, an edge arc; $bigedges$, an array of all edges $(EA.v, v') \in E$ such that $\measuredangle\pi(v' - EA.v) < EA.\alpha_1$; θ, the minimum angle defined by any three vertices in $\pi(V)$

Output: EA_ℓ and EA_r, edge arcs satisfying the properties in Theorem 3.47

1: $n_v \leftarrow |EA.verts|$

2: $mid \leftarrow \lceil \frac{n_v}{2} \rceil$

3: $\alpha \leftarrow \measuredangle\pi(EA.verts[mid] - EA.v) - \theta/2$

4: $s \leftarrow e^{i(\alpha - \frac{\pi}{2})}$

5: $m_\ell \leftarrow$ 1-Indeg$(EA.v, s) - |\{b \in bigedges \mid \measuredangle b < \pi + \alpha\}|$

6: $m_r \leftarrow EA.count - m_\ell$

7: $EA_\ell \leftarrow$ edge arc where $EA_\ell.v = EA.v$, $EA_\ell.\alpha_1 = \alpha$, $EA_\ell.\alpha_2 = EA.\alpha_2$, $EA_\ell.verts = EA.verts[: mid]$, and $EA_\ell.count = m_\ell$

8: $EA_r \leftarrow$ edge arc where $EA_r.v = EA.v$, $EA_r.\alpha_1 = EA.\alpha_1$, $EA_r.\alpha_2 = \alpha$, $EA_r.verts = EA.verts[mid + 1 :]$, and $EA_r.count = m_r$

9: **return** (EA_ℓ, EA_r)

in $EA.verts$ are above v and half are below v with respect to the direction s. This allows us to create a new edge arcs corresponding to each half; see Figure 3.8. The properties of Algorithm 3.12 are described in the following theorem.

Theorem 3.47 (Arc Splitting)**.** *Algorithm 3.12 uses one diagram and* $\Theta(\log n + d + \Pi)$ *time to split* EA *into two new edge arcs* EA_ℓ *and* EA_r *with the properties:*

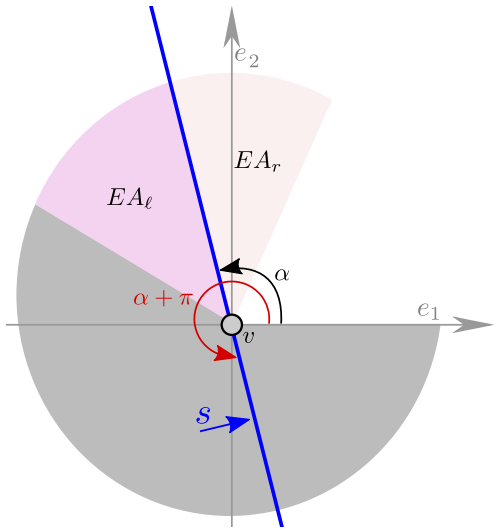


Figure 3.8: The splitting of edge arc EA into EA_ℓ and EA_r, as in Algorithm 3.12. The large gray region is the region containing all edges of *bigedges*. That is, all edges whose angle with the positive e_1-axis is at least $EA.\alpha_1$. On Line 5 of the algorithm, we compute the number of edges in EA_ℓ by first computing the indegree of $EA.v$ in direction s from the diagram in direction s, then we subtract the number of edges in *bigedges* that are below the height of $EA.v$ in direction s (i.e., below the blue line). By the pigeonhole principal, we find $EA_r.count = EA.count - EA_\ell.count$.

(i) Sets $EA_r.verts$ and $EA_\ell.verts$ partition the vertex set $EA.verts$ so vertices in $EA_\ell.verts$ come before those in $EA_r.verts$, with respect to clockwise ordering around $EA.v$.

(ii) $|EA_\ell.verts| = \lceil \frac{1}{2}|EA.verts| \rceil$.

(iii) $|EA_r.verts| = \lfloor \frac{1}{2}|EA.verts| \rfloor$.

Proof. For the runtime, we walk through the algorithm and analyze the time and diagram complexity of each line. In Lines 1–3, we find the angle α that splits $EA.verts$ into two equal sets, then in Line 4 compute a direction s orthogonal to α. See Figure 3.8. Lines 1–4 use no diagrams and can be done in constant time when restricting our attention to the (e_1, e_2)-plane. However, we need s to be a direction in $\mathbb{R}^d$ (as opposed to only in the (e_1, e_2)-plane), so the computation takes $\Theta(d)$ time. (Note that with some clever data structures,

this $\Theta(d)$ could be reduced to constant time. For example, we could require vectors in $\mathbb{R}^2$ are automatically padded with 0's to become vectors in $\mathbb{R}^d$ when needed. However, this is out of the scope of the real RAM model of computation.) Specifically, s is the vector

$$\begin{aligned} s &= e^{\frac{1}{2}i(2\alpha-\pi-\theta)} \\ &= \left(\cos\left(\alpha - \frac{1}{2}\pi - \frac{1}{2}\theta\right), \sin\left(\alpha - \frac{1}{2}\pi - \frac{1}{2}\theta\right), 0, 0, \ldots, 0\right). \end{aligned} \tag{3.9}$$

To compute m_ℓ in Line 5, we compute 1-Indeg(v, s) then subtract the cardinality of the set $S := \{b \in \mathit{bigedges} \mid \measuredangle b < \pi + \alpha\}$, where $\measuredangle b$ is taken to mean the angle b makes with the e_1-axis, when viewed as a vector with $EA.v$ as the origin. By Lemma 3.45, we compute 1-Indeg(v, s) via the oracle using one diagram and $\Theta(\log n + \Pi)$ time. Since *bigedges* is sorted and since s lies in the (e_1, e_2)-plane, we can find the set S in $\Theta(\log(|\mathit{bigedges}|))$ time. The subtraction in Line 5 takes constant time, as does Line 6.

In Lines 7 and 8, we create two edge arc objects. The time complexity of creating them is proportional to the size of the objects themselves. All attributes of edge arc objects, except the array of vertices (*verts*), are constant size. By construction, $EA_\ell.\mathit{verts}$ and $EA_r.\mathit{verts}$ split $EA.\mathit{verts}$ into two sets, which can be done naïvely in $\Theta(d|EA.\mathit{verts}|)$ time by walking through $EA.\mathit{verts}$ and storing each one explicitly. We improve this to $\Theta(\log|EA.\mathit{verts}|)$ time if we have a globally accessible array of vertices (sorted cw around v) and simply compute the pointers to the beginning and end of the sub-arrays corresponding to the *verts* attributes of the new edge arc objects. In total, Algorithm 3.12 and takes $\Theta(d+\log n+\Pi+\log(|\mathit{bigedges}|)+1+\log(|EA.\mathit{verts}|)) = \Theta(\log n + d + \Pi)$ time and uses uses one diagram.

Now that we have walked through the algorithm and established the runtime and diagram complexity, we prove correctness. To do so, we first show that EA_ℓ and EA_r are edge arc objects. In particular, this means showing that they have the correct values for *count* and *verts*. We prove this for EA_ℓ; the proof for EA_r follows a similar argument.

EA_ℓ.count: We must show that $EA_\ell.count$ is the number of edges in EA_ℓ incident to $EA_\ell.v$. By Lemma 3.44, the value returned from 1-Indeg$(EA.v, s)$ counts all edges incident to $EA.v$ and below $s \cdot EA.v$ in direction s. By Lemma 3.46, this is exactly the total number of edges in EA_ℓ plus edges $(EA.v, v') \in E_v$ for which $s \cdot v' < s \cdot EA.v$. Thus, by subtracting $|\{b \in bigedges \mid \measuredangle b < \pi + \alpha\}|$ from 1-Indeg$(EA.v, s)$ on Line 5, we are left with m_ℓ, the number of edges incident to $EA.v$ contained in EA_ℓ. Setting $EA_\ell.count = m_\ell$ on Line 7, we see that $EA_\ell.count$ is correct.

EA_ℓ.verts: We must show that $EA_\ell.verts$ contains an array of all vertices contained in EA_ℓ radially ordered clockwise. This follows from the fact that $EA.verts$ is all vertices contained in EA ordered clockwise, so when we restrict $EA.verts$ to $EA.verts[: mid]$ on Line 7, we are eliminating vertices not contained in EA_ℓ, so $EA_\ell.verts$ is correct.

Next, we prove Statement (i). Recall that $EA.verts$ orders the vertices in decreasing angle with e_1. In Line 3, $\measuredangle\pi(EA.verts[mid] - EA.v)$ is the angle made by $EA.v$ with the middle vertex. We tilt this angle by $\theta/2$ on Line 3 to obtain angle α. By construction of α,

$$\measuredangle\pi(EA.verts[mid] - EA.v) > \alpha.$$

By definition of θ, the angle α satisfies:

$$\alpha > \measuredangle\pi(EA.verts[mid+1] - EA.v).$$

Since the array $EA.verts$ is sorted, all vectors in the set $\pi(EA.verts[: mid] - EA.v)$ have an angle of at least α and all vectors in $\pi(EA.verts[: mid] - EA.v)$ have an angle of at most α.

By Lines 1–2 and Line 5, we know $EA_\ell.verts$ has the first $m = \lceil \frac{1}{2}|EA.verts| \rceil$ vertices in $EA.verts$, so Statement (ii) holds. Statement (iii) follows from Statements (i) and (ii). □

In Algorithm 3.13, we use Algorithm 3.12 to find all outgoing edges from a given vertex. In particular, the algorithm maintains a stack of edge arc objects. When processing an edge arc (the while loop in Lines 4–16), we are determining which of the vertices in *verts* form edges with v. If an edge arc has *count* $= 0$, it contains no edges, and it can be ignored (Lines 6–8). If it has *count* exactly equal to the number of vertices in *verts*, each vertex in *verts* must form an edge with v (Lines 8–11). Otherwise, the edge arc is split in half using Algorithm 3.12 and each half is put on the stack to be processed (see Figure 3.9).

Theorem 3.48 (Finding Edges Above a Vertex)**.** *Algorithm 3.13 finds the sorted array of edges above v using $\Theta(\deg(v)\log n)$ diagrams in $\Theta((\deg(v)\log n)(\log n + d + \Pi))$ time.*

Proof. First, we analyze the time complexity of the algorithm and the number of diagrams it requires. By Lemma 3.45, Line 1 can be computed in $\theta(\log n)$ time (since we are given the diagram and do not need an additional oracle query). Storing $A.verts$ by storing a pointer to V_v, we initialize *eastack* and E_v in Lines 2 and 3 in constant time.

Algorithm 3.13: `UpEdges`$(v, V_v, in_v, \theta, \hat{\rho}(e_2))$

Input: $v \in V$; V_v, array of all vertices in V above v, ordered clockwise; in_v, array of all incoming edges of v, sorted radially clockwise; θ, the minimum angle formed by any three vertices in $\pi(V)$; and $\hat{\rho}(e_2)$, the augmented persistence diagram in direction e_2

Output: array of all outgoing edges of v

1: $\mathit{indeg} \leftarrow$ indegree of v in direction $-e_2$.

2: $\mathit{eastack} \leftarrow$ a stack of edge arc objects, initialized with a single edge arc A, where $A.v = v$, $A.\alpha_1 = 0$, $A.\alpha_2 = \pi$, $A.\mathit{count} = \mathit{indeg}$, and $A.\mathit{verts} = V_v$

3: $E_v \leftarrow \emptyset$

4: **while** *eastack* is not empty **do**

5: $EA \leftarrow \mathit{eastack.pop}()$

6: **if** $EA.\mathit{count} = 0$ **then**

7: **Continue** to top of while loop

8: **else if** $EA.\mathit{count} = |EA.\mathit{verts}|$ **then**

9: Append $v \times EA.\mathit{verts}$ to E_v, in order

10: **Continue** to top of while loop

11: **else**

12: $(EA_\ell, EA_r) \leftarrow$ `SplitArc`$(EA, in_v \cup E_v, \theta)$

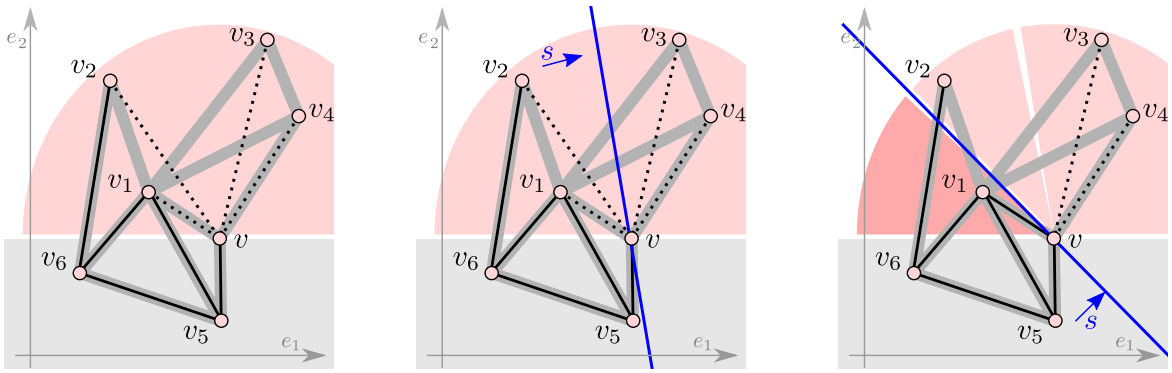
13: Push EA_r onto *eastack*

14: Push EA_ℓ onto *eastack*

15: **end if**

16: **end while**

17: **return** E_v



(a) At vertex v. (b) Split edge arc. (c) Split again.

Figure 3.9: We demonstrate one step of Algorithm 3.13. (a) By assumption, we initially know $[v_5, v] \in E$. From Line 1 of Algorithm 3.13 we also know that two of the four vertices above v are adjacent to v. Thus, we create an edge arc object EA with $EA.count = 2$, and $EA.verts = (v_1, v_2, v_3, v_4)$. (b) In Algorithm 3.12, we choose a direction s such that half of the vertices in EA are below v. We use this split to create two edge arcs, EA_r and EA_ℓ, corresponding to the pink shaded regions on the right and left of the blue line defined by s. We push EA_r onto a stack to be processed later and focus on the arc EA_ℓ. Since two edges contribute to v's indegree in direction s and one is the known edge $[v_5, v]$, we have $EA.count = 2 - 1 = 1$. (c) Next, we find a new direction s that splits $EA_\ell.verts$ into two sets of size one. We push the set above s onto our stack. The edge arc containing only v_1 also has $EA.count = 2 - 1 = 1$, so $[v_1, v] \in E$. After all steps of Algorithm 3.13 are applied to find the edges above a particular vertex, Algorithm 3.13 is then applied to the next highest vertex, eventually processing every vertex in V in a sweep (Algorithm 3.14).

To analyze the complexity of the loop in Lines 4–16, we first note that this is a radial binary multi-search. When processing an edge arc, we decide whether all edges have been found or if we need to split the edge arc. If there is only one edge in the arc (i.e., $EA.count = 1$), then this loop is a binary search for an edge, using the angle with e_1 in the (e_1, e_2)-plane as the search key. When $EA.count > 1$, we search for all edges, finding them in decreasing angle order (since arcs with larger angles are added after arcs of smaller angles). The if statement in Lines 8–11 is where the edges are added to E_v. Note that this shortcuts additional edge arc splitting by stopping the process once we find that the number of edges in the arc is equal to the number of potential vertices that can form the edges. As a result, each edge

above v contributes to $O(\log n)$ edge arcs being added to *eastack* and, in the case that every other vertex is incident to an edge with v, we have $\Theta(\log n)$ edge arcs added to the stack. All operations in the while loop are constant time, except splitting the edge arc object in Line 12, which uses one diagram and takes $\Theta(\log n + d + \Pi)$ time.

The complexity of Algorithm 3.13 is dominated by the complexity of the while loop: the algorithm uses $\Theta(\deg(v) \log n)$ augmented persistence diagrams and takes takes $\Theta((\deg(v) \log n)(\log n + d + \Pi))$ time.

To prove correctness of this algorithm, we track how edges are stored. In this algorithm, *eastack* is a stack of edge arc objects. We abuse notation slightly and say that an edge $(v', v) \in E$ is in *eastack* if there exists an edge arc $A \in$ *eastack* such that $\measuredangle(v', v) \in (A.\alpha_1, A.\alpha_2)$. In this case, we also say (v', v) is in A. We claim the while loop has the following loop invariant:

(i) every outgoing edge from v is either in *eastack* or is in the list of edges E_v;

(ii) the edges of E_v are radially sorted clockwise order;

(iii) the edge arcs in *eastack* are non-overlapping and in radially clockwise order; and,

(iv) for all $(v', v) \in E_v$ and $(v'', v) \in$ *eastack*, the angle $\measuredangle(v', v)$ is larger than the angle $\measuredangle(v'', v)$.

Entering the loop, E_v is empty and *eastack* is a single edge arc accounting for the entire half space above v. Thus, all outgoing edges are in *eastack* and E_v is empty, so the loop invariant is initially true. Denote the edge arc that is processed in iteration j by EA^j. Suppose entering

iteration j, the loop invariant holds. In Line 5, we pop EA^j from *eastack*. There are three cases to consider. First, if $EA^j.count = 0$, there are no edges in EA^j, so we enter the if clause in Line 6 and continue to the top of the while loop. Since no edges were removed from *eastack* when EA^j was popped, we do not add any edges to E_v nor do we change any ordering of the previously sorted edges or edge arcs in *eastack*. Thus, the loop invariant is maintained. Second, if $EA^j.count = |EA^j.verts|$, we enter the else-if clause in Line 8. Here, we know the vertices in the current edge arc are in a one-to-one correspondence with the edges in EA^j, so we add these edges in sorted order to E_v in Line 9; since the loop iteration was true upon entering the loop, these edges are between the edges already in E_v and the remaining edges in *eastack*. Thus, (iii) and (iv) are maintained. Furthermore, since we add the edges of EA^j in radially clockwise order, E_v is still radially sorted in clockwise order, satisfying (ii). Since all edges of EA^j are now in E_v, (i) is satisfied. So again, the loop invariant is maintained. The third and final case is that $EA.count \neq 0$ and $EA^j.count \neq |EA^j.verts|$. Then, we split the current edge arc on Line 12, pushing the two new halves onto the stack in a radially clockwise order, maintaining (iii). In this case, E_v remains as it was going into the loop and the edges in *eastack* remain the same, so the loop invariant is maintained. To finalize the proof of partial correctness, suppose the loop ends and that the loop invariant is true. Since the loop ends, *eastack* is empty. By (i) and (ii), all outgoing edges are in E_v in clockwise order. The runtime tells us the algorithm terminates, thus, the algorithm is correct. □

Finally, our main algorithm (Algorithm 3.14) is a sweepline algorithm that considers vertices in increasing order of their e_2-coordinates and finds all outgoing edges of each vertex.

Algorithm 3.14: `FindEdges`(V)

Input: V, array of all vertices in the unknown graph

Output: E, array of all edges in the unknown graph

1: $\hat{\rho} \leftarrow$ `Oracle`$(-e_2)$

2: $E \leftarrow \{\}$

3: *vertsabove* $\leftarrow$ for each $v \in V$, an array clockwise ordering all vertices in V that are above v

4: $\theta \leftarrow$ min angle defined by any three vertices of $\pi(V)$

5: **for** v in V, in increasing height in direction e_2 **do**

6: *inedges* $\leftarrow$ clockwise sorted array of edges in E incident to v

7: $E+=$ `UpEdges`$(v, \mathit{vertsabove}[v], \mathit{inedges}, \theta, \hat{\rho})$

8: **end for**

9: **return** E

Theorem 3.49 (Edge Reconstruction)**.** *Let (V, E) be a graph GP-immersed in $\mathbb{R}^d$. Given V, Algorithm 3.14 reconstructs E using $\Theta(m \log n)$ augmented persistence diagrams in $\Theta(n^2 + m \log n(\log n + d + \Pi))$ time.*

Proof. We first analyze the runtime and diagram count for Algorithm 3.14 by walking through the algorithm line-by-line. In Line 1, we ask the oracle for the diagram in direction $-e_2$, which takes $\Theta(\Pi)$ time. In [8, Theorem 14 (Edge Reconstruction)], simultaneously find the cyclic ordering around all vertices in $\Theta(n^2)$ time by Lemmas 1 and 2 of [47]. In Line 3, we do that as well; however, we do not store vertices that are above v in the array *vertsabove*$[v]$, and thus this line takes $\Theta(n^2)$ time. We note that such

a cyclic ordering exists around each vertex by Assumption 2(ii). Once we have *vertsabove*, to find the minimum angle defined by any three vertices of V, we check all angles between vectors $\mathit{vertsabove}[v][i] - v$ and $\mathit{vertsabove}[v][i+1] - v$ in Line 4 in $\Theta(n+m)$ time.

The for loop in Lines 5–8 is repeated n times, once for each vertex in V. To determine the order of processing the vertices in V, we follow the births in $\hat{\rho}_0$, in decreasing order (since $\hat{\rho}_0$ corresponds to the lower-star filtration in direction $-e_2$). Thus, finding the order takes $\Theta(n)$ time. In each iteration, we compute the incoming edges (those whose other vertex is below v) in Line 6 followed by all outgoing edges (those whose other vertex is above v) in Line 7. By Assumption 2(iii), every edge is either incoming or outgoing with respect to direction e_2. So, all edges incident to v are in E once E is updated in Line 6. By Theorem 3.48, the call to Algorithm 3.13 on Line 7 for a vertex v takes $\Theta((\deg(v)\log n)(\log n + d + \Pi))$ time and uses $\Theta(\deg(v)\log n)$ diagrams. Summing over all vertices, the loop in Lines 5–8 takes

$$\begin{aligned}\sum_{v\in V} &\Theta((\deg(v)\log n)(\log n + d + \Pi)) \\ &= \Theta(m\log n(\log n + d + \Pi))\end{aligned}$$

time and uses $\sum_{v\in V}\Theta(\deg(v)\log n) = \Theta(m\log n)$ augmented persistence diagrams.

In total, we find that Algorithm 3.14 takes $\Theta(\Pi) + \Theta(n^2) + \Theta(n+m) + \Theta(n) + \Theta(m\log n(\log n + d + \Pi)) = \Theta(n^2 + m\log n(\log n + d + \Pi))$ time and uses $\Theta(1) + \Theta(m\log n) = \Theta(m\log n)$ diagrams.

Next, we prove the correctness of Algorithm 3.14 (i.e., that all edges are found). In order to process vertices in order of their heights in the e_2 direction, we first sort them in Line 5. For $1 \leq j \leq n$, let v_j be the jth vertex in this ordering. To show that Algorithm 3.14 finds all edges in E, we consider the loop invariant (LI): when we process v_j, all edges with maximum vertex height equal or less than the height of v_j are known. The LI is trivially true for v_1. We now assume that it is true for iteration j, and show that it must be true for iteration $j+1$. By assumption, all edges (v_i, v_j) with $1 \leq i < j$ are known, and so by Theorem 3.48, Algorithm 3.13 finds all edges (v_k, v_j), where $k > j$, and we add them to the edge set E. Note that, by assumption, all edges (v_x, v_i) for $1 \leq i \leq j$ are also already known, and so the invariant is maintained. Thus, after the loop terminates, all edges are found. □

3.49.1 Putting it Together: Full Reconstruction

The results of Section 3.46.2 are related to just part of the full process of reconstruction, since reconstruction begins with no knowledge of the underlying simplicial complex. Identifying the location of all vertices is the first step, and is one that has been previously examined in detail. In particular, Belton et al. provide an algorithm to reconstruct V in $\Theta(dn^{d+1}+d\Pi)$ time and $d+1$ oracle queries, assuming stricter general position assumptions; see [8, Algorithm 1 & Theorem 9]. One could also consider using the vertex reconstruction methods described in Section 3.8.1, which is suitable for weaker general position assumptions than Assumption 2 (we opt not to focus on this choice here simply because Section 3.8.1 did not address the runtime of the reconstruction). An appropriate vertex reconstruction method taken together with Theorem 3.49, gives us the following for a full reconstruction process.

Theorem 3.50 (Graph Reconstruction)**.** *Let G be an unknown graph immersed in $\mathbb{R}^d$. If we use a vertex reconstruction method that requires $\Theta(\mathbb{V}_{time})$ time and $\Theta(\mathbb{V}_{dgm})$ diagrams, we can reconstruct G using $\Theta(\mathbb{V}_{dgm} + d + m \log n)$ diagrams in $\Theta(\mathbb{V}_{time} + n^2 + m \log n(\log n + d + \Pi))$ time.*

We omit a proof of Theorem 3.50, as it simply combines the results of vertex reconstruction methods (such as [8, Theorem 9]) and Theorem 3.49 of the previous section. Note that, taking the stricter general position assumptions of [8], Theorem 3.50 gives us a runtime of $\Theta(dn^{d+1} + d\Pi + n^2 + m \log n(\log n + d + \Pi))$ with $\Theta(d + m \log n)$ oracle queries.

Observing that the methods presented here are immediately applicable in the reconstruction of one-skeletons of general simplicial complexes, we have the following corollary:

Corollary 3.51 (One-Skeleton Reconstruction)**.** *Let K be an unknown simplicial complex GP-immersed in $\mathbb{R}^d$. A vertex reconstruction algorithm that requires $\Theta(\mathbb{V}_{time})$ time and $\Theta(\mathbb{V}_{dgm})$ diagrams taken together with Algorithm 3.14 of the current chapter reconstruct the one-skeleton of K using $\Theta(\mathbb{V}_{dgm} + d + m \log n)$ augmented persistence diagrams in $\Theta(\mathbb{V}_{time} + n^2 + m \log n(\log n + d + \Pi))$ time.*

Finally, we note that embedding a graph (or simplicial complex) in $\mathbb{R}^2$ is a special case, as $m = O(n)$ and d is constant. This speeds up edge reconstruction as well as potentially, vertex reconstruction. For instance, with the stricter general position assumptions, the method of [8, Theorem 6] gives a vertex reconstruction of a graph embedded in $\mathbb{R}^2$ that uses three diagrams and $\Theta(n \log n + \Pi)$ time. Broadening our view again to general vertex reconstruction algorithms, we obtain a result for plane graph reconstruction:

Corollary 3.52 (Reconstruction in $\mathbb{R}^2$)**.** *Suppose we use a vertex reconstruction method that requires $\Theta(\mathbb{V}_{time})$ time and $\Theta(\mathbb{V}_{dgm})$ diagrams. We can use an oracle to reconstruct the one-skeleton of an unknown simplicial complex embedded in $\mathbb{R}^2$ using $O(\mathbb{V}_{dgm} + n \log n)$ diagrams and $O(\mathbb{V}_{time} + n^2 + n \log n(1 + \Pi))$ time.*

3.53 Discussion

In this chapter, we begin by providing sufficient conditions for a faithful discretization of the APHT. For a simplicial complex GP-immersed in $\mathbb{R}^d$, with n_0 vertices, n simplices, and dimension κ, the discretization is of size $O(n2^\kappa + d)$ and it can be computed in $O(\log(n_0 + d)(d^2 n_0^d + nd + nn_0 2^\kappa) + n2^\kappa d^3)$. Since only the presence and dimension of filtration events are used (and not birth/death pairing information), the techniques presented in the chapter apply to *any* dimension-returning transform, such as the Augmented Betti Curve Transform. Furthermore, we give an adaptation of these methods to show that the set of directions we define is also a discretization of the AECCT. Finally, we show that the discretization is stable with respect to multiple types of perturbations.

In this chapter, we also provide a faster algorithm for reconstructing a graph immersed in $\mathbb{R}^d$. We use fewer persistence diagrams than presented in alternate approaches. For example, the algorithm that we present for edge reconstruction (when the vertex locations are known) uses $\Theta(m \log n)$ diagrams. In contrast, [8, Theorem 16] uses $n^2 - n$ diagrams. Note that, for a very dense edge set, that is, when $m = \Theta(n^2)$, the method in [8, Theorem 16] uses fewer diagrams. However, if $m = O(n)$, as is common in many complexes, the representation computed in this chapter has fewer diagrams. Moreover, we emphasize that the number of diagrams is not exponential in the ambient dimension.

One might hope to use binary search strategies to reconstruct a simplicial complexes, but the methods presented here are unique to one-skeletons. Radially ordering higher dimensional simplices is not well-defined, and this issue prevents the methods presented here from being immediately transferable. On the other hand, with the representations in this chapter being output-sensitive (as opposed to testing if every set of $k+1$ vertices is a k-simplex), we have hope for the discretization of the (A)PHT of a simplicial complex immersed in $\mathbb{R}^d$ being proportional to the size of the complex itself.

CHAPTER FOUR

ORDERING CLASSES OF TOPOLOGICAL DESCRIPTORS

4.1 Introduction

Recent developments in shape reconstruction and comparison call for the use of many different types of topological descriptors. The use of different descriptors to represent a shape motivates our main question - *how can we compare topological descriptors in terms of their ability to uniquely correspond to shapes?* First, we define a framework and relation for comparing general topological descriptors via their *strength equivalence classes* (Section 4.2 and Section 4.5). Although we focus on six commonly studied topological descriptors, the relation we define to compare strength classes of descriptors could be applied to any set of topological descriptors using any types of filtration. In Section 4.15 we give a partial order on the strength classes of our set of six descriptors, including a strict inequality that separates the subset of augmented descriptor strength classes. Finally, we provide lower bounds on the cardinality of representative sets for both non-augmented descriptors (Section 4.22.1) and augmented (Section 4.27.1). Although we provide strict lower bounds for both, the main contribution of this section is Ω bounds that strongly suggest non-augmented descriptors are generally much weaker than augmented descriptors. This work sets up a rigorous theory that will allow for future comparisons and analysis of topological descriptors.

4.2 Preliminary Considerations

Although we focus on a specific set of topological descriptors in subsequent sections, we begin by setting up a theory suitable for comparing general topological descriptors. In this section, we explore considerations this general viewpoint requires; in particular, considerations related to the cardinalities or existence of faithful sets of descriptors.

First, we need a definition of a general topological descriptor. In some contexts, we may view topological descriptors as arising solely from persistence modules, in which case, topological descriptors need not have any connection to even an underlying simplicial complex. However, since we define our relations based on the ability of descriptor types to represent particular filtrations of simplicial complexes, we take the following viewpoint.

Definition 4.3 (Topological Descriptor)**.** *A* topological descriptor *is a map whose domain is the collection of finite filtered geometric simplicial complexes.*

We compare descriptor types by quantifying the number of descriptors needed to fully represent a shape, that is, to be faithful for the underlying shape. Recall that we say a (parameterized) set of descriptors is *faithful* for a simplicial complex K if only K could have generated that same set of descriptors (see Definition 3.2).

Specific bounds on the cardinality of faithful sets exist in related work for some descriptors. For instance, from Chapter 3, we know that for any simplicial complex $K \in \mathbb{R}^d$ (with standard general position assumptions), there is a set of $O(n2^\kappa + d)$ augmented descriptors (comprised of either $\hat{\chi}$'s, $\hat{\beta}$, $\hat{\rho}$'s, or a combination of the three) that is faithful for K (see, e.g., Theorem 3.27). Given additional assumptions on the curvature of K, Curry

et al. upper bound the number of χ's needed to form a representative set [20].

When a descriptor of type D corresponds to a parameter value p and simplicial complex K, we use the notation $\mathrm{D}(K, p)$. We may instead write $\mathrm{D}(p)$ when the simplicial complex is clear from context. Furthermore, just as in Chapter 3, we refer to the set of descriptors parameterized by P using the notation $\mathrm{D}(K, P)$. We make an effort to use P for an arbitrary parameter set and S when the parameter set is specifically a subset of the sphere of directions (e.g., when a descriptor corresponds to a lower-star filtration).

4.3.1 Zoo of Descriptor Examples

We take a moment to explore a variety of descriptor types highlighting certain situations that may arise when comparing the cardinalities of faithful sets of descriptors. Note that without the reliance on general position assumptions, many simplicial complexes do not have a faithful set of some types of topological descriptors; Lemma 4.16 gives such a construction. There are also many examples of topological descriptors that are not capable of faithfully representing most simplicial complexes (even simplicial complexes satisfying strict general position assumptions), such as the following.

Example 1 (Descriptor Type Only Faithful for Convex Point Clouds)**.** *Consider a descriptor that returns the coordinates of the lowest vertex (or vertices) in a lower-star filtration and the cardinaltiy of the vertex set, but no other information. Any set of vertices that defines the corners of a convex region can be faithfully represented by this descriptor, but this descriptor type is incapable of faithfully representing any other type of simplicial complex.*

We can also construct descriptor types that are simply never able to form faithful sets.

Example 2 (Descriptor Incapable of Faithfully Representing Any Complex)**.** *Consider the trivial descriptor type that returns zero for all sublevel sets in a lower-star filtration. Although it is an invariant of the filtration, it can not faithfully represent any simplicial complex.*

Finally, we can also find instances of topological descriptors that are able to faithfully represent a simplicial complex, but with a set no smaller than uncountably infinite.

Example 3 (Descriptor with Uncountable Minimum Faithful Sets)**.** *Let $K \in \mathbb{R}^d$ be a simplicial complex satisfying the general position Assumption 3 and let D be a descriptor type parameterized by $\mathbb{R}^d$ that is constant over a filtration and is defined by*

$$D(K, (x_1, x_2, \dots x_d)) = \begin{cases} 1 & \textit{if } K \cap (x_1, x_2, \dots x_d) \neq \emptyset \\ \\ 0 & \textit{else.} \end{cases}$$

Then $D(K, \mathbb{R}^d)$ is the minimum faithful set for K.

4.3.2 Cardinality Conventions

We would still like to include these (rather pathological) descriptor type/simplicial complex pairs in our analysis. In particular, we would like to consider descriptors that are not able to form faithful sets for simplicial complexes as being in *weaker* strength classes than those that can, those that require uncountaly infinite sets as being in weaker classes than those that require countably infinite sets, and those that require countably infinite sets as being in weaker classes than those whose minimum faithful sets are finite.

We incorporate the situation that a descriptor cannot faithfully represent a simplicial complex as a specific case in our definitions and proofs. If the minimum cardinality of a

faithful set is uncountably infinite, we say it has size $\aleph_1$. If the minimum cardinality of a faithful set is countably infinite, we say it has size $\aleph_0$. As is standard, we take $\aleph_0 < \aleph_1$.

In the following lemma, we show if a descriptor type has a finite faithful set for the vertex set of a complex, the minimum faithful set for the entire complex is also finite.

Lemma 4.4 (Sufficicient Conditions for Finite Faithful Set)**.** *Let $K \in \mathbb{R}^d$ be a finite simplicial complex and let* D *be a type of topological descriptor that can faithfully represent K. Suppose there exists a set of descriptors of type* D *that is faithful for K_0 with finite cardinality. Then there exists a finite faithful set of type* D*.*

Proof. From the hypotheses, we know there is a finite set $\mathrm{D}(K, P_0)$ that is faithful for K_0 and a set $\mathrm{D}(K, P)$ that is faithful for K. We will use these two sets to build a *finite* faithful set for the entire simplicial complex, K. Since $\mathrm{D}(K, P)$ is faithful for K, we have

$$\bigcap_{p \in P} \left\{ K' \in \mathbb{R}^d \mid \mathrm{D}(K', p) = \mathrm{D}(K, p) \right\} = \{K\}. \tag{4.1}$$

Similarly, since $\mathrm{D}(K, P_0)$ is faithful for K_0, we define

$$B := \bigcap_{p \in P_0} \left\{ K' \in \mathbb{R}^d \mid \mathrm{D}(K', p) = \mathrm{D}(K, p) \right\}$$

and see that $B \subseteq \{K' \mid K'_0 = K_0\}$, i.e., B is a subset of all simplicial complexes built out of the vertices of K. In particular, we note that this set is finite; since $|K_0| = n_0$ is finite, there are a finite number of simplicial complexes we can build over this set of vertices.

Suppose there is some $K^1 \in B$ such that $K^1 \neq K$ (if not, then $\mathrm{D}(K, P_0)$ is a finite

faithful set for K, and we are done). Then there must be some parameter $p^1 \in P$ for which we have $K^1 \notin \{K' \in \mathbb{R}^d \mid \mathrm{D}(K', p) = \mathrm{D}(K, p)\}$. Otherwise, the intersection in Equation (4.1) would contain K^1 in addition to just K, and, by Lemma 1.18, $\mathrm{D}(K, P)$ would not be faithful.

Since B is a finite set, there are finite $K^i \neq K$ contained in B, and by similar reasoning, for each such K^i we can find a direction $p^i \in P$ whose set of possible complexes does not include K^i. Denote the number of complexes not equal to K that are contained in B by n. Then the set $P_* = P_0 \cup \{p^1, p^2, \ldots, p^n\}$ is finite, and $\mathrm{D}(K, P_*)$ faithfully represents K. $\square$

4.5 Relating Descriptors

We are now ready to begin developing tools for comparing topological descriptor strengths. When discussing the cardinality of a set $\mathrm{D}(K, P)$, we generally use the notation $|P|$ for brevity; it should be noted that the size of a set of descriptors is equivalent to the size of its parameter set, i.e., $|\mathrm{D}(K, P)| = |P|$. First, we define what we mean when we say two topological descriptors have equal *strength*, and show this is an equivalence relation.

Definition 4.6 (Strength Equivalence Classes of Topological Descriptors)**.** *Let A and B be types of topological descriptors. Suppose that, for all finite geometric simplicial complexes K, one of the following is true:*

- *The minimal faithful set of type A and the minimal faithful set of type B have the same cardinality, or;*
- *There are no faithful sets of type A or B for K.*

Then we say the topological descriptor types A and B have equal strength*. This defines an equivalence class of strength, i.e., we write $[A] = [B]$.*

Example 4. *Consider the topological descriptor denoted $-\rho$ that takes a lower-star filtration in direction s, and produces $\rho(-s)$, the persistence diagram in direction $-s$. Although it is not generally the case that $\rho(s) = \rho(-s)$, we observe that a faithful set $\rho(K, S)$ has the same cardinality as the faithful set $-\rho(K, -S)$, and if a simplicial complex has no faithful set of type ρ, it will have no faithful set of type $-\rho$. Thus, $[\rho] = [-\rho]$.*

Lemma 4.7. *The equality relation of Definition 4.6 is well defined on sets of strength equivalence classes.*

Proof. Let A, B, and C be types of topological descriptors.

1. *Reflexivity:* It is clear that $[A] = [A]$.

2. *Symmetry:* Suppose that $[A] = [B]$. If there is some simplicial complex for which A cannot form a faithful set, since $[A] = [B]$, neither can B, so it must be that $[B] = [A]$. Suppose then we are in the case that both A and B can form faithful sets for a given simplicial complex, K. Then the minimal faithful sets of type A and of type B for K have the same cardinality, and thus, $[B] = [A]$.

3. *Transitivity:* Suppose $[A] = [B]$ and $[B] = [C]$. In the trivial case, if there are, WLOG, no faithful sets of type A, the first equality implies there are no faithful sets of type B, which means there are no faithful sets of type C, so $[A] = [C]$. Suppose then that there is a simplicial complex for which there are faithful sets for all three descriptor types. Then the minimal faithful sets for any finite simplicial complex have the same size among types A and B, and among types B and C. Thus, we see they have the same cardinality among types A and C, and so $[A] = [C]$.

We have shown equality of strength classes satisfies reflexivity, symmetry, and transitivity, so is well-defined. □

Next, we define what it means for the strength class of a topological descriptor type to be *weaker* than another, and show this is indeed a well-defined ordering relation.

Definition 4.8 (Relation $\preceq$)**.** *Let A and B be two types of topological descriptors. Suppose that, for every geometric simplicial complex K, one of the following is true;*

- *If $A(K, P_A)$ is a faithful set, then there is a faithful set $B(K, P_B)$ with $|P_B| \leq |P_A|$, or;*
- *There is no faithful set of type A for K.*

Then we say the class of A is weaker than *the class of B, denoted $[A] \preceq [B]$.*

Lemma 4.9. *The relation $\preceq$ is well defined.*

Proof. In order to show that $\preceq$ is well defined, we must show that it satisfies the definition of a relation. Let A, B, and C be types of topological descriptors.

1. *Reflexivity:* Clearly, $[A] \preceq [A]$.

2. *Antisymmetry:* Suppose that $[A] \preceq [B]$ and $[B] \preceq [A]$. If there are no faithful sets of types A or B for any simplicial complex, we immediately have $[A] = [B]$. The case there are faithful sets for one descriptor type but not the other would contradict one of the assumed inequalities. Thus, we proceed assuming both A and B can form faithful sets for some simplicial complex, K. From the first hypotheses we know that if $A(K, P_A)$ is a faithful set for K, there is a faithful set $B(K, P_B)$ with $|P_B| \leq |P_A|$. From the

second hypotheses, we know that there is a faithful set $A(K, P'_A)$ with size $|P'_A| \leq |P_B|$. Continuing this logic, we create a sequence of sizes of faithful sets with alternating types that is monotonically decreasing. However, since this is a sequence of integers that cannot reach zero, it must be that the minimum size of faithful set of type A and B are equal, and thus, $[A] = [B]$.

3. *Transitivity:* Suppose that $[A] \preceq [B]$ and $[B] \preceq [C]$. In the trivial cases (i.e., if there is no faithful set of type A or no faithful sets of types A and B), it is clear that $[A] \preceq [C]$. Suppose then that faithful sets exist for all three descriptor types. From the first hypothesis, if $A(K, P_A)$ is faithful, we can find a faithful set $B(K, P_B)$ such that $|P_A| \geq |P_B|$. From the second hypothesis, there exists a faithful set $C(K, P_C)$ such that $|P_B| \geq |P_C|$. Thus, we have $|P_A| \geq |P_C|$, so that $[A] \preceq [C]$, as desired.

We have shown the relation $\preceq$ on strength classes satisfies reflexivity, antisymmetry, and transitivity, so is well-defined. □

This relation implies the following definition.

Definition 4.10 (Relation $\prec$)**.** *Let A and B be two types of topological descriptors. We say the class of A is* strictly weaker than *the class of B (denoted $[A] \prec [B]$) if $[A] \preceq [B]$ and $[A] \neq [B]$. That is, if $[A] \preceq [B]$ and additionally there exists a geometric simplicial complex for which the minimum faithful set of type B is strictly smaller than the minimum faithful set of type A, or for which there exists a faithful set of type B but not of type A.*

Although the above relations are well defined, the following cautionary lemma shows that not all topological descriptors are comparable.

Lemma 4.11 (Incomparable Strength Classes)**.** *There exist incomparable strength classes of topological descriptor types.*

Proof. We define a topological descriptor D and compare its strength class to the class of the descriptor type $\hat{\rho}$. Let D be a directional topological descriptor that returns: (1) the exact coordinates of the lowest vertex (or vertices) of a lower-star filtration, and (2) the cardinality of the total vertex set (as in Example 1). Let $v_1 = (0, 0)$ and $v_2 = (0, 1)$.

First, consider the simplicial complex $K = [v_1]$. Then a single descriptor of type D generated from any direction is faithful for K. However, since $K \subset \mathbb{R}^2$, any faithful set of $\hat{\rho}$'s must have at least two linearly independent directions to recover both coordinates of K.

Next, consider the simplicial complex $K' = [v_1, v_2]$. No set of descriptors of type D is faithful for K' (it cannot distinguish between K' and the simplicial complex consisting of the two disconnected vertices v_1 and v_2 without an edge). However, two $\hat{\rho}$'s suffice to form a faithful set for K'; for example, the set $\hat{\rho}(K', \{e_1, e_2\})$. Thus, if $\mathrm{D}(K, S_A)$ and $\hat{\rho}(K, S_{\hat{\rho}})$ are minimal faithful sets, we see that $|S_{\mathrm{D}}(K)| < |S_{\hat{\rho}}(K)|$ but $|S_{\mathrm{D}}(K')| > |S_{\hat{\rho}}(K')|$. Thus, although we have shown $[\mathrm{D}] \neq [\hat{\rho}]$, they are incomparable. □

Now that we have a way of ordering (comparable) descriptor classes, we are ready to discuss determining ordering. It is often difficult to directly determine and compare minimal faithful sets in general, so it is useful to have a variety of tools that allow us to determine how descriptors are related. We conclude this section by defining reduction of one descriptor to another, and show this is another valid strategy for determining equivalence class order.

Definition 4.12 (Reducing a Descriptor to Another)**.** *Let A and B be two types of topological descriptors, let P be a family of filtration parameters, and let K be a simplicial complex. We say B can be* reduced *to A if there exists an algorithm such that, for each $p \in P$, the input of $B(K, p)$ into the algoirthm produces the output $A(K, p)$.*

We can also view the preceeding definition through a more functorial perspective. If we think of A and B as assignments that map a simplicial complex and parameter pair to some targets, B can be reduced to A if A factors through B (if the codomain of A and B agree), or through B composed with a map between the codomain of B to the codomain of A.

Lemma 4.13. *Let A and B be two types of directional topological descriptors. Then if B can be reduced to A, we have $[A] \preceq [B]$.*

Proof. If, for some simplicial complex K, there is no faithful set of type A, we have $[A] \preceq [B]$ by definition. Therefore, suppose that for some simplicial complex K, both B and A can form faithful sets. Let $B(K, P_B)$ be faithful and have minimum possible cardinality. Assume towards a contradiction, that there exists a faithful set $A(K, P_A)$ such that $|P_A| < |P_B|$. Then, since we can reduce B to A, the set $B(K, P_A)$ is faithful. But then $|B(K, P_A)| < |B(K, P_B)|$, contradicting the minimality of $B(K, P_B)$. Thus, any faithful set of type A is at least as large as $B(K, P_B)$, so $[A] \preceq [B]$, as desired. $\square$

4.14 The Topological Descriptors

Any partial order has an underlying set. The set on which we build our partial order is a set of strength equivalence classes of six popular topological descriptors. In Chapter 1, we carefully defined these descriptors. Here, we briefly remind the reader of their key attributes.

- Euler characteristic curves (χ's). From χ, we see the Euler characteristic of subcomplexes in a filtration every time it changes.

- Betti curves (β's). From β, we see the Betti numbers of subcomplexes in a filtration every time they change.

- Persistence diagrams (ρ's). From ρ, we see when non-trivial homology classes appear and disappear throughout a filtration, and how these birth/death events are paired.

- Augmented Euler characteristic curves ($\hat{\chi}$'s). From $\hat{\chi}$, we see the total number of even and odd simplices throughout a filtration.

- Augmented Betti curves ($\hat{\beta}$'s). From $\hat{\beta}$, we see the total number of i-simplices throughout a filtration, for all $i \geq 0$.

- Augmented persistence diagrams ($\hat{\rho}$'s). From $\hat{\rho}$, we see when homology classes are appear and disappear throughout a filtration, and how these birth and death events are paired. This includes homology classes that have trivial lifespan, corresponding to the appearance of boundaries that were not previously cycles.

See Figure 1.1 for an example of each descriptor. Although each of these topological descriptor types may correspond to abstract one-parameter persistence modules, or from any type of filtration, for the purposes of comparing the ability of these descriptor types to represent shapes, we consider the descriptors only as corresponding to a lower-star filtration of a finite simplicial complex. When a descriptor of type D is generated from a lower-star filtration in direction s over a simplicial complex K, we use the notation $\mathrm{D}(K, s)$ and

may write $\mathrm{D}(s)$ when the simplicial complex is clear from context. Furthermore, just as in Chapter 3, we refer to the set of descriptors parameterized by S using the notation $\mathrm{D}(K, S)$.

4.15 Partial Order

In this section, we present and prove our main theorem, a partial order on six common topological descriptors. Throughout this section, we may use e_i as shorthand notation for the ith unit standard basis vector. We begin with a lemma that highlights a stark deficit of our three non-augmented descriptors, and will serve as evidence for the reasonableness of instituting a general position assumption.

Lemma 4.16 (Non-Augmented Equality)**.** *Without reliance on any general position assumptions, we have* $[\chi] = [\beta] = [\rho]$.

Proof. Consider the simplicial complex K that is a single edge in $\mathbb{R}^2$ with vertices at $(0, 0)$ and $(1, 1)$. No set of χ's, β's, or ρ's can faithfully represent K, since the descriptors from any direction corresponding to K are identical to the descriptors from any direction corresponding to any simplicial complex with vertices $(0, 0)$ and $(0, 1)$ defining the boundary of its convex hull and any number of colinear points forming a string of colinear edges. Thus, no set of χ's, β's, or ρ's is faithful. Indeed, the edges of any simplicial complex can not be uniquely represented by these three descriptors, regardless of dimension or presence/absence of other dimensions of simplices. Thus, we see that without reliance on general position assumptions, the three non-augmented descriptors χ, β, and ρ are incapable of forming faithful sets for any simplicial complex, and are therefore in the same equivalence class of strength. $\square$

We obtain a richer analysis of the descriptors of Section 4.14 using the general position assumption of Section 3.8.1 of Chapter 3 (Assumption 3), which also has the benefit of being more reflective of assumptions commonly taken in practice.

Assumption 3. (General Position) *In what follows, we say a point set $V \subset \mathbb{R}^d$ is in* general position *if, for all k with $1 \leq k \leq d+1$, every subset of V of size k is affinely independent.*

Proceeding with this assumption, we are now ready to order the descriptors of Section 4.14. We begin by observing that augmenting a descriptor only makes it stronger.

Lemma 4.17. $[\chi] \leq [\hat{\chi}]$, $[\beta] \leq [\hat{\beta}]$, *and* $[\rho] \leq [\hat{\rho}]$.

Proof. Recall that each augmented descriptor contains information about instantaneous changes in topological properties (e.g., instantaneous changes in Euler characteristic, Betti number, or homology). It is clear when events are instantaneous, and we can recover each non-augmented counterpart by removing the information of instantaneous changes, a reduction as in Definition 4.12. Then by Lemma 4.13, we have the desired relations. □

We also use reduction to confirm the following relations.

Lemma 4.18. $[\chi] \preceq [\beta] \preceq [\rho]$ *and* $[\hat{\chi}] \preceq [\hat{\beta}] \preceq [\hat{\rho}]$

Proof. The proof follows directly from a reduction argument. We can reduce any $\rho(s)$ to $\beta(s)$ by 'forgetting' the relationship between birth and death events. We can then reduce $\beta(s)$ to $\chi(s)$ by taking the alternating sum of points on the curves in $\beta(s)$. A nearly identical argument shows the relationship between augmented versions of these descriptors. □

We note that the reduction from ρ to χ was observed in [20, Prop 4.13].

Lemma 4.19. *For* $\mathrm{D} \in \{\chi, \beta, \rho\}$ *and* $\hat{\mathrm{D}} \in \{\hat{\chi}, \hat{\beta}, \hat{\rho}\}$*, we have* $[\mathrm{D}] \neq [\hat{\mathrm{D}}]$.

Proof. Let K be the simplicial complex consisting of a single edge embedded in $\mathbb{R}^2$ with vertex coordinates $(0,0)$ and $(0,1)$. From $\hat{\mathrm{D}}(e_1)$, we see that there is only one connected component with a single e_1-coordinate. From $\hat{\mathrm{D}}(e_2)$, we see that there simplices at two e_2 heights, whose values we know. Then K is the only simplicial complex that could have generated both $\hat{\mathrm{D}}(e_1)$ and $\hat{\mathrm{D}}(e_2)$. Thus, the set $\hat{\mathrm{D}}(K, \{e_1, e_2\})$ is faithful.

Next, consider the descriptor type D. For any $s \in \mathbb{S}^1$, the descriptor $\mathrm{D}(s)$ contains exactly one event and thus can only uniquely correspond to one coordinate of K at a time. Since K can be expressed with a minimum of three coordinate values, it is not possible for any faithful set of descriptors of type D to have cardinality less than three. Thus, since $|S_{\mathrm{D}}| > |S_{\hat{\mathrm{D}}}|$, we have shown $[\mathrm{D}] \neq [\hat{\mathrm{D}}]$, as desired. $\square$

Note that the previous lemma can also be proven using the example of any chain of edges with three or more vertices colinear, which by Lemma 4.16 can never have faithful set of non-augmented descriptors, but does have a faithful set of augmented descriptors. However, the proof of Lemma 4.19 uses a simplicial complex that does not violate standard general position assumptions, and is therefore preferable.

So far, many of our arguments have made use of the fact that augmented descriptors have an event at the height of every vertex, but have not made use of the different types of topological invariants recorded by the various descriptor types. This, combined with the fact that the sets of directions described in Chapter 3 formed faithful sets of $\hat{\rho}$'s, $\hat{\beta}$'s, or $\hat{\chi}$'s (so, in particular, these faithful sets all have the same cardinality), might make it seem as though our

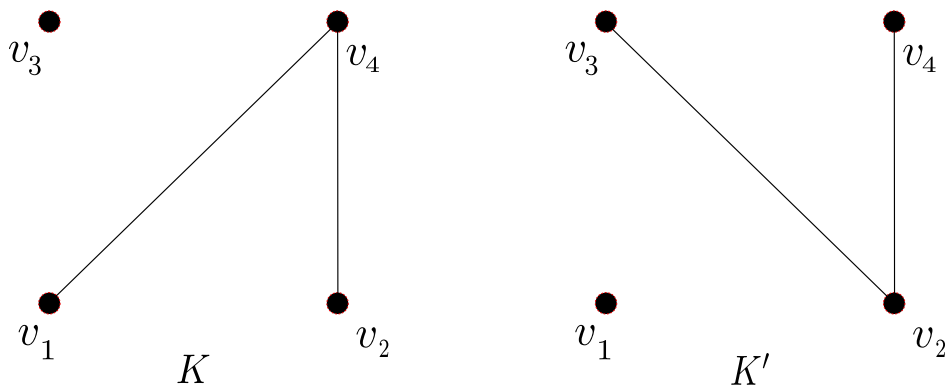


Figure 4.1: The two simplicial complexes considered in the proof of Lemma 4.20

augmented descriptors are all in the same strength class. Particularly because the methods of Chapter 3 were immediately transferable between $\hat{\rho}$'s and $\hat{\beta}$'s (indeed, and to any other augmented dimension-returning descriptors), it is tempting to suppose the two descriptor types $\hat{\rho}$ and $\hat{\beta}$ are in the same equivalence class of strength However, this is an incorrect leap of logic; in particular, because the faithful sets of Chapter 3 are almost certainly never minimal. The following lemma gives an instance where the topological invariant matters, rather than just the dimension or presence of events, when forming a minimal faithful set.

Lemma 4.20. $[\hat{\beta}] \prec [\hat{\rho}]$.

Proof. We know by Lemma 4.18 that $[\hat{\beta}] \leq [\hat{\rho}]$. It therefore suffices to show that $[\hat{\beta}] \neq [\hat{\rho}]$, that is, there exists a simplicial complex for which the cardinality of minimal faithful sets of $\hat{\beta}$'s and $\hat{\rho}$'s differ, and we proceed with this aim.

Consider the simplicial complex $K \subset \mathbb{R}^2$ consisting of: $v_1 = (0,0), v_2 = (0,1), v_3 = (1,0)$, and $v_4 = (1,1)$ with edges $[v_1, v_4]$ and $[v_2, v_4]$, as in the left side of Figure 4.1.

We first claim that $\hat{\rho}$'s generated from the e_1 and e_2 directions are sufficient to form a faithful set for K, i.e., a faithful set of cardinality two. In the degree zero homology

of $\hat{\rho}(e_1)$, we see two births at height zero, two (instantaneous) births at height one, and no other births. This tells us K has four vertices, two with e_1-coordinates of 0 and two with e_1-coordinates of 1. The same is true in $\hat{\rho}(e_2)$, so we know K_0 precisely. Since $\hat{\rho}(e_1)$ saw two births at height zero with infinite lifespan, we know there is no edge from v_1 and v_3. From $\hat{\rho}(e_2)$, we see a (non-instantaneous) zero-dimensional death at height one, meaning we know K must have two edges, either $[v_1, v_3]$ and $[v_2, v_3]$, or $[v_1, v_4]$ and $[v_2, v_4]$. Because $\hat{\rho}(e_1)$ sees two non-instantaneous zero-dimensional births at height zero, we know there cannot be two edges adjacent to v_3, so we are able to determine K exactly and thus have a faithful set.

We next claim that no two augmented Betti curves are sufficient to form a faithful set, i.e., a faithful set of $\hat{\beta}$ must have cardinality greater than two. Suppose, to the contrary, that s_1 and s_2 are two directions such that $\hat{\beta}(K, \{s_1, s_2\})$ is a faithful set. We first observe that, WLOG, $s_1 \in \{e_1, -e_1\}$ and $s_2 \in \{e_2, -e_2\}$. If not, the arrangement of filtration lines from s_1 and s_2 has more than four two-way intersections (unless $s_1 = -s_2$, which clearly does not correspond to a faithful set). An arrangement of filtration lines containing more than four two-way intersections is not sufficient to recover K_0; certainly the degree zero information $\hat{\beta}_0(s_1)$ and $\hat{\beta}_0(s_2)$ alone is not. The degree one curves, while telling us the height of edges, only confirms that there is a vertex at the height of the edge and some vertex below, information we already had from $\hat{\beta}_0$. Thus, WLOG, $s_1 \in \{e_1, -e_1\}$ and $s_2 \in \{e_2, -e_2\}$. However, for each of these four directions, the associated augmented Betti curve is not able to distinguish K from K' as in the right side of Figure 4.1. For instance, both $\hat{\beta}(K, e_1)$ and $\hat{\beta}(K', e_1)$ see two vertices at height zero, and two vertices and two edges at height one, meaning $\hat{\beta}(K, e_1) = \hat{\beta}(K', e_1)$. The other possible values of s_1 and s_2 are similar.

Thus, we have found a faithful set of $\hat{\rho}$'s with cardinality two, but have shown there is no faithful set of $\hat{\beta}$ with cardinality two, and the claim follows. $\square$

Combining the results from above, we arrive at our main theorem.

Theorem 4.21 (Partial Ordering)**.** *The following diagram is a correct partial order on strength classes of topological descriptors.*

$$\begin{tikzcd}
{[\hat{\chi}]} \arrow[r, "\leq"] & {[\hat{\beta}]} \arrow[r, "<"] & {[\hat{\rho}]} \\
{[\chi]} \arrow[u, "<"] \arrow[r, "\leq"] & {[\beta]} \arrow[u, "<"] \arrow[r, "\leq"] & {[\rho]} \arrow[u, "<"]
\end{tikzcd}$$

4.22 Bounds on Faithful Sets

This section deals with lower bounds on the size of faithful sets of descriptors (of the types listed in Section 4.14), both non-augmented (Section 4.22.1) and augmented (Section 4.27.1). Each subsection begins with a statement of strict lower bounds, but the more interesting results are finding examples of simplicial complexes whose minimal faithful set is large. Thus, the main theorems of this section provide big omega bounds on the worst-case minimum size of faithful sets.

4.22.1 Non-Augmented Descriptor Bounds

A defining feature of non-augmented descriptors is that there is not necessarily information recorded at the height of every vertex in a filtration. (We explored this feature in the case of χ's specifically in [28].) For instance, the closer a feature becomes to colinear, coplanar, etc., the smaller the range of directions that can detect the feature becomes. Difficulty detecting the presence or absence of structures that are near to the same affine

subspace puts greater restrictions on the ability of non-augmented descriptors to represent a simplicial complex. We make this observation precise in the following definition and lemma.

Definition 4.23 (Simplex Envelope)**.** *Let $K \subset \mathbb{R}^d$ be a simplicial complex and let $\sigma \in K$. Let S be a set of directions. Then, we define the* envelope of σ*, denoted P_σ^S, as the intersection of (closed) supporting halfspaces*

$$P_\sigma^S = \bigcap_{s \in S} \{p \in \mathbb{R}^d \mid s \cdot p \geq \min_{v \in \sigma}(s \cdot v)\}.$$

If S is clear from context, we write P_σ. By the dimension *of P_σ, we mean the largest dimension of ball that can be contained entirely in P_σ. By ∂P_σ, we mean the boundary of P_σ.*

Remark 2. *Intersections of closed halfspaces is a well-studied phenomenon in convex geometry. The simplex envelopes of Definition 4.23 have connections to topics such as convex cones, support functions, etc. See, e.g., [22, 48] for further information. In particular, [48, Thm 3.1.1, Cor 3.1.2] establish that a simplex envelope in the sense of Definition 4.23 with respect to the entire sphere of directions is the simplex itself.*

Note that since P_σ^S is the intersection of convex regions, it is itself a convex region. Furthermore, since, with respect to each $s \in S$, the height of each point of σ is greater than or equal to its minimum vertex, we see that P_σ^S contains σ. See Figure 4.2.

We use simplex envelopes to define a necessary condition for a set of non-augmented descriptors to be faithful for a simplicial complex.

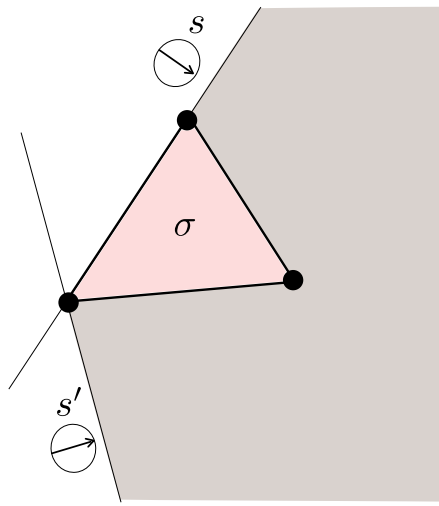


Figure 4.2: Given $S = \{s, s'\}$, the envelope for σ is shown in grey (and includes the area covered by σ).

Lemma 4.24 (Envelope Dimensions for Faithful Non-Augmented Sets)**.** *Let K be a simplicial complex embedded in $\mathbb{R}^d$, and for $D \in \{\chi, \beta, \rho\}$ and $S \subseteq \mathbb{S}^{d-1}$, let $\mathrm{D}(K, S)$ faithful. Then, for any maximal simplex $\sigma \in K$, the dimension of P_σ equals the dimension of σ.*

Proof. Let k denote the dimension of σ, and let c denote the dimension of P_σ. First, we observe that since σ is contained in P_σ, we must have $k \leq c$. The claim is trivial when $k = d$, so we proceed assuming $k < d$ and assume to the contrary that $k < c$.

We claim that in this case, 1) at every interior point of σ, there is a normal vector that ends in the interior of P_σ, and 2) for some such vector, the endpoint p is contained in the interior of P_σ, the point p is higher than the lowest vertex of σ with respect to each $s \in S$.

The first part of the claim, 1), is true since otherwise, P_σ would be k-dimensional. 2) is true since each halfspace defining P_σ contains the lowest vertex (or vertices) of σ with respect to the corresponding direction. Thus, since p is in the interior of P_σ, it must be higher than this lowest vertex (or vertices) with respect to the corresponding direction.

Now consider the simplicial complex K', defined as having all the simplices of K in addition to the simplex formed by joining p to σ, i.e., the simplex $p * \sigma$. We claim that, for any $s \in S$, we have $\mathrm{D}(K', s) = \mathrm{D}(K, s)$. First, we note that since $p * \sigma$ deformation retracts onto σ, K' has the same homology as K. Next, we observe that $\mathrm{D}(K', s)$ and $\mathrm{D}(K, s)$ cannot differ by more than a connected component birth/death; higher dimensional differences would require more than the join of a point with an existing face.

Finally, since p is lower than the lowest vertex of σ with respect to any $s \in S$, the simplex $p*\sigma \in K'$ does not correspond to any connected component birth or death in $\mathrm{D}(K', s)$ that was not present in $\mathrm{D}(K, s)$. Thus, since $\mathrm{D}(K, S)$ is faithful, we must have $k = c$. $\square$

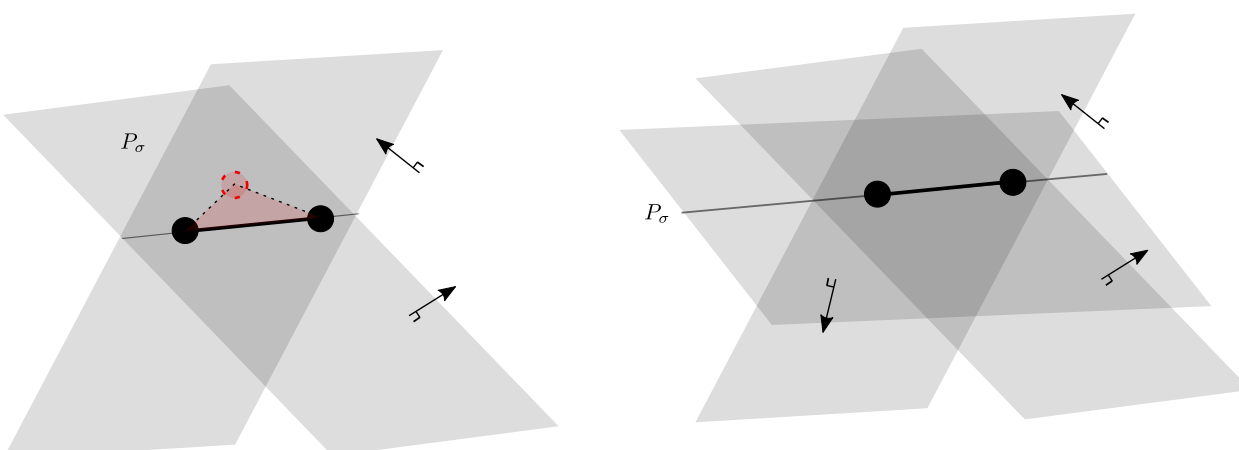


Figure 4.3: With only two directions in our direction set on the left side of the figure, we see that the envelope P_σ for the maximal edge σ is still three-dimensional (here, P_σ is the unbounded region above σ between the two planes). This means that we could construct an adversarial two-simplex (shown in pink) contained entirely in P_σ that would not be detected by either non-augmented descriptors in the set. On the right, the addition of a third direction reduces P_σ to a linear subspace.

A straightforward extension of Lemma 4.24 gives the following quantitative result.

Corollary 4.25 (Count of Non-Augmented Descriptors Per Simplex)**.** *Let $K \subset \mathbb{R}^d$ be a simplicial complex,* $\mathrm{D} \in \{\chi, \beta, \rho\}$*, and $S \subseteq \mathbb{S}^{d-1}$. If* $\mathrm{D}(K, S)$ *is faithful, then for each maximal simplex $\sigma \in K$ of dimension $k < d$, then there are at least $d - k + 1$ directions in S that are perpendicular to σ. Furthermore, if σ has dimension $k < d - 1$, these directions are additionally pairwise linearly independent with one another.*

Proof. Since $\mathrm{D}(K, S)$ represents K, we know by Lemma 4.24 that P_σ must be k-dimensional. Recall that P_σ is an intersection of closed half spaces. Then, for P_σ to be k-dimensional, it is a standard result in manifold theory that this intersection requires $d - k + 1$ hyperplanes, and if $k < d - 1$, these hyperplanes must not be parallel. The result follows. □

The following corollary is a simple extension of both Lemma 4.24 and Corollary 4.25.

Corollary 4.26 (Strict Lower Bound)**.** *Let $K \subset \mathbb{R}^d$ be a simplicial complex,* $\mathrm{D} \in \{\chi, \beta, \rho\}$*, and $S \subseteq \mathbb{S}^{d-1}$. Suppose that* $\mathrm{D}(K, S)$ *is faithful. Then $|S| \geq d + 1$, and this bound is strict.*

This bound is met whenever $K \in \mathbb{R}^d$ is (for example) a single vertex. However, the above bound is generally not met, and the minimum cardinality of a representative set of non-augmented descriptors is generally much higher. In practice, the requirement that each maximal k-simplex in a simplicial complex corresponds to $d - k + 1$ perpendicular directions is a huge hindrance when forming faithful sets (especially when compared to augmented descriptors, for which there is no such requirement). Counteracting the need for perpendicular directions is the fact that, as d increases, more simplices span common

hyperplanes, and thus perpendicular directions can increasingly be shared. The following organizes these observations into a more precise statement.

Theorem 4.27 (Lower Bound for Worst-Case Non-Augmented Descriptor Complexity)**.** *Let* $\mathrm{D} \in \{\chi, \beta, \rho\}$*. Then the worst-case cardinality of a minimal descriptor set is* $\Omega(d + n_1)$*.*

Proof. Suppose that $K \subset \mathbb{R}^d$ for $d > 2$ is a graph comprised of $n_1 > d - 1$ edges. By Lemma 4.24, since $\hat{\rho}(K, S)$ is faithful, the envelope of each maximal edge σ has the property $P_\sigma^S \setminus \overline{\sigma} = \emptyset$, i.e., the envelope contains the edge and no additional area. A necessary condition for this requirement is that the envelope of each edge must be one-dimensional. Furthermore, by Corollary 4.25, another necessary condition is that, for every edge σ, S must contain $d - 1 + 1 = d$ pairwise linearly independent directions perpendicular to σ. Let $S^\perp$ be a subset of directions in S satisfying the conditions of perpendicularity and one-dimensional envelopes such that $S^\perp$ has minimal cardinality.

We will bound the size of $S^\perp$ by considering how such a set could be constructed. Naïvely, we could try constructing $S^\perp$ by choosing d pairwise linearly independent directions for each edge, so that $|S^\perp| = n_1 d$. However, this is generally more directions than we need to satisfy the condition that each edge has d directions perpendicular to it; note that in $\mathbb{R}^d$, collections of $m < d$ edges are all parallel to a common m-plane (considered as vectors, the edges span this common plane), and thus, a single direction may be perpendicular to many edges at once. We also note that d pairwise linearly independent directions all perpendicular to an edge do not necessarily define a one-dimensional envelope for an edge. The combinatorial challenge then, is to find the minimum number of directions in $S^\perp$

that satisfy the requirement that all envelopes are one-dimensional, while noting that some directions may be perpendicular to more than one edge at a time.

We construct $S^{\perp}$ as follows. First, note that there is a $(d - n_1 - 1)$-sphere's worth of directions perpendicular to all edges in the set. In particular, since $d - n_1 - 1 > 0$, this is an uncountably infinite set of perpendicular directions. Choose any $d - 1$ pairwise linearly independent directions from this sphere to be included in the set $S^{\perp}$.

We need an additional perpendicular direction for each edge to bring the total for each edge up to d. However, in order to ensure the envelopes of each edge are one-dimensional, these additional directions must not be perpendicular to any hyperplane defined by subsets of more than one edge. This means we must consider a total of n_1 additional directions, so that $S^{\perp}$ has cardinality $d - 1 + n_1$. Since $|S^{\perp}|$ lower bounds $|S|$, we find $|S| \in \Omega(d + n_1)$. $\square$

Remark 3 (Why Bound n_1?)**.** *Bounding the number of edges for our construction in the proof of Theorem 4.27 may seem counterintuitive. After all, shouldn't more edges require more directions?*

In fact, in the asymptotic analysis, we see an interesting instance of this increased need for directions that are perpendicular to each edge being balanced by the increase of subsets of edges that are parallel to common $(d - 1)$-planes, and the increasing ability of a single direction to "count" for multiple edges' totals.

To construct $S^{\perp}$, we choose directions perpendicular to sets of $d - 1$ edges, choosing the sets so the envelopes of each edge are one-dimensional. In order to ensure each edge has d directions perpendicular to it, we perform this near-partition of the total edge set d times,

meaning that the asymptotic size of $S^{\perp}$ is (surprisingly),

$$d\left\lceil \frac{n_1}{d-1} \right\rceil \in \Omega(n_1).$$

In particular, the increased presence of subsets of edges parallel to the same hyperplane yields an asymptotic lower bound that is independent of d.

4.27.1 Augmented Descriptor Bounds

We now shift to augmented descriptors. We begin with the strict lower bound on the size of a representative set of non-augmented descriptors.

Lemma 4.28 (Strict Lower Bound)**.** *Let $K \in \mathbb{R}^d$ and let $\hat{\mathrm{D}} \in \{\hat{\chi}, \hat{\beta}, \hat{\rho}\}$. Suppose that $\hat{\mathrm{D}}(K, S)$ is faithful. Then $|S| \geq d$, and this bound is strict.*

Proof. No vertex in K can be described using fewer than d coordinates. Thus, a set of descriptors of type $\hat{\mathrm{D}}$ with cardinality less than d cannot be faithful for K_0, let alone K. To see that this bound is strict, consider the case where K is a single vertex; augmented descriptors generated by any d pairwise linearly independent directions will form a faithful set for the vertex (e.g., $S = \{\hat{\mathrm{D}}(K, e_1), \hat{\mathrm{D}}(K, e_2), \ldots, \hat{\mathrm{D}}(K, e_d)\}$). □

In back-of-the-envelope experiments, we find that minimum faithful sets of augmented descriptors for simple simplicial complexes often have cardinality $d + 1$. Independence from the size of the simplicial complex is not too surprising, since augmented descriptors have events corresponding to each simplex. In future work, we hope to classify the simplicial complexes for which this minimum cardinality of $2d - 1$ is met. It was through these

experiments, however, that we arrived at a surprising construction; a family of simplicial complexes whose minimum cardinality is linear in the number of vertices. The remainder of this section discusses this construction.

In what follows, we will use $\alpha_{i,j}$ to denote the angle that vector $v_j - v_i$ makes with the x-axis. When discussing the size of angles, we always assume they take value in $[0, 2\pi)$. We establish two preliminary observations. The first is a straightforward exercise in geometry.

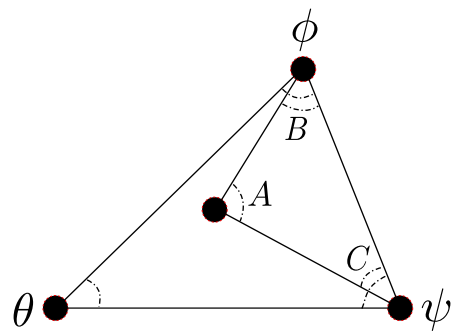


Figure 4.4: Nested triangles as discussed in Lemma 4.29

Lemma 4.29. *Consider a pair of nested triangles as in Figure 4.4 Then angle A is larger than θ, $\phi - B$, and $\psi - C$.*

Proof. Adding angles in the larger triangle, we see $\theta + \phi + \psi = \pi$. Then,

$$\theta + (\phi - B) + B + (\psi - C) + C = \pi$$

$$A + \theta + (\phi - B) + B + (\psi - C) + C = A + \pi$$

$$(A + B + C) + \theta + (\phi - B) + (\psi - C) = A + \pi$$

$$\pi + \theta + (\phi - B) + (\psi - C) = A + \pi$$

$$\theta + (\phi - B) + (\psi - C) = A.$$

All the terms in the last line are positive, meaning A is larger than θ, $\phi - B$, and $\psi - C$. $\square$

Our second observation is a specific instance of the general phenomenon that a simplicial complex stratifies the sphere of directions based on vertex order.

Remark 4. *Suppose that a simplicial complex $K \in \mathbb{R}^2$ contains an unconnected edge $[v_1, v_2]$. Then v_1 will be seen as a birth event in $\hat{\rho}(K, s)$ for all s in the half of $\mathbb{S}^1$ defined by the interval $H = [\alpha_{1,2} - \pi, \alpha_{1,2} + \pi]$ (so that for $s \in H$, we have $s \cdot v_1 \geq s \cdot v_2$) and as an instantaneous event for $s \in H^C$.*

We are now ready to describe the building block that forms the family of simplicial complexes used in our bound.

Construction 2 (Clothespin Motif)**.** *Let $K \subset \mathbb{R}^2$ be a simplicial complex with a vertex set $\{v_1, v_2, v_3, v_4\}$. Suppose that only v_3 is in the interior of the convex hull of $\{v_1, v_2, v_4\}$, and that the edge set consists of $[v_1, v_2]$ and $[v_3, v_4]$. See the left-hand side of Figure 4.5 for an illustration of this construction.*

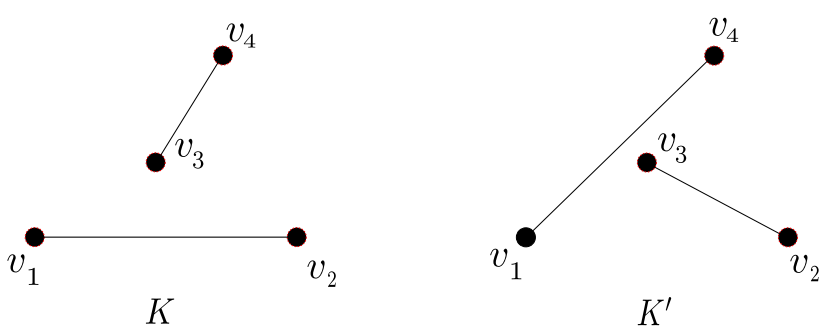


Figure 4.5: The two simplicial complexes considered in the proof of Lemma 4.30.

Construction 2 was built specifically for the following necessary condition for faithful sets of augmented descriptors. We state this condition in terms of $\hat{\rho}$'s in the following lemma.

Lemma 4.30 (Clothespin Representability)**.** *Let K be a clothespin motif, as in Construction 2, and suppose that $\hat{\rho}(K, S)$ is faithful. Then we must have at least one direction $s \in S$ such that the angle formed between s and the e_1-axis lies in the wedge-shaped region $[\alpha_{3,4} - \pi, \alpha_{2,3} - \pi] \cup [\alpha_{3,4} + \pi, \alpha_{2,3} + \pi]$.*

Proof. Let $K' \subset \mathbb{R}^2$ be a simplicial complex with the same vertex set as K, but with edges $[v_1, v_4]$ and $[v_2, v_3]$ (as in the right side of Figure 4.5). Each vertex of K and K' can correspond to either a birth event or an instantaneous event depending on the direction of filtration. Since $\hat{\rho}(K, S)$ must contain at least one $\hat{\rho}$ generated by a direction such that some event differs when filtering over K' rather than K. Otherwise, we would have $\hat{\rho}(K, S) = \hat{\rho}(K', S)$, meaning $\hat{\rho}(K, S)$ would not be faithful.

We proceed by considering each vertex v_i individually and determining subsets $R_i \subset \mathbb{S}^1$ such that, whenever $s \in R_i$, the event at $s \cdot v_i$ is different when filtering over K versus K', but for $s_* \notin R_i$, the type of event at $s_* \cdot v_i$ is the same between the two graphs. Figure 4.6 shows these regions, and in what follows, we define them precisely.

First, consider v_1. By Remark 4, $v_1 \in K$ corresponds to a birth event for all directions in the interval $B = [\alpha_{1,2} - \pi, \alpha_{1,2} + \pi]$, but $v_1 \in K'$ corresponds to a birth event for all directions in the interval $B' = [\alpha_{1,4} - \pi, \alpha_{1,4} + \pi]$. Then we write $R_1 = (B \setminus B') \cup (B' \setminus B)$, which is the wedge-shaped region such that for any $s \in R_1$, the type of event associated to $v_1 \in K$ and $v_1 \in K'$ differ, meaning $\hat{\rho}(K, s) \neq \hat{\rho}(K', s)$.

Using this same notation, we will find the wedge shaped region R_i for vertex $i \in [2, 3, 4]$ such that any direction from R_i generates $\hat{\rho}$'s that have different event types at vertex v_i

when filtering over K versus K'. A nearly identical argument shows that

$$R_1 = [\alpha_{1,2} - \pi, \alpha_{1,4} - \pi] \cup [\alpha_{1,2} + \pi, \alpha_{1,4} + \pi]$$

$$R_2 = [\alpha_{2,3} - \pi, \alpha_{2,1} - \pi] \cup [\alpha_{2,3} + \pi, \alpha_{2,1} + \pi]$$

$$R_3 = [\alpha_{3,4} - \pi, \alpha_{2,3} - \pi] \cup [\alpha_{3,4} + \pi, \alpha_{2,3} + \pi]$$

$$R_4 = [\alpha_{1,4} - \pi, \alpha_{3,4} - \pi] \cup [\alpha_{1,4} + \pi, \alpha_{3,4} + \pi]$$

(where we have again included R_1 for a complete list). Thus, any direction in the region defined by $W = \cup_{i=1}^{4} R_i$ would generate distinct $\hat{\rho}$'s when filtering over K versus K' and any direction in the region W^C would not. That is, for any $s \in W^C$, we have $\hat{\rho}(K, s) = \hat{\rho}(K', s)$, and for any $s_* \in W$, we have $\hat{\rho}(K, s_*) \neq \hat{\rho}(K', s_*)$.

We claim that $W = R_3$, i.e., exactly the region described in the lemma statement. This is a direct corollary to Lemma 4.29; the angles swept out by each regions correspond to the angles formed by pairs of edges in K and K'; in particular, the angle $\measuredangle v_2 v_3 v_4$ is the largest and geometrically contains the others. This means the extremal boundaries over all R_i's are formed by the angles $\alpha_{2,3} \pm \pi$ and $\alpha_{3,4} \pm \pi$, the defining angles of R_3. This means that $R_1, R_2, R_4 \subseteq R_3 = W$ and we have shown our claim. $\square$

We refer to this wedge shaped region W as a clothespin's *region of observability*. Crucially, this region is defined by the angle made by v_2, v_3, and v_4, meaning a different embedding of these vertices could result in a smaller region.

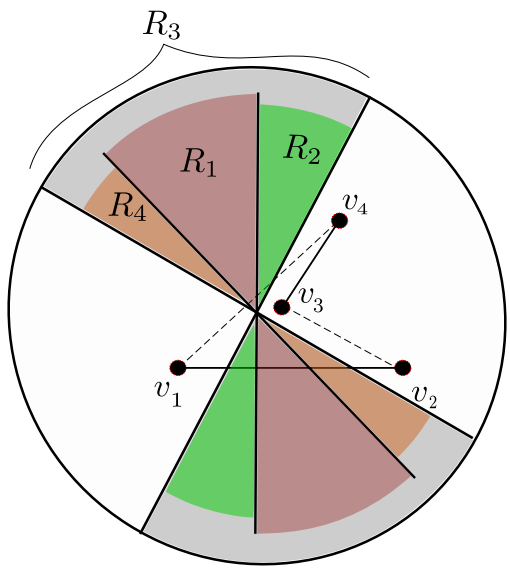


Figure 4.6: The regions of observability described in the proof of Lemma 4.30. K is shown as solid black edges and K' as dashed edges. For any lower-star filtration in a direction contained in R_i, the event at vertex v_i differs when considering K or K', thus, such directions are able to distinguish K from K'. Note that any direction outside the regions of observability (i.e., the non-shaded portions of the circle) is not able to distinguish K from K'.

Remark 5 (W Can be Arbitrarily Small)**.** *As the angle $\measuredangle v_2v_3v_4$ approaches zero, the region of observability, W, described in the proof of Lemma 4.30 also approaches zero.*

To construct a family of simplicial complexes, each of which must have at least $\Theta(n_0)$ augmented descriptors to form a faithful set, we use Remark 5 to knit together clothespin motifs (Construction 2) in the following way.

Construction 3 (Clothespins on a Semicircular Clothesline)**.** *Let $K_m \subset \mathbb{R}^2$ be a simplicial complex formed by m copies of Construction 2 (m clothespin motifs) such that the regions of observability for each clothespin do not overlap. Note this is possible for any m by Remark 5.*

See Figure 4.7 for an example of K_m for $m = 4$. Next, we show that this construction implies a lower bound on the number of $\hat{\rho}$'s needed to fully represent a simplicial complex.

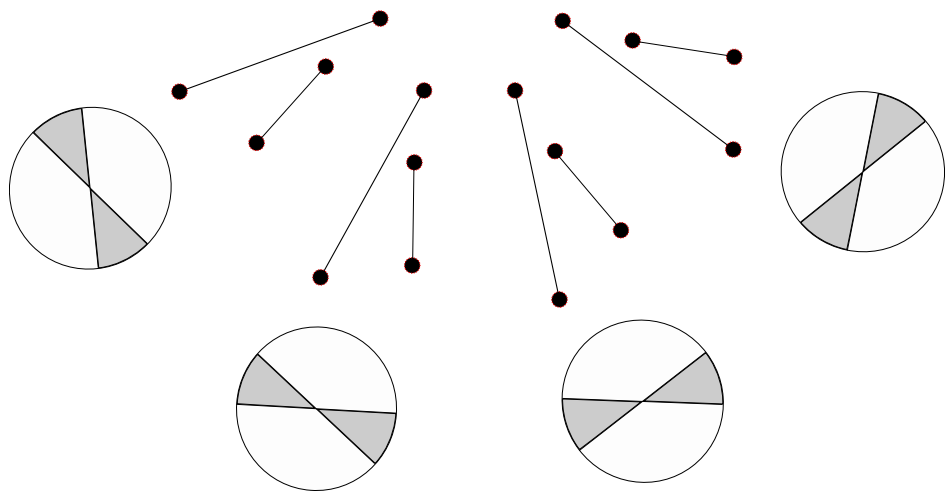


Figure 4.7: An example of Construction 3 for $m = 4$. The regions of observability are shown below each clothespin. By construction, each of these four double wedges define disjoint regions of $\mathbb{S}^2$.

Lemma 4.31 (Lower Bound for Worst-Case $\hat{\rho}$'s Complexity)**.** *Let K_m be as in Construction 3 and suppose $\hat{\rho}(K_m, S)$ is a faithful set. Then S must contain at least one direction in each of the m regions of observability, meaning that $|S| \geq m = n_0/4$. Thus, the size of a faithful set of $\hat{\rho}$'s for K_m is $\Omega(n_0)$.*

Note that by Theorem 4.21, the above lemma implies the following theorem.

Theorem 4.32 (Lower Bound for Worst-Case Augmented Descriptor Complexity)**.** *Let $\hat{\mathrm{D}} \in \{\hat{\chi}, \hat{\beta}, \hat{\rho}\}$. Then the worst-case cardinality of a minimal descriptor set is $\Omega(n_0)$.*

4.33 Discussion

In this chapter, we provide a framework for comparing abstract topological descriptors by their ability to represent simplicial complexes. Although this framework is general, we focus on six topological descriptors that are particularly relevant to applications and related work; the Euler characteristic curve, the Betti curve, the persistence diagram, the augmented Euler characteristic curve, the augmented Betti curve, and the augmented

persistence diagram. We provide a partial order on this set of six descriptors, including showing the strict inequality, $\hat{\beta} < \hat{\rho}$.

We then identify strict lower bounds for both non-augmented and augmented descriptors in the set of six, as well as asymptotic lower bounds for worst-case complexity of the non-augmented and augmented descriptors. The lower bounds for worst-case complexity quantify the observation that the non-augmented descriptors in our set of six are much weaker than the augmented descriptors, since they require many perpendicular directions to each maximal simplex. Furthermore, we are surprised to find a construction that requires a faithful set of augmented descriptors to have cardinality linear in the number of vertices: since augmented descriptors have events at the height of each vertex, one might suppose (and find to be generally true) that the size of a faithful set of augmented descriptors only depends on the embedding dimension, d.

In ongoing work, we hope to classify the simplicial complexes for which minimal faithful sets of augmented descriptors are independent of the size of the complex.

CHAPTER FIVE

ZIG-ZAG PERSISTENCE MODULES: COSHEAVES AND K-THEORY

5.1 Introduction

In this chapter, we demonstrate the utility of viewing persistent phenomena from the perspective of constructible (co)sheaves. In particular, we demonstrate how cosheaves provide a convenient interpretation of augmented descriptors of persistence modules and how cosheaves are a convenient setting for constructing invariants via algebraic K-theory. The present is in the same spirit of the program we first employed in [36], namely, applying stratified mathematics and higher algebra to TDA.

The use of cosheaves in TDA goes back at least to, e.g., [33,35]. The work of Curry and collaborators (such as work with Patel [21]), serves as an inspiration for our own perspectives. The key idea interpolating between persistence modules and constructible cosheaves is that of a stratified space. A persistence module $\{V_i\}_{i\in I}$ is obtained by sampling (or otherwise selecting a subset of) an underlying parameter space. In this chapter, we only consider persistence modules over locally finite posets. For concreteness, in the one-parameter case, consider $I \subset \mathbb{R}_{\geq 0}$ as a discrete selection of "instances" in our one-dimensional ray of "time." As our persistence module only changes at elements of I, it is locally constant on $\mathbb{R} \setminus I$, which is the defining property of a constructible cosheaf.

Constructible cosheaves are particularly nice mathematical objects for several reasons, chief among them is their equivalence to representations of the so called *entrance path*

category; this is known as "the" Exodromy Theorem. Any stratified space has an associated entrance path category and in good cases (e.g., when the space is a combinatorial or cubical manifold), the entrance path category is a straightforward combinatorial object; in many cases, it is simply a poset. The idea of *exodromy* goes back to communications by MacPherson, and proofs in different settings appear in the work of Curry and Patel [21], Treumann [60], Lurie [44], and Barwick with Glasman and Haine [6].

Given a parameter space (and a choice of sampling instances), exodromy allows us to consider all persistence modules/constructible sheaves on that space as a category of functors. Such categories of functors then inherit desirable properties from the target category. For instance, if we consider modules valued in vector spaces, the category of persistence modules is naturally an Abelian category. Abelian categories are the home of homological and homotopical algebra, so we are free to apply the tools of algebraic topology/homotopy theory, e.g., algebraic K-theory. The combinatorial nature of the entrance path category makes K-theory computations manageable and allows us to consider connections with other persistent invariants such as Euler curves and persistence diagrams.

We first introduce our constructions and results in the setting of one-dimensional parameter spaces. The resulting persistence modules are the zig-zag persistence modules of Carlsson and de Silva [14], which includes the more typically seen monotone (standard) modules. We then extend our methods and compute the K-theory of multi-parameter modules, building on the one-parameter results as a base case for an induction argument.

Readers interested in topological transforms of Chapter 3 and Chapter 4 may note that topological transforms are a special type of persistence module themselves, where,

depending on conventions, the parameter space is the stratified sphere of directions or the product of the sphere of directions with $\mathbb{R}$. Our computation of K-theory for d-parameter zig-zag persistence modules is phrased only for modules over the parameter space $\mathbb{R}^d$, but in future work we hope to extend the computation to arbitrary conically stratified d-manifolds. Thus, the results of this chapter may be useful for future work in the computation of other invariants of topological transforms as well as of more general persistence bundles (see [39]).

5.2 Why K-theory?

In the later part of this chapter, we compute the K-theory of the category of zig-zag modules. Here, we briefly overview why K-theory is a useful invariant.

K-theory began as simply as group completion of a monoid. Indeed, let $(M, \oplus)$ be a commutative monoid and define $K_0(M, \oplus)$ to be the unique (up to isomorphism) Abelian group, equipped with a monoid homomorphism from M, satisfying the universal property: for any Abelian group A and homomorphism (of monoids) $\varphi\colon M \to A$, there exists a unique group homomorphism factorization through $K_0(M, \oplus)$. This universal property is described as the *universal Euler characteristic* and is conveyed diagrammatically as follows:

$$\begin{tikzcd} (M,\oplus) \arrow[r, "\forall"] \arrow[d] & A \\ K_0(M,\oplus) \arrow[ur, dashed, "\exists!"'] & \end{tikzcd}$$

For instance, let $\mathcal{V}$ be the isomorphism classes of finite dimensional vector spaces over the field $\mathbb{R}$ (with direct sum) and $\varphi : \mathcal{V} \to \mathbb{Z}$ the rank function, then we have an induced map $K_0(\mathcal{V}) \to \mathbb{Z}$, which happens to be an isomorphism. Expanding this example to complexes, let $\mathcal{C}$ denote isomorphism classes of bounded complexes of $\mathbb{R}$-vector spaces. The natural extension of the rank function is the Euler characteristic, which again factors

uniquely through $K_0(\mathcal{C})$. (In the topological setting, the Chern character of a vector bundle is an example of such an additive map.) The universal property of K_0 extends to categories equipped with a symmetric monoidal structure, since isomorphism classes of objects in such a category form a commutative monoid.

K-theory is more than just a single Abelian group, but rather a spectrum, $\mathbb{K}(\mathcal{C})$, associated to a category (equipped with additional structure). Recall that spectra are the central objects of homotopy theory. The homotopy groups of $\mathbb{K}(\mathcal{C})$ define the K-groups of $\mathcal{C}$, i.e., $K_n(\mathcal{C}) := \pi_n(\mathbb{K}(\mathcal{C}))$. To first approximation, spectra can be thought of as the objects that parametrize cohomology theories. As such, they, so K-theory in particular, admit a wealth of computational tools, refined structures, and interpretations from algebraic topology. Cohomology theories are also the natural home for obstruction/anomaly theory and in this way, K-theory has become a central tool in topology (index theory, finiteness obstructions) and algebraic number theory (class field theory).

When refined to the level of spectra, K-theory has a remarkable additive structure with respect to split short exact sequences of categories. (We discuss this in Section 5.22, see also [11].) Combined with its property as the universal Euler characteristic, K-theory is the *universal additive invariant* of (Waldhausen) categories.

5.2.1 Flavors and History of K-theory

There are several constructions of K-theory; we trace the history to Waldhausen's construction, the tool we use in the present chapter. See the canonical texts of Rosenberg [53] and Weibel [63] for historical references and more details on the development of K-theory.

The genesis of K-theory came in the late 1950's and early 1960's through the work

of Grothendieck in complex (algebraic) geometry and Atiyah and Hirzebruch in topology. Algebraic K-theory (the kind used in this chapter) is an extension of Grothendieck's ideas to build a family of functors from rings to Abelian groups $K_i : \mathsf{Ring} \to \mathsf{Ab}$. While Grothendieck only defined K_0, suitable definitions for K_1 and K_2 were found by the mid 1960's; the contributions of Bass, Schanuel, and Milnor are most notable. (Bass and Karoubi also gave definitions of negative K-theory, $K_{-n}(R)$.) Milnor had given a definition of higher K-groups, though this *Milnor* K-theory is only a summand of the now accepted definition of higher K-theory. Defining higher K-groups was a major open problem in the early 1970's, which was first solved by Dan Quillen, and is known as the $+$-construction. Given a ring, R, Quillen produced a space, $BGL(R)^+$, whose homotopy groups recover/define the K-theory of R.

Quillen quickly followed his $+$-construction with the Q-construction. The Q-construction takes as input an exact category, $\mathcal{C}$, e.g., the category of finitely generated projective modules for a ring, and outputs a space, $\Omega BQ\mathcal{C}$, whose homotopy groups define K-theory. Quillen used the Q-construction to prove many fundamental results in algebraic K-theory that restricted to those for rings, as he also proved that $+ = Q$, that is, the Q-construction is a strict generalization of $+$-construction for rings.

The next revolution in algebraic K-theory came through Waldhausen's work in manifold topology [62]. Published in 1985, Waldhausen gave a construction that takes as input categories with structure that generalizes that of exact categories—nowadays called Waldhausen categories—and outputs a spectrum (the basic building block of homotopy theory) whose homotopy groups define the corresponding K-groups. (Segal some 15 years earlier used his Γ-objects to produce a K-theory spectrum in certain cases.) Waldhausen's

construction is often referred to as the $S_\bullet$-construction and we give a brief overview in Section 5.22. The $S_\bullet$-construction is a strict extension of the Q-construction. Perhaps most significantly, the $S_\bullet$-construction satisfies an additivity result for split short exact sequences; this result has become a central tool in algebraic K-theory.

Finally, we note that K-theory has been extended to the higher categorical/homotopical algebraic setting as well. The work of Blumberg, Gepner, and Tabuada [11] proves that K-theory satisfies certain universal properties, such as additivity, (and hence is essentially uniquely defined by such properties) in this setting.

5.2.2 K-theory and persistence

Through the work of Patel, Bubenik, and collaborators, K-theoretic considerations have started to appear in the TDA literature. Patel considered the Grothendick group, i.e., K_0, of one-dimensional persistence modules valued in symmetric monoidal categories [49]. Subsequently, in [13], Bubenick and Milićević show that the category of persistence modules over any preorder is Abelian. The key idea—which we use below as well—is that functor categories inherit many of the properties of the target category, so if the target is Abelian or Grothendieck, i.e., AB5 with a generator, then the functor category whose domain is a preorder (or any small category) is Abelian or Grothendieck. It would be interesting to apply Quillen's Q-construction to these categories of persistence modules and compare the resulting K-theories to our results below. (We note [13] contains much more than just outlined, e.g., the authors prove an embedding theorem in the vein of the Gabriel–Popescu Theorem.)

More relevant for us is the recent article [12] by Bubenik and Elchesen. In this work, the group completion of the monoid of persistence diagrams is described, i.e., $K_0(\mathsf{Diag})$ is

defined (semi-)explicitly. Points in diagrams are counted with multiplicity, so the binary operation is simply induced by the map $+\colon \mathbb{N}_0 \times \mathbb{N}_0 \to \mathbb{N}_0$, "overlaying" a diagram on top of another. The input data for the construction of Bubenik and Elchesen is somewhat flexible, so one can talk about diagrams (and their group completions) indexed by the entire first quadrant, the integers, etc. We make contact with this work in Section 5.67.1 below.

5.2.3 What we do

We have aimed to illustrate the connection between persistence modules and cosheaves and the utility of this interplay. To this end, we accomplish the following.

coSheaves from filtrations The relevance of cosheaves in TDA has been advocated by Curry and others for a number of years. In Section 5.28, we give explicit constructions of constructible cosheaves associated to persistence modules. We are particularly interested in persistence modules arising from index filtrations of spaces. In Section 5.34.0.1, we describe the *augmented filtration cosheaf*, which records both non-instantaneous and instantaneous events. We flag the recent work of Berkouk, Ginot, and Oudot [9] where level-set persistence is recast in terms of sheaves over $\mathbb{R}$.

Equivalence Theorem We prove an equivalence of categories between a localization of the category of zig-zag modules a la Carlsson and de Silva and constructible cosheaves on $\mathbb{R}^d$. The explicit statement of the result is Theorem 5.42. This result is stated in passing in the recent work of Curry and Patel [21, Example 6.3], and we make it explicit with proof. We hope the proof is as interesting to the reader as the result, though it uses techniques that are different from the rest of the chapter.

One motivation for this result is to argue that our K-theoretic computations that follow

deserve to be called *the* K-theory of zig-zag and multi-parameter modules.

K-theory of Multi-Parameter Zig-Zag Modules In Section 5.51, we define and compute K-theory of persistence modules, viewed as constructible cosheaves on a stratified parameter space. We use Waldhausen's $S_\bullet$ construction of K-theory. A key input is *additivity*, in this case with respect to strata. For instance, in the cases considered, the group K_0 is the free abelian group on the strata of the parameter space (Theorems 5.60 and 5.64). We further investigate this result in the specific instances of both Vect-valued persistence modules and persistence modules valued in pointed sets. The higher K-groups do not vanish but rather are given by the algebraic K-theory of fields and/or the sphere spectrum (in the cases Vect- and pointed set-valued modules, respectively). In ongoing work, we aim to interpret these higher K-groups as arising from data.

Euler Manifolds and Virtual Diagrams We conclude the body of the chapter by connecting our K-theoretic work back to recent work in TDA. First, we show how the Euler characteristic curve of a persistence module (or higher parameter analogues like Euler characteristic surface, etc.) has a natural interpretation as a class in K-theory. This is as expected, e.g., Kashiwara and Schapira [41] prove that K_0 is isomorphic to constructible functions via a local Euler index. With this observation, we define an *Euler class* for arbitrary persistence modules regardless of dimension in Definition 5.67. Lastly, Section 5.67.1 constructs a group homomorphism from K_0 of persistence modules to Bubenik and Elchesen's abelian group of virtual persistence diagrams.

Conventions

We assume the reader has some familiarity with algebraic topology, and freely use concepts from Hatcher's standard text [38]. Throughout, we will let $\mathsf{Vect}_{\mathbb{F}}$ be the category of finite dimensional vector spaces over the field $\mathbb{F}$ and linear maps. Much of our work does not depend on making a choice of field and we simply use the notation Vect. Unless otherwise noted, we will assume all stratified spaces are combinatorial or cubulated manifolds equipped with their native stratification, notions we define in the next section.

5.3 Constructible coSheaves

This section is a terse introduction to terminology and notation used. Examples and further details are abundantly available, e.g., [21] or [36].

5.3.1 Stratified/Constructible Basics

Definition 5.4. *Let $(\mathcal{P}, \leq)$ be a poset. The* upward closed topology *on $(\mathcal{P}, \leq)$ is defined as follows: $U \subseteq \mathcal{P}$ is open if and only if for all $u \in U$, $\mathcal{P}_{u\leq} := \{p \in \mathcal{P} \mid u \leq p\} \subseteq U$.*

The upward closed topology is also known as the *Alexandrov* topology.

Definition 5.5. *A* stratified topological space *is a triple $(X \xrightarrow{\phi} \mathcal{P})$ consisting of*

- *a paracompact, Hausdorff topological space, X,*
- *a poset $\mathcal{P}$, equipped with the upward closed topology, and*
- *a continuous map $X \xrightarrow{\phi} \mathcal{P}$.*

Any topological space is stratified by the terminal poset consisting of a singleton set. Moreover, the simplices of a simplicial complex, K, come equipped with the structure of a

poset, and we call the resulting stratification of K the *native stratification*, denoted $\mathsf{Nat}(\mathsf{K})$.

Definition 5.6. *Given a stratified topological space $\phi : X \to \mathcal{P}$, and any $p \in \mathcal{P}$, the p-*stratum*, X_p, is defined as*

$$X_p := \phi^{-1}(p).$$

Example 5. *For $n \in \mathbb{N}$, let $[n]$ denote the totally ordered set $\{0 < 1 < \cdots < n\}$. Consider a stratified circle, $S^1 \to [1]$, stratified by v, a single vertex, and α, the arc that is the complement of v. This example is illustrated in Figure 5.1. (So the map $\phi : S^1 \to [1]$ is given by $v \mapsto 0$ and $\alpha := S^1 \setminus \{v\} \mapsto 1$.) The 0-stratum is the vertex v and the 1-stratum is the arc α, i.e., $S^1_0 = \{v\}$ and $S^1_1 = S^1 \setminus \{v\} = \alpha$.*

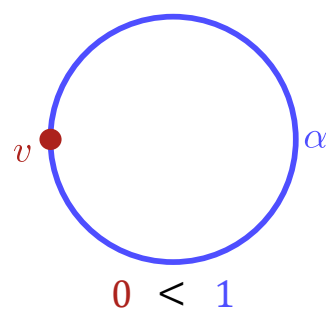


Figure 5.1: A stratified circle, $S^1 \to [1]$ as in Example 5, where $v \mapsto 0$ and $\alpha \mapsto 1$.

Definition 5.7. *A* map of stratified topological spaces $(\phi\colon X \to \mathcal{P})$ *to* $(\psi\colon Y \to \mathcal{Q})$ *is a pair of continuous maps (f_1, f_2) making the following diagram commute.*

$$\begin{tikzcd} X \arrow[r, "f_1"] \arrow[d, "\phi"'] & Y \arrow[d, "\psi"] \\ \mathcal{P} \arrow[r, "f_2"'] & \mathcal{Q}. \end{tikzcd}$$

A map between stratified topological spaces is a stratified homeomorphism *if it admits a two-sided stratified inverse.*

Definition 5.8 (Conically Stratified Space)**.** *Let $X \to \mathcal{P}$ be a stratified space and consider a point $x \in X_p \subseteq X$. The space X is* conically stratified at x *if there exists an open neighborhood, U_x, of x and a stratified homeomorphism $U_x \cong Z \times CY$ where Z is a topological space and CY is the cone on a space Y stratified by $\mathcal{P}_{>p}$. A space is called* conically stratified *if it is conically stratified at all of its points.*

Definition 5.9. *Let L be a polyhedron, so every point admits a neighborhood which is a finite union of simplices. Recall that a map $f\colon L \to \mathbb{R}^n$ is* piecewise linear (PL) if there exists a stratification of L such that restricted to each strata f is linear.

Definition 5.10. *A* piecewise linear (PL) manifold *is a topological manifold that admits an atlas where transition functions are piecewise linear.*[1]

Completely analogously to smooth manifolds, PL manifolds form a category with morphisms being PL maps and isomorphisms being PL homeomorphisms.

In [37], we discussed *combinatorial manifolds*, or triangulated PL manifolds since our main object of study was one-parameter zig-zag modules, which naturally endow $\mathbb{R}$ with a triangulation – in the one-dimensional setting, a triangulation is also a cubulation. In the present work however, we shift our focus d-parameter zig-zag modules, and therefore shift our focus to cubulated PL manifolds.

The terms *cubical manifold* and *cubulated manifold* are used with various conventions in the literature. Cubical complexes are sometimes considered in the specific case that they have non-positive curvature, or to are *foldable* (meaning they admit a stratified map onto

[1] Admiting a PL atlas is equivalent to specifying a class of trangulations of the underlying manifold which is stable under subdivision.

the n-cube). See, e.g., [5, 56, 65]. We take no such additional constraints, and by *cubical manifold*, we simply mean the following.

Definition 5.11 (Cubical Manifold)**.** *A* cubical manifold X *is a cubulated PL manifold. In other words, a cubical manifold is a PL manifold X along with a cubical complex $\mathfrak{C}$ and a PL homeomorphism $\mathfrak{C} \to X$. Note that the manifold X inherits a native stratification from $\mathfrak{C}$.*

Remark 6. *As discussed in [4], every Whitney stratified manifold is conically stratified. In particular, a cubical manifold X is conically stratified.*

For further details on PL manifolds stratified by polyhedra, see [54].

Definition 5.12 (Connected Ambient Stratification [36])**.** *Let $(S \xrightarrow{\varphi} \mathcal{Q})$ be a stratified space, $S \hookrightarrow X$ a topological embedding, and $\pi_0(X \setminus S) = \mathcal{A}$. Define the poset, $\mathcal{Q}^{\pitchfork}$, as the set $\mathcal{Q} \amalg \mathcal{A}$, subject to the following generating relations:*

1. *The relations of $\mathcal{Q}$;*
2. *For $\ell \in \mathcal{Q}$ and $\alpha \in \mathcal{A}$, $\ell \leq \alpha$ if and only if $\varphi_S^{-1}(\ell) \subseteq \overline{\alpha}$, i.e., the ℓ-stratum is in the closure of the connected component indexed by α.*

There is an obvious extension of the map φ, $\psi_S \colon X \to \mathcal{Q}^{\pitchfork}$ and we call this stratification the connected ambient stratification*. We denote this stratified space by $(X, S)^{\pitchfork}$.*

Example 6. *Consider a discrete subset $I \subset \mathbb{R}$. The resulting stratified space, $(\mathbb{R}, I)^{\pitchfork}$, is a cubical (and combinatorial) manifold.*

The stratified space of Example 6 is a main object of study in the one-parameter setting. In Section 5.28, Definition 5.30, we extend this type of construction to the d-fold product of such stratified spaces, and is our main object of study, generally.

Definition 5.13. *Let X be a topological space, $\mathsf{Op}(\mathsf{X})$ the poset of open sets in X, and Ω a category. A* precosheaf on X valued in Ω *is a functor $\mathcal{F}\colon \mathsf{Op}(\mathsf{X}) \to \Omega$. A precosheaf is furthermore a cosheaf if, for each open set $U \subseteq X$ and any open cover of U, $\{U_i \to U\}$, there is an equivalence in Ω given by*

$$\mathsf{colim}\left(\coprod_{i,j} \mathcal{F}(U_i \cap U_j) \rightrightarrows \coprod_i \mathcal{F}(U_i)\right) \xrightarrow{\sim} \mathcal{F}(U).$$

For what remains, we will assume Ω is a category where cosheafification exists. (Cosheafification is quite subtle, even compared to its dual notion of sheafification.) In particular, we will later focus on the case that $\Omega = \mathsf{Vect}$ or $\Omega = \mathsf{Set}_*$.

Lemma 5.14. *Let $\mathcal{B}$ be a basis for the topology of a space X and let $\mathcal{F}$ be a cosheaf on the poset determined by $\mathcal{B}$. Then, up to isomorphism, there is a unique extension of $\mathcal{F}$ to a cosheaf defined on all open sets of X.*

Proof. Let $U \subseteq X$ be open. Then there exist basis elements $B_1, \ldots, B_n \in \mathcal{B}$ such that $\cup_{i=1}^n B_i = U$. We extend the cosheaf via the assignment $\mathcal{F}(U) = \mathcal{F}(\cup_{i=1}^n B_i)$, where $\mathcal{F}$ of the union is defined via the coequalizer,

$$\mathcal{F}(\cup_{i=1}^n B_i) := \mathsf{colim}\left(\coprod_{i,j} \mathcal{F}(B_i \cap B_j) \rightrightarrows \coprod_i \mathcal{F}(B_i)\right).$$

Any other collection of basis elements $B'_1, \dots B'_m \in \mathcal{B}$ such that $\cup_{i=1}^m B'_i = U$ would also be an open cover of $\cup_{i=1}^n B_i$, and since $\mathcal{F}$ satisfies the cosheaf condition for $\mathcal{B}$ and by uniqueness of colimits, $\mathcal{F}(\cup_{i=1}^m B'_i) \cong \mathcal{F}(\cup_{i=1}^n B_i)$. Thus, the extension of $\mathcal{F}$ from being defined only on $\mathcal{B}$ to a cosheaf defined for all X is unique up to isomorphism.

To further see that this extension is a cosheaf (rather than just a precosheaf), we observe that any open cover of U can be rewritten as a collection of unions of basis elements. Since $\mathcal{F}$ satisfies the cosheaf condition for basis elements, it therefore satisfies the cosheaf condition for any open cover of U. Since U and our choices of covers were arbitrary, we have shown that $\mathcal{F}$ extends uniquely to a cosheaf defined over all of X. $\square$

Note that the idea of the preceeding lemma can also be thought more concisely of in terms of a Kan extension diagram:

$$\begin{tikzcd} \mathcal{B} \arrow[r, "\mathcal{F}"] \arrow[d, hook] & V \\ \mathsf{Op(X)}. \arrow[ur, dashed, "\exists!"'] & \end{tikzcd}$$

We end with two additional definitions related to (pre)cosheaves.

Definition 5.15. *Let $M \to \mathcal{P}$ be a stratified space (not necessarily conical or simplicial) and $\mathcal{F}$ a cosheaf on M. The cosheaf, $\mathcal{F}$, is* constructible *if it is locally constant when restricted to any stratum of $M \to \mathcal{P}$, i.e., given $p \in \mathcal{P}$ and $x \in M_p$ there exists a neighborhood $p \in U \subseteq M_p$ such that $\mathcal{F}|_U$ is constant.*

Definition 5.16 (Costalk at a Point)**.** *Let $\mathcal{F}$ be a (pre)cosheaf on X and let $p \in X$. Then the* costalk of $\mathcal{F}$ at p *is defined by*

$$\mathcal{F}_p := \lim_{U \ni p} \mathcal{F}(U).$$

5.16.1 Operations on coSheaves

Given a continuous map $\xi\colon X \to Y$, there is in an induced functor on the posets of opens $\hat{\xi}\colon \mathsf{Op(Y)} \to \mathsf{Op(X)}$ given by preimages with respect to ξ.

Definition 5.17. *Let $\xi\colon X \to Y$ be a continuous map and $\mathcal{F}$ a (pre)cosheaf on X. The* pushforward *of $\mathcal{F}$, $\xi_*\mathcal{F}$, is the (pre)cosheaf on Y given by $\xi_*\mathcal{F} := \mathcal{F} \circ \hat{\xi}$.*

There is a contravariant functor as well associated to a map $\xi\colon X \to Y$: the pullback $\xi^*\colon \mathsf{coShv(Y)} \to \mathsf{coShv(X)}$. As a continuous map is not necessarily an open map, ξ^* is (slightly) more involved to define: it is the limit over opens containing $\xi(U)$ for $U \subseteq X$ an open. Only pushforwards appear below.

Example 7. *Let $p \in X$ be a point in the topological space X and $i\colon p \hookrightarrow X$ the inclusion map. Let $\mathcal{F}$ be a cosheaf on X, then $i^*\mathcal{F}$ is the costalk at p of $\mathcal{F}$, $\mathcal{F}_p$. Let W be a cosheaf on p, then i_*W is a* skyscraper cosheaf *on X.*

Example 8. *Let $\xi\colon (\phi\colon X \to \mathcal{P}) \to (\psi\colon Y \to \mathcal{Q})$ be a stratified map and $\mathcal{F}$ a constructible cosheaf on X. Although the pullback of a constructible cosheaf is always constructible, it is not necessarily the case that $\xi_*\mathcal{F}$ is constructible on Y.*

- *Consider the inclusion* $\iota\colon [0,1/2) \hookrightarrow [0,1)$ *and let* $\underline{V}$ *be a nonzero constant cosheaf on* $[0,1/2)$*. Further stratify* $[0,1/2)$ *and* $[0,1)$ *with zero-stratum* $\{0\}$ *and one-stratum* $(0,1/2)$ *(resp.* $(0,1)$*). The cosheaf* $\iota_*\underline{V}$ *is not locally constant on* $(0,1)$ *as*

$$\iota_*\underline{V}(0,1/4) = V, \qquad \textit{while} \qquad \iota_*\underline{V}(3/4,1) = 0.$$

- *Constructibility is preserved by pushforwards in certain cases. Let* $C\colon [0,4] \to [0,3]$ *be the "elementary collapse" of the interval* $[2,3]$*, i.e.,*

$$C(t) = \begin{cases} t, & \textit{if } 0 \le t < 2 \\ 2, & \textit{if } 2 \le t \le 3 \\ t-1, & \textit{if } 3 < t \le 4 \end{cases} .$$

  The map C *is a stratified map with respect to the (connected) ambient stratifications induced by* $\{0,1,2,3,4\} \subset [0,4]$ *and* $\{0,1,2,3\} \subset [0,3]$*. Let* $\mathcal{F}$ *be any constructible cosheaf on* $[0,4]$*. It is straightforward to verify that* $C_*\mathcal{F}$ *is constructible on* $[0,3]$*.*

5.17.1 Entrance Paths and Their Representations

Given a stratified space, $M \to \mathcal{P}$, an entrance path is a continuous path in M such that it for all time it stays in a stratum or enters into a deeper (with respect to $\mathcal{P}$) stratum.

Definition 5.18 (Entrance Path Category)**.** *Let* $M \to \mathcal{P}$ *be a stratified space. The* entrance path category *of* $M \to \mathcal{P}$*, denoted* $\mathsf{Ent}(\mathsf{M},\mathcal{P})$ *has, as objects, the points of* M *and, as morphisms, (elementary) homotopy classes of entrance paths.*

Theorem 5.19 (Theorem 6.1 of [21])**.** *Let $M \to \mathcal{P}$ be a conically stratified space and V a category. There is an equivalence of categories*

$$\mathsf{cShv}^{\mathsf{V}}_{\mathsf{cbl}}(\mathsf{M}, \mathcal{P}) \xrightarrow{\sim} \mathsf{Fun}(\mathsf{Ent}(\mathsf{M}, \mathcal{P}), \mathsf{V})$$

between constructible cosheaves on M and representations of its entrance path category.

Definition 5.20 (Combinatorial Entrance Path Category)**.** *Let $M \to \mathcal{P}$ be a cubical manifold. The* combinatorial entrance path category*, $\mathsf{Ent}_{\Delta}(\mathsf{M}, \mathcal{P})$ has as objects the strata of M and a morphism $\mathfrak{c} \to \mathfrak{c}'$ whenever $\mathfrak{c}'$ is a face of $\mathfrak{c}$.*

Lemma 5.21. *Let $M \to \mathcal{P}$ be a cubical manifold. Then there is an equivalence of categories $\mathsf{Ent}(\mathsf{M}, \mathcal{P}) \xrightarrow{\sim} \mathsf{Ent}_{\Delta}(\mathsf{M}, \mathcal{P})$.*

Proof. Define a functor $F : \mathsf{Ent}(\mathsf{M}, \mathcal{P}) \to \mathsf{Ent}_{\Delta}(\mathsf{M}, \mathcal{P})$, where the image of a point $x \in \mathsf{Ob}(\mathsf{Ent}(\mathsf{M}, \mathcal{P}))$ is the unique (open) cube $\mathfrak{c}$ containing x, and the image a morphism $x \to y$ is the combinatorial entrance path from $F(x) \to F(y)$ (well-defined since the cube containing y must be a face of the cube containing x in order to be an entrance path). We claim that F is fully faithful and essentially surjective. Again, let $x \in \mathfrak{c}$, $y \in \mathfrak{c}'$, so that $\mathfrak{c}'$ is a face of $\mathfrak{c}$. Since $\mathfrak{c}'$ and $\mathfrak{c}$ are face-coface pairs of a non-degenerate cubulation, the subspace $\mathfrak{c}' \cup \mathfrak{c}$ deformation retracts onto $\mathfrak{c}'$, meaning there is a unique homotopy class of entrance paths from x to y. Furthermore, there is a unique morphism $\mathfrak{c} \to \mathfrak{c}'$ in $\mathsf{Ent}_{\Delta}(\mathsf{M}, \mathcal{P})$, i.e., F is fully faithful. Next, we observe that, for any cube $\mathfrak{c} \in \mathsf{Ob}(\mathsf{Ent}_{\Delta}(\mathsf{M}, \mathcal{P}))$, we can always find a point x so that $F(x) = \mathfrak{c}$ (for example, let x be the barycenter of $\mathfrak{c}$). That means we have shown F is also essentially surjective, and thus gives an equivalence of categories. $\square$

Remark 7. *One might hope that there is an equivalence of entrance and combinatorial entrance path categories for a larger class of stratifications. However, even when a space is stratified by a "degenerate" cubulation, or even triangulation, this equivalence does not generally hold. For instance, consider the stratified space shown in Figure 5.1,* $S^1 \to [1]$*, stratified by* v*, a single vertex, and its complement* α*, an open arc. Let* $x \in \alpha$*. Then there are two distinct homotopy classes of entrance paths from* $x \to v$ *in* $\mathsf{Ent}(\mathsf{M}, \mathcal{P})$*, but only one combinatorial entrance path given by the face relation in* $\mathsf{Ent}_{\Delta}(\mathsf{M}, \mathcal{P})$*.*

One useful interpretation of Lemma 5.21 is that the data of a constructible cosheaf on a cubical manifold is simply a specification of costalks on the stratifying poset and a specification of linear maps between them.

5.22 Waldhausen's K-theory

In this section, we outline Waldhausen's construction of algebraic K-theory. In particular, we build to a fundamental additivity result known as Waldhausen Additivity. The work of Waldhausen was first published in 1985 [62]. Our notation follows the much more recent article of Fiore and Pieper [32].

Definition 5.23 (Waldhausen Category)**.** *A Waldhausen category,* C*, is a category equipped with a subcategory of weak equivalences,* $w(\mathsf{C})$*, a subcategory of cofibrations,* $co(\mathsf{C})$*, and a distinguished zero object. Further, the triple* $(\mathsf{C}, co(\mathsf{C}), w(\mathsf{C}))$ *must satisfy*

1. *Every isomorphism in* C *is a cofibration;*

2. *Each object* $c \in \mathsf{C}$ *is cofibrant, i.e., the unique map* $0 \to c$ *is a cofibration;*

3. *Cokernels exist and define cofibration sequences; and*

4. Weak equivalences glue along cofibrations. That is, given the following diagram,

$$\begin{tikzcd} \mathsf{A} \arrow[d, "w.e."] & \mathsf{B} \arrow[l] \arrow[r, "cof."] \arrow[d, "w.e."] & \mathsf{C} \arrow[d, "w.e."] \\ \mathsf{A}' & \mathsf{B}' \arrow[l] \arrow[r, "cof."] & \mathsf{C}', \end{tikzcd}$$

the induced map between pushouts $\mathsf{C} \amalg_{\mathsf{B}} \mathsf{A} \to \mathsf{C}' \amalg'_{\mathsf{B}} \mathsf{A}'$ *is a weak equivalence.*

Example 9 (Vect)**.** *Consider the category* Vect*, with* $w(\mathsf{Vect})$ *the of vector spaces with isomorphisms,* $co(\mathsf{Vect})$*, the subcategory of vector spaces with injective maps, and distinguished zero object as the zero vector space. Condition 1. of Definition 5.23 is satisfied since every isomorphism is injective. Condition 2. is satisfied since every map out of the zero vector space is injective. Condition 3. is met since cokernels clearly exist in* Vect*, and the sequence* $V \xrightarrow{f} W \to coker(f)$ *is a cofiber sequence. Finally, we consider Condition 4. Recall that, for* $f : \mathsf{B} \to \mathsf{C}$*,* $g : \mathsf{B} \to \mathsf{A}$ *and for all* $b \in B$*, a pushout in vector spaces is given by* $\mathsf{C} \amalg_{\mathsf{B}} \mathsf{A} \cong \mathsf{A} \oplus \mathsf{C}/(f(b) \sim g(b))$*.*

Let $f' : \mathsf{B}' \to \mathsf{C}'$*,* $g' : \mathsf{B}' \to \mathsf{A}'$*,* $\phi : A \to A'$*, and* $\psi : C \to C'$ *denote the maps indicated in the diagram above. The induced map between pushouts is the unique map making the cube of pushout diagrams commute. Specifically, it is the map* $\mathsf{A} \oplus \mathsf{C}/(f(b) \sim g(b)) \to \mathsf{A}' \oplus \mathsf{C}'/(f'(b') \sim g'(b'))$ *defined by sending the equivalence class* $[a, c]$ *to the equivalence class* $[\phi(a), \psi(c)]$*. Since both* ϕ *and* ψ *are isomorphisms, and by commutativity of all faces on the cube of pushouts, this map is an isomorphism.*

Thus, Vect *with the given subcategories of weak equivalences and cofibrations is an example of a Waldhausen category.*

Waldhausen categories are a more general setting for algebraic K-theory than Abelian and exact categories. In particular, if R is a commutative ring and $\mathcal{M}$ is the category of finitely generated modules, then declaring weak equivalences to be isomorphisms and cofibrations to be monomorphisms makes $\mathcal{M}$ into a Waldhausen category. The following lemma addresses functor categories and is a straightforward verification.

Lemma 5.24. *Let* A *be a Waldhausen category and* D *a small category. The category of functors* $\mathsf{Fun}(\mathsf{D}, \mathsf{A})$ *is a Waldhausen category where*

(z) The zero object $Z \in \mathsf{Fun}(\mathsf{D}, \mathsf{A})$ *is the constant functor to the zero object in* A*;*

(w) A natural transformation $\eta\colon F \Rightarrow G$ *is a weak equivalence if and only if for each* $d \in \mathsf{D}$*, the map* $\eta_d\colon F(d) \to G(d)$ *is an isomorphism; and*

(c) A natural transformation $\alpha\colon F \Rightarrow G$ *is a cofibration if and only if for each* $d \in \mathsf{D}$*, the map* $\alpha_d\colon F(d) \to G(d)$ *is a monomorphism.*

Given a Waldhausen category C, one can compute the K-theory of C using *Waldhausen's* $S_\bullet$ *construction*, which we briefly outline here. Define $S_n\mathsf{C}$ to be the category with objects length-n sequences of cofibrations in C, and morphisms natural transformations between sequences. Furthermore, for each pair of objects $A_i, A_j \in S_n\mathsf{C}$, we have a choice of $A_{ij} = A_i/A_j$ that is compatible with other choices of quotients. One can show that each $S_n\mathsf{C}$ is an example of a Waldhausen category itself, and furthermore, ranging over n, we can form a *simplicial* Waldhausen category, denoted $S_\bullet\mathsf{C}$ by defining boundary maps $S_n\mathsf{C} \to S_{n-1}\mathsf{C}$. This is known as the $S_\bullet$ *construction* (see [63, Ch. IV]).

The simplicial Waldhausen category $S_\bullet \mathsf{C}$ has a subcategory, $wS_\bullet \mathsf{C}$, of weak equivalences. The *K-theory spectrum of* C is defined to be the Ω-spectrum whose *nth* space is given by

$$\mathbb{K}(\mathsf{C})_n := |w \underbrace{S_\bullet S_\bullet \cdots S_\bullet}_{n \text{ iterates}} \mathsf{C}|,$$

i.e., the realization of the subcategory of weak equivalences of the n-fold (degreewise) application of the $S_\bullet$ construction.

Definition 5.25 ((Split) Exact Sequences)**.** *Let* A, E, *and* B *be Waldhausen categories. A sequence of exact functors*

$$\mathsf{A} \xrightarrow{i} \mathsf{E} \xrightarrow{f} \mathsf{B}$$

is exact *if*

1. *The composition $f \circ i$ is the zero map to* B*;*

2. *The functor i is fully faithful; and*

3. *The functor f restricts to an equivalence between* E/A *and* B.[2]

A sequence, as above, is split *if there exist exact functors*

$$\mathsf{A} \xleftarrow{j} \mathsf{E} \xleftarrow{g} \mathsf{B}$$

that are adjoint to i and f and such that the unit of the adjunction, $\mathrm{Id}_{\mathsf{A}} \Rightarrow j \circ i$*, and the counit of the adjunction,* $f \circ g \Rightarrow \mathrm{Id}_{\mathsf{B}}$*, are natural isomorphisms.*

[2]Here, E/A is the full subcategory of E on objects e such that for all $a \in \mathsf{A}$ the hom set $\mathsf{E}(i(a), e)$ is a point.

Definition 5.26 (Standard Split SES)**.** *A split SES of Waldhausen categories*

$$\begin{tikzcd} \mathsf{A} \arrow[r, "i"] & \mathsf{E} \arrow[l, bend left, "j"] \arrow[r, "f"] & \mathsf{B}, \arrow[l, bend left, "g"] \end{tikzcd}$$

is standard *if*

1. *For each $e \in \mathsf{E}$, the component of the counit, $(i \circ j)(e) \to e$, is a cofibration;*

2. *For each cofibration $e \hookrightarrow e'$ in E, the induced map $e \amalg_{(i \circ j)(e)} (i \circ j)(e') \to e'$ is a cofibration; and*

3. *If $a \to a' \to 0$ is a cofiber sequence in A, then the first map is an isomorphism.*

The following is one of the fundamental theorems of algebraic K-theory and is a critical tool for what follows.

Theorem 5.27 (Waldhausen Additivity)**.** *Let*

$$\begin{tikzcd} \mathsf{A} \arrow[r, "i"] & \mathsf{E} \arrow[l, bend left, "j"] \arrow[r, "f"] & \mathsf{B}, \arrow[l, bend left, "g"] \end{tikzcd}$$

be a standard split SES of Waldhausen categories. Then the functors i and g induce an equivalence of spectra

$$\mathbb{K}(i) \vee \mathbb{K}(g) \colon \mathbb{K}(\mathsf{A}) \vee \mathbb{K}(\mathsf{B}) \xrightarrow{\sim} \mathbb{K}(\mathsf{E}).$$

5.28 Persistence Modules, Persistence Cosheaves, and Filtrations

We now introduce our main actors: persistence modules and persistent cosheaves. We first consider constructible cosheaves that arise from common types of persistence modules/filtrations. We construct these cosheaves in a way that is compatible with traditional models of the specific module/filtration in question. We finish with an

equivalence result relating the category of d-parameter zig-zag grid modules to the category of constructible cosheaves over a stratified $\mathbb{R}^d$.

5.28.1 Persistent Definitions

Many of the following definitions were briefly given in Chapter 1; we restate and elaborate them here for clarity. Recall from Definition 1.2 that a *persistence module* is a functor $P : \mathbb{P} \to \Omega$, where $\mathbb{P}$ is a poset category. To make the connection to the usual conventions in persistent homology and TDA clearer, for the remainder of this chapter we might consider Ω to be $\mathsf{Vect}_{\mathbb{F}}$, the category of vector spaces over a field $\mathbb{F}$, and by $\mathbb{F}^i$, we mean a free vector space of dimension i in $\mathsf{Vect}_{\mathbb{F}}$. We will often omit the subscript $\mathbb{F}$ from the notation and simply write Vect.

The previous definition is general – in what follows, we will first restrict our attention to one-parameter persistence modules before shifting to multi-parameter modules. There are two flavors of such modules common in the literature: zig-zag [14] and monotone (standard) persistence modules [68] (see also Chapter 1). We note that monotone persistence modules are most commonly called simply 'persistence modules;' we have added the word 'monotone' to emphasize their distinction from more general modules.

In this chapter, we are particularly interested in zig-zag grid modules, which may be *multi-parameter* modules. We remind the reader of these distinctions. To any locally finite poset, there is an associated graph: its Hasse diagram (see Chapter 1, Section 1.1). Here, we consider the Hasse diagram of a poset as a one-dimensional simplicial complex.

Definition 5.29. *Let $\mathbb{P}$ be a poset.*

- *If* $\mathbb{P}$ *is the product of* d *(non-trivial) linear orders, then a representation* $P\colon \mathbb{P} \to \Omega$ *is a* d-parameter monotone grid persistence module.

- *The poset* $\mathbb{P}$ *is* d-parameter zig-zag grid *if its Hasse diagram is homeomorphic the Hasse diagram of the product of* d *(non-trivial) linear orders. A representation of a zig-zag grid poset* $P\colon \mathbb{P} \to \Omega$ *is a* d-parameter zig-zag grid persistence module.

Note that all monotone grid persistence modules are additionally zig-zag, but zig-zag modules are not generally monotone. Letting $d = 1$ in the above definitions, we recover traditional one-parameter modules. Generally, we may use the phrase *multi-parameter* for arbitrary d. Particularly in this section, we sometimes focus on zig-zag grid persistence modules whose underlying poset is specifically the product of one-parameter modules. We explicitly state when this is the case. In later sections (e.g., Section 5.51), we broaden our view to general zig-zag grid persistence modules.

In previous work [37] we viewed one-parameter zig-zag persistence modules as cosheaves over the parameter space $\mathbb{R}$. When shifting to d-parameter modules, we therefore consider the parameter space $\mathbb{R}^d$. We define the parameter space for a d-dimensional zig-zag grid persistence module as follows.

Definition 5.30 (Stratified d-Parameter Space and its Entrance Path Category)**.** *Let* I *be a zig-zag grid poset, whose Hasse diagram is homeomorphic to the Hasse diagram of the product poset* $I_1 \times I_2 \times \ldots \times I_d$*, where each* $I_n \subset \mathbb{R}$ *is a zig-zag poset with discrete elements.*

We then define the stratification

$$(\mathbb{R}^d; I) := \prod_{n=1}^{d} (\mathbb{R}, I_n)^{\wedge\!\wedge}.$$

We are often interested in the combinatorial entrance path category of $(\mathbb{R}^d; I)$ *(see Definition 5.20). We abuse notation and write* $\mathsf{Ent}_{\Delta}(\mathbb{R}^{\mathsf{d}}; \mathsf{I})$ *for this category.*

Remark 8. *We caution the reader and emphasize that* I *is distinct from the stratifying poset of* $(\mathbb{R}^d; I)$*. The stratifying poset is implied (if not explicitly notated) by the definition;* $(\mathbb{R}^d; I)$ *has the stratification/cubulation that* I *naturally endows on* $\mathbb{R}^d$*, where, e.g., zero-strata correspond to objects of* I *and one-strata correspond to the generating relations of* I*. This distinction is illustrated in the following example.*

Example 10 (Stratified Two-Parameter Space and its Stratifying Poset)**.** *In Figure 5.2, the poset* $I = [2]^2$ *defines a stratification* $(\mathbb{R}^2; I)$*. Note that the stratifying poset is distinct from* I*; while* I *has nine objects, the stratifying poset of* $(\mathbb{R}^2; I)$ *has 49 objects (corresponding to each cube of the cubulation).*

Our goal is to define a cosheaf on a stratified $\mathbb{R}^d$. We will be helped by the following two lemmas regarding minimum zero-strata.

Lemma 5.31 (Minimum Zero-Stratum)**.** *Let* $I_1, I_2, \ldots I_d$ *be one-parameter zig-zag posets with each* $I_n \subset \mathbb{R}$ *discrete, so that* $I := I_1 \times I_2 \times \ldots \times I_d$ *is a d-parameter poset. Equip* $\mathbb{R}^d$ *with the stratification* $(\mathbb{R}^d; I)$*. Then every i-stratum has a zero-stratumn as a face that is minimum with respect to the partial order* I*.*

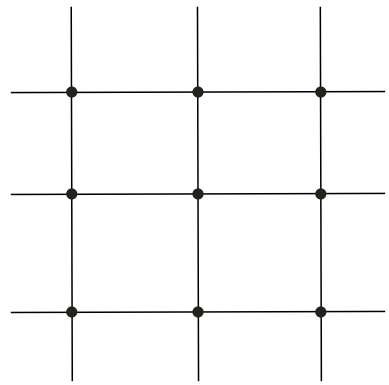


Figure 5.2: The poset $I = [2]^2$ defines a natural cubulation of $\mathbb{R}^2$, which then leads to a natural stratification on $\mathbb{R}^2$, denoted $(\mathbb{R}^2; I)$. Each strata of $(\mathbb{R}^2; I)$ is a cube of the associated cubulation.

Proof. When considered as a poset (sub)category, each finite i-stratum is the product of i intervals, i.e., is equivalent to $[1]^i$. The zero-stratum corresponding to $\{0\} \times \{0\} \times \ldots \times \{0\}$ is clearly the unique minimum. Infinite i-strata can be considered as equivalent to $[I]^i$, but where up to i terms in the product have been replaced with copies of the ray $[0, \infty)$ or $(-\infty, 0]$ (and these replacements are all of the first type or all of the second type). In either case, the zero-stratumn corresponding to $\{0\} \times \{0\} \times \ldots \times \{0\}$ is still the unique minimum zero-stratum of the stratification. $\square$

Lemma 5.32 (Mimimum Zero-Strata for Face/Cofaces are Comparable)**.** *Let $I_1, I_2, \ldots I_d$ be one-parameter zig-zag posets with each $I_n \subset \mathbb{R}$ discrete, so that $I := I_1 \times I_2 \times \ldots \times I_d$ is a d-parameter poset. Equip $\mathbb{R}^d$ with the stratification $(\mathbb{R}^d; I)$. For strata $\mathfrak{c}'$ and $\mathfrak{c}$ such that $\mathfrak{c}'$ is a face of $\mathfrak{c}$, the minimum zero-strata of $\mathfrak{c}$ is less than or equal to the minimum zero-strata of $\mathfrak{c}'$ with respect to the partial order I.*

Proof. We begin with the case of finite strata. Consider the finite strata $\mathfrak{c}$ as the product $[i]^i$ where $i \in [0, 1, \ldots, d]$ is the dimension of $\mathfrak{c}$, so that the minimum vertex of $\mathfrak{c}$ corresponds to $\{0\} \times \{0\} \times \ldots \times \{0\}$. Since $\mathfrak{c}'$ is a face of $\mathfrak{c}$, the minimum vertex of $\mathfrak{c}'$ corresponds to some point of $[I]^i$ as well. Since $\{0\} \times \{0\} \times \ldots \times \{0\}$ is minimal, it is comparable to (indeed, less than or equal to) every object of $[1]^i$, and the result follows.

The case of infinite strata is similar; if $\mathfrak{c}$ is an infinite i-strata, we can express it as the i-fold product of copies of $[1]$ and copies of either $[0, \infty)$ or $(-\infty, 0]$. The minimum vertex of $\mathfrak{c}$ still corresponds to the point $\{0\} \times \{0\} \times \ldots \times \{0\}$. A face $\mathfrak{c}'$ of $\mathfrak{c}$ may be either finite or infinite, but in either case, the minimum vertex of $\mathfrak{c}$ is less than or equal to the minimum vertex of $\mathfrak{c}'$ by similar reasoning. $\square$

Using the previous two lemmas as tools in our geometric and ordering arguments, we are ready to formally define persistence modules as cosheaves. By Lemma 5.14, to define a cosheaf on $\mathbb{R}^d$, it suffices to define a cosheaf on a basis of the topology on $\mathbb{R}^d$.

First, we give a cosheaf theoretic interpretation of the notion of multi-parameter zig-zag grid modules in the specific case that the underlying poset is the product of one-parameter zig-zag posets. We call this cosheaf *propagated* because the functor is entirely determined by the ordering of and assignments to zero-strata; the values over higher-dimensional strata are propagated from the minimum adjacent zero-stratum.

Construction 4 (The Propagated Persistence Cosheaf on $\mathbb{R}^d$)**.** *Given $P : I \to \Omega$ with $I = I_1 \times \ldots \times I_d$ where each $I_n \subset \mathbb{R}$ is discrete, we define the* the propagated persistence cosheaf $F_P \colon \mathsf{Opens}(\mathbb{R}^{\mathsf{d}}) \to \Omega$ *as follows. Let $B_\epsilon \subset \mathbb{R}^d$ be a metric ϵ-ball so that $2\epsilon > 0$ is less*

than the distance between any pair of zero-strata. By construction, the minimum-dimensional stratum that intersects B_ϵ is unique and, furthermore, by Lemma 5.31, (the closure of) each strata has a minimum zero-stratum with respect to the partial order I. For B_ϵ that intersects with a minimum-dimensional stratum whose minimum zero-stratum is $i \in I$, we assign

$$F_P(B_\epsilon) = P(i).$$

Next, we describe the assignment of morphisms, $F_P(B_\epsilon \hookrightarrow B'_\epsilon)$. If B_ϵ and B'_ϵ intersect the same minimum-dimensional stratum, then $F_P(B_\epsilon \hookrightarrow B'_\epsilon) = Id_{F(B_\epsilon)}$. Suppose instead that the minimum strata intersecting B'_ϵ is distinct from the minimum strata intersecting B_ϵ. Call these strata $\mathfrak{c}'$ and $\mathfrak{c}$, respectively. Suppose the minimum zero-strata of $\mathfrak{c}'$ is $i \in I$ and the minimum zero-strata of $\mathfrak{c}$ is $j \in J$. Since the epsilon balls were constructed to be smaller than any positive-dimensional strata and since $B_\epsilon \subseteq B'_\epsilon$, i, the minimum strata intersecting the two balls will have a face/coface relation, and thus, by Lemma 5.32, i is comparable to j. Then $F_P(B_\epsilon \hookrightarrow B'_\epsilon) = (P(i) \to P(j))$ if $i < j$ (or $(P(j) \to P(i))$ if $j < i$). Note that if $i = j$, this is again the identity. See Figure 5.3 for an illustration in the one-parameter case.

Despite its title, Construction 4 is an instance of an explicit precosheaf, and, so far, we have not given justification for its status as a cosheaf. The following remark discusses our reasoning for being somewhat cavalier with showing the cosheaf condition is satisfied.

Remark 9 (Cosheafification)**.** *Unlike its dual notation of* sheafification*, the existence of* cosheafification *(extending a precosheaf to a cosheaf) is subtle. However, cosheafification exists for precosheaves valued in nice-enough categories, such as* Vect *and* Set_**, our primary*

cases of interest. See e.g., [64] for the Set*-valued case, [19] for the* Vect*-valued case, and [50] for a broader discussion. We will generally assume our cosheaves are valued in categories for which cosheafification exists, and proceed defining cosheaves by specifying values on a basis.*

Finally, we observe that the cosheaf F_P is locally constant on strata, so it defines a constructible cosheaf on the stratified space $(\mathbb{R}^d; I)$.

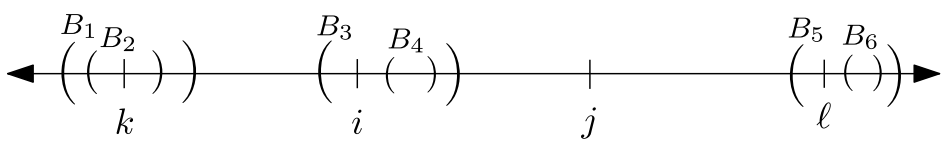


Figure 5.3: Examples of open intervals occurring in Construction 4

Example 11. *Suppose that I is the poset $k > i > j < l$, where $k, i, j,$ and l are ordered with the standard ordering on $\mathbb{R}$ as in Figure 5.3. Then $F_P(B_2 \hookrightarrow B_1) = Id_{F_P(B_2)} = Id_{P(k)}$, $F_P(B_4 \hookrightarrow B_3) = (P(j) \to P(i))$, and $F_P(B_6 \hookrightarrow B_5) = Id_{P(\ell)}$.*

5.32.1 Filtered Spaces and Cosheaves

Next, we discuss how persistence-modules and persistence module cosheaves relate to filtrations of spaces. General filtrations were defined in Chapter 1 (Definition 1.6), and lower-star filtrations were the focus of Chapters 3 and 4. Here, we revisit the definitions of various filtrations, as well as discuss the relationships between them.

5.32.1.1 Monotone and Index Filtrations

Here, we recall the definitions of monotone and index filtrations given in Chapter 1. Briefly, a *one-parameter monotone filtration* is a sequence of nested subcomplexes indexed by one parameter. If a d-parameter filtration looks like a one-parameter monotone filtration

when restricted to any parameter, we say it is a *d-parameter monotone filtration*. *Index filtrations* are a special type of monotone filtration where each inclusion map adds exactly one simplex. Given a one-parameter monotone filtration, $\{L_{m_j}\}_{m_j \in M}$ with filter function f, consider a total ordering of simplices, $\sigma_1, \sigma_2, \ldots, \sigma_n$, such that if either $f(\sigma_i) < f(\sigma_j)$ or $\sigma_i \prec \sigma_j$, then $i < j$. Then we say the index filtration $L_j := \{\sigma_i \mid i \leq j\}$ is *compatible with the monotone filtration*. Suppose we have a d-parameter monotone and a d-parameter index filtration. If, by restricting each filtration to a single parameter at a time, we can obtain a one-parameter index filtration that is compatible to the corresponding one-parameter monotone filtration, then we say d-parameter index filtration is compatible with the d-parameter monotone filtration.

Example 12 (Standard (One-Parameter Monotone) Persistence)**.** *If we take $I \subset \mathbb{R}$ to be the indexing set of a filtration, then there is a natural way to view I as a poset with the standard ordering of $\mathbb{R}$. Passing to homology in degree n defines an associated monotone persistence module via the assignment $i \mapsto H_n(L_i)$. The propagated persistence cosheaf on $\mathbb{R}$ (see Construction 4) is easy to describe. Indeed, for a single one-stratum, we have $F(i,j) = H_n(L_i)$ and that the costalk of F at a zero-stratum i is $H_n(L_i)$.*

Example 13 (Two-Parameter Monotone Persistence)**.** *The left side of Figure 5.5 illustrates a two-parameter monotone filtration, where $I = [2]^2$.*

5.32.1.2 From Index Persistence Modules to Monotone Persistence Modules

We begin by discussing the one-parameter case. Suppose that $\{L_{m_j}\}_{m_j \in M}$ is a discrete one-parameter monotone filtration with filter function f and $\{L'_i\}_{i \in [1,n]}$ is a compatible index

filtration. Then for every $m_j \in M \setminus \{m_0\}$, there is some maximum interval $[\ell, r)$ for $\ell, r \in \{1, 2, \ldots, n\} \cup \{\pm\infty\}$ such that $f(\sigma_\ell) = f(\sigma_r) = m_j$ (where, whenever $r = \infty$ or $r = -\infty$, we define $\sigma_\infty := \sigma_n$ and $\sigma_{-\infty} = \sigma_0$, respectively). These intervals cover $\mathbb{R}$, and every interval corresponds to a unique $m_j \in M$. We define a map of stratified spaces, $C : (\mathbb{R}; [1, \ldots, n]) \to (\mathbb{R}; M \setminus \{m_0\})$ that maps intervals with a particular value under the filter function f to intervals with that same value under f. We provide a precise definition below.

Definition 5.33. *Suppose that $a \in [\ell, r)$, where $[\ell, r)$ is the associated interval for some parameter value $m_j \in M$. Three cases arise: if $-\infty < \ell, r < \infty$, we assign*

$$C(a) = \begin{cases} m_j, \ \text{if } a < r - 1 \\ \\ m_j(r - a) + m_{j+1}(a - (r - 1)), \ \text{if } a \geq r - 1 \end{cases} . \tag{5.1}$$

if $[\ell, r) = [-\infty, 1)$, we assign

$$C(a) = am_1 \tag{5.2}$$

and if $[\ell, r) = [n, \infty)$, we assign

$$C(a) = \frac{am_p}{n}. \tag{5.3}$$

Lemma 5.34. *C is a stratified map.*

We could similarly define a map from d-parameter index filtrations to their compatible

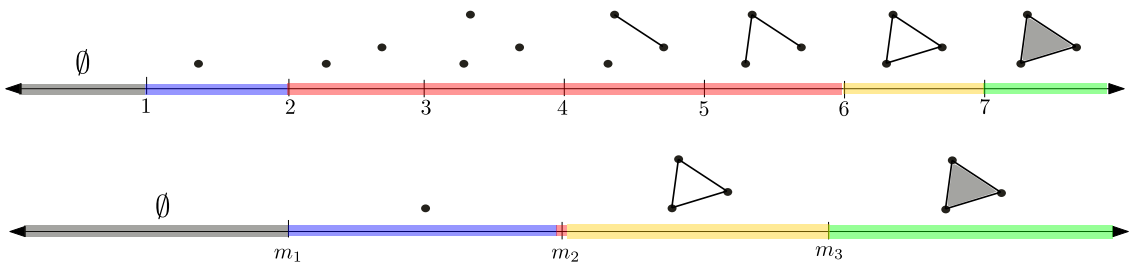


Figure 5.4: An example of the map C. The relevant interval for the point, e.g., m_2 is $[2, 7)$, since the image of each simplex added in that interval under the filter function f is m_2. Then $[2, 6)$ is mapped to m_2 and $[6, 7)$ is mapped to $[m_2, m_3)$.

d-parameter monotone filtrations. This could be done by performing a version of the map C described in Definition 5.33 for one parameter at a time, colapsing the hyper-rectangular fibers over intervals to $(d-1)$-planes. We conjecture that such a map is independent of the order in which we collapse parameters, but leave a proof of this conjecture to future work.

5.34.0.1 One-Parameter Augmented Descriptors via Index Filtrations

In this subsection, we shift our focus to the one-parameter setting; all filtrations considered are one-parameter filtrations. The d-parameter case is more subtle, and is discussed in the next section (Section 5.36.0.1).

Let $\{L_{m_j}\}_{m_j \in M}$ and $\{L'_i\}_{i \in [0,n]}$ be a discrete monotone filtration and compatible index filtration, respectively (as in the previous section). Given a monotone filtration, we are perhaps interested in the so-called *instantaneous events* that are captured in augmented topological descriptors, a remnant of the fact that many standard algorithms to produce descriptors for monotone filtrations are often actually employing compatible index filtrations. For example, an instantaneous n-dimensional homology event at time m_j records an n-boundary that was not mapped from a boundary or cycle in the inclusion $L_{m_{j-1}} \hookrightarrow L_{m_j}$.

Note that many applications of TDA, such as the classic application of manifold learning through a Vietoris-Rips filtration, discard events with a short lifespan because they may be attributed to noise, so non-augmented persistence diagrams are the traditional tool of choice (see [16]). However, recent developments in areas such as shape comparison and inverse TDA problems (see, e.g., [8], Chapter 3, Chapter 4) rely on instantaneous events for efficient representation of simplicial or cubical complexes, particularly when the filtration used is directional (e.g., height filtration, lower-star filtration, etc.)

We aim to track both instantaneous and non-instantaneous events at every step of a monotone filtration. We introduce A_n to account for instantaneous events (the extra data of an augmented module). Let β_{m_j} denote the free group on n-dimensional boundaries of L_{m_j} and let $\kappa_{m_{j-1}}$ denote the kernel of the map on n-dimensional homology induced by the inclusion $L_{m_{j-1}} \hookrightarrow L_{m_j}$. Since $\kappa_{m_{j-1}}$ corresponds to all cycles of $L_{m_{j-1}}$ that become boundaries in L_{m_j}, i.e., all cycles of $L_{m_{j-1}}$ that map to elements of β_{m_j}, the subgroup $\kappa_{m_{j-1}}$ can naturally be identified with a subgroup of β_{m_j}. Furthermore, since boundaries of $L_{m_{j-1}}$ are mapped injectively to boundaries of L_{m_j}, the subgroup $\beta_{m_{j-1}}$ is also naturally identified with a subgroup of β_{m_j}. Then, we define:

$$A_n(L_{m_j}) = \beta_{m_j}/(\beta_{m_{j-1}} + \kappa_{m_{j-1}}). \tag{5.4}$$

Note that, since $A_n(L_{m_j})$ is a quotient of free groups, and since the generators of $\beta_{m_{j-1}}$ and $\kappa_{m_{j-1}}$ are subsets of the generators of β_{m_j}, $A_n(L_{m_j})$ is free. It may be helpful to think of $A_n(L_{m_j})$ as the free group on n-boundaries of L_{m_j} that are not the images of boundaries or

cycles in $L_{m_{j-1}}$. An instantaneous event in a monotone filtration is the appearence of an n-boundary that was not a boundary or cycle in the previous step of the filtration, meaning the rank of $A_n(L_{m_j})$ is the number of points (counting multiplicity) on the diagonal (m_j, m_j) in the corresponding standard n-dimensional augmented persistence diagram. We can also view A_n as a repackaging of the "entire" information in index filtrations, independent of the choice of compatible index filtration. The connection between the A_n's of a monotone filtration and the data of compatible index filtrations is made explicit in the following lemma.

Lemma 5.35. *Suppose that $\{L_{m_j}\}_{m_j \in M}$ is a discrete monotone filtration corresponding to a filter function f and $\{L'_i\}_{i\in[0,n]}$ is any compatible index filtration. Let κ'_i denote the kernel of the map induced on homology in degree n by the inclusion $L'_i \hookrightarrow L'_{i+1}$. Furthermore, let $\kappa'_{\hookrightarrow}$ denote the kernel of the map induced on homology in degree n by the composition of inclusions $L'_{(\min C^{-1}(m_j))-1} \hookrightarrow \ldots \hookrightarrow L'_{\max C^{-1}(m_j)}$, where C is as in Definition 5.33 and Figure 5.4. Then:*

$$A_n(L_{m_j}) \cong \left(\bigoplus_{i=(\min C^{-1}(m_j))-1}^{\max C^{-1}(m_j)} \kappa'_i \right) / \kappa'_{\hookrightarrow}. \tag{5.5}$$

Proof. Recall that A_n as defined in Equation 5.4 is a free group, so we proceed by first showing the right side of Equation 5.5 is also a free group, and next, showing we have the desired isomorphism through a counting argument.

We observe that generators of $\kappa'_{\hookrightarrow}$ correspond to cycles of $L'_{(\min C^{-1}(m_j))-1}$ that become boundaries somewhere along the composition of inclusions. Consider such a cycle and suppose that L'_i is the last subcomplex in the filtraton where this cycle is still not a boundary. Then this cycle is naturally identified with a generator of κ'_i, since the cycle becomes a

boundary in the inclusion $L'_i \hookrightarrow L'_{i+1}$. This is true for each generator of $\kappa'_{\hookrightarrow}$, so we may view $\kappa'_{\hookrightarrow}$ as a subgroup of the sum in Equation 5.5. Since the right side of Equation 5.5 is a quotient of free groups, where the generators of $\kappa'_{\hookrightarrow}$ are a subset of generators of the sum, the right side of the equation is a free group.

Since both the left and right side of Equation (5.5) are free groups, it suffices to proceed by showing the left and right side have equal rank. Each step in an index filtration adds a single simplex, so either $\kappa'_i \cong \mathbb{F}^0$ (if the simplex added does not fill in any n-cycle) or $\kappa'_i \cong \mathbb{F}^1$ (if the simplex added in $L'_i \hookrightarrow L'_{i+1}$ fills in an n-cycle). Thus, the direct sum in the equation above has nontrivial terms only for values of i such that $L'_i \hookrightarrow L'_{i+1}$ witnesses the death of n-cycles in the index filtration. Recall that $[\min C^{-1}(m_j), \max(C^{-1}(m_j)) + 1)$ is the maximum interval whose image under the filter f is m_j. This means that, shifting to the left, we can identify $L'_{(\min C^{-1}(m_j))-1} = L_{m_{j-1}}$ and $L'_{\max C^{-1}(m_j)} = L_{m_j}$. Thus, every boundary of L_{m_j} that was not present as a boundary in $L_{m_{j-1}}$ is introduced or becomes a boundary in some step of the index filtration between the values $(\min C^{-1}(m_j)) - 1$ and $\max C^{-1}(m_j)$, which means terms of the direct sum above are nontrivial only when boundaries are created. This is exactly the count of boundaries introduced in the inclusion $L_{m_{j-1}} \hookrightarrow L_{m_j}$, i.e., it is $\beta_{m_j}/\beta_{m_{j-1}}$, using the notation previously introduced in the paragraph above and in Equation (5.4). However, recall that $A_n(L_{m_j})$ does not account for boundaries that fill in a cycle from a previous step in the filtration. Thus, we quotient out by the kernel of the composition of maps between $\min C^{-1}(m_j)$ and $\max(C^{-1}(m_j)) + 1$. This kernel is generated by boundaries and cycles of $L_{m_{j-1}}$ that are mapped to boundaries in L_{m_j}. Since $L'_{(\min C^{-1}(m_j))-1} = L_{m_{j-1}}$ and $L'_{\max C^{-1}(m_j)} = L_{m_j}$, and since the index filtration is compatible with the monotone

filtration, we see that $\kappa'_{\hookrightarrow} \cong \kappa_{m_{j-1}}$. Thus, the rank of the right side of Equation 5.5 is exactly the rank of the A_n as defined in Equation 5.4, and since we established both are free groups, we have shown the desired isomorphism. $\square$

Example 14. *Suppose that $\{L_{m_i}\}$ and $\{L'_i\}$ are monotone and index filtrations as in the bottom and top of Figure 5.4. Then $A_0(L_{m_2}) \cong \beta_{m_2}/(\beta_{m_1} + \kappa_{m_1}) \cong \mathbb{F}^2/\mathbb{F}^0 \cong \mathbb{F}^2$. Computed using the identification of Lemma 5.35, we see that, indeed, the values agree, since $\bigoplus_{i=1}^{6} \kappa'_i/\kappa'_{\hookrightarrow} \cong (\mathbb{F}^0 \oplus \mathbb{F}^1 \oplus \mathbb{F}^1/\mathbb{F}_0) \cong \mathbb{F}^2$.*

Finally, we are able to define the following cosheaf, which organizes the information of both instantaneous and non-instantaneous events.

Definition 5.36 (Augmented Filtration Cosheaf on $\mathbb{R}$)**.** *Let $\{L_{m_j}\}_{m_j \in M}$ be a discrete monotone filtration of a simplicial complex K, and suppose $\mathbb{R}$ is equipped with the stratification $(\mathbb{R}; M \setminus \{m_0\})$. We define the* augmented filtration cosheaf on $\mathbb{R}$, *$F_A : \mathsf{Opens}(\mathbb{R}) \to \mathsf{Vect}$, on metric ϵ-balls as follows.*

$$F_A(U) = \begin{cases} H_n(L_{m_j}) \oplus A_n(L_{m_j}), \ \textit{if } m_{j-1} \in U \\ H_n(L_{m_j}) \ \textit{if } U \subset (m_j, m_{j+1}) \ \textit{or } U \subset (m_j = m_p, \infty) \\ H_n(L_{m_1}) \ \textit{if } U \subset (-\infty, m_1) \end{cases} .$$

Observe that the above definition implies that the costalk at a zero-stratum m_{j-1} of $(\mathbb{R}; M \setminus \{m_0\})$ is $H_n(L_{m_{j-1}}) \oplus A_n(L_{m_j})$.

Remark 10. *For an index filtration $\{L'_i\}_{i\in I}$, any new n-cycles introduced through the map $L'_{i-1} \hookrightarrow L'_i$ are not n-boundaries, since the boundaries and interiors of simplices are added at distinct filtration events. Thus, $A_n(L'_i)$ is trivial, so the augmented filtration cosheaf corresponding to an index filtration is equivalent to its non-augmented filtration cosheaf.*

An instance of the previous remark is illustrated by following example.

Example 15. *Let $\{L'_i\}$ be the index filtration in the top of Figure 5.4. Notice that the one-dimensional costalk of the non-augmented filtration cosheaf at 7 is $H_1(L'_7) \cong \mathbb{F}^0$, which is isomorphic to $H_1(L'_7) \oplus A_1(L'_7) \cong \mathbb{F}^0 \oplus \mathbb{F}^1/\mathbb{F}^1 \cong \mathbb{F}^0$.*

The stratified map C define above provides a clean interpolation between the augmented, non-augmented, and index cosheaves associated to a filtration.

Proposition 1. *Let $\mathcal{F}_M$ and $\mathcal{F}_A$ be the non-augmented and augmented filtration cosheaves for some one-parameter monotone filtration $\{L_{m_j}\}_{m_j\in M}$ and let $\mathcal{F}_I$ be the filtration cosheaf for a compatible index filtration $\{L_i\}_{i\in[0,n]}$. Let $C : (\mathbb{R}; [1, \ldots, n]) \to (\mathbb{R}; M \setminus \{m_0\})$ be the map of stratified spaces as above. Then,*

(i) We have an isomorphism of cosheaves $C_*\mathcal{F}_I \cong \mathcal{F}_M$;

(ii) Let $U \subset \mathbb{R}$ be open such that $U \cap M = \emptyset$, then $\mathcal{F}_M(U) \cong \mathcal{F}_A(U)$.

Proof. That $\mathcal{F}_M$ and $\mathcal{F}_A$ agree on one-strata follows directly from their definitions (they can differ at zero-strata). In claim (i), there are two parts: that $C_*\mathcal{F}_I$ is constructible and that $C_*\mathcal{F}_I$ is isomorphic to $\mathcal{F}_M$. To prove the first, note that C is a composition of

"elementary collapses" as described in Example 8, so by functoriality $C_*\mathcal{F}_I$ is constructible. The second part of (i) is an explicit unwinding of the definition of the pushforward. $\square$

5.36.0.1 Multi-Parameter Augmented Descriptors

Given that definitions of monotone filtrations and their compatible index filtrations can be extended to multi-parameter filtrations, one might hope to find a parallel story to Section 5.34.0.1 in the multi-parameter setting. Defining an augmented multi-parameter filtration cosheaf is more nuanced, since an instantaneous event in one parameter may not be instantaneous in another; see Figure 5.5.

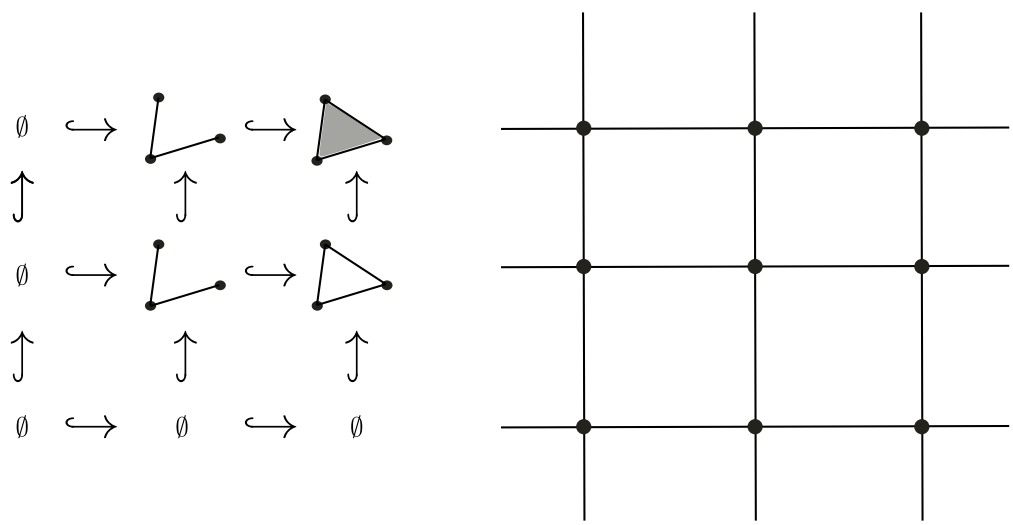


Figure 5.5: A bifiltration (left) illustrating one nuance when defining an augmented filtration cosheaf over stratified $\mathbb{R}^2$ (right). In final event of the filtration (at $(2,2)$), we see the instantaneous birth and death of a cycle coming from the horizontal inclusion map, but no such instantaneous event coming from the vertical map. There is therefore no natural assignment of a single vector space to the zero-strata at $(2,2)$ that captures this behavior.

There are many possible ways to handle the issue of events being instantaneous in some parameters but not others. We outline the following solution that extends a Vect-valued

persistence module to one valued in grVect with bounded support, where, for $1 \leq i \leq d$, the value assgned to the ith vector space is the value assigned to the one-parameter submodule arising from resricting the filtration to parameter i as in Definition 5.36 (and the values assigned to other levels in the grading are zero). Let $L_{m_i^1,\ldots,m_j^\ell,\ldots,m_k^d}$ be the simplicial complex corresponding to filtration index $(m_i^1,\ldots,m_j^\ell,\ldots m_k^d) \in I$, where $I \subset \mathbb{N}^d$ is the indexing set of the filtration. Define $A_n^\ell(L_{m_i^1,\ldots,m_j^\ell,\ldots m_k^d})$ as in Equation (5.4), but using the n-boundaries and n-cycles of the map

$$L_{m_i^1,\ldots,m_{j-1}^\ell,\ldots m_k^d} \hookrightarrow L_{m_i^1,\ldots,m_j^\ell,\ldots m_k^d}.$$

That is, the map into $L_{m_i^1,\ldots,m_j^\ell,\ldots m_k^d}$ in the ℓth direction. Then, our augmented filtration cosheaf is defined similarly to Definition 5.36, but using A_n^ℓ in the ℓth level of the image, for $\ell \in [1,\ldots,d]$. For open sets that do not contain a zero-strata, we propose a propogation similar to the one described in Construction 4, with the same values for each grade.

Example 16. *Consider the bifiltration shown in Figure 5.5. Denote $\mathbb{F}^i[n]$ to mean the graded vector space containing $\mathbb{F}^i$ in degree n and the zero vector space elsewhere. The value of the augmented filtration cosheaf for an ϵ-ball containing the point $(2,2)$ when considering first homology is $\mathbb{F}^1[1]$, since inclusion map into $L_{2,2}$ in the first parameter (e_1) direction sees an instantaneous one-cycle birth and death, while the inclusion map in the second parameter direction (e_2) direction does not.*

An ϵ-ball containing the vertical one-strata $\{(2,1),(2,2)\}$ has value $\mathbb{F}^1[1] \oplus \mathbb{F}^1[2]$ (since both parameters see a one-cycle for this parameter range) and an ϵ-ball containing the horizontal one-strata $\{(1,2),(2,2)\}$ gets assigned the zero vector space in all levels (since

no parameters see a one-cycle for this parameter range).

5.37 An Equivalence Result

In this subsection we make explicit the relationship between zig-zag modules as put forth by Carlsson–Zamorodian and representations of the entrance path category of $(\mathbb{R}^d; \mathbb{N}^d)$. In the process we will need to equip our zig-zag grid modules with additional structure, which we call "markings."

Let Poset_I be the category of posets with Hasse diagrams homeomorphic to Hasse diagrams of products of d-linear orders, and whose underlying sets are at most countable. Morphisms in Poset_I are surjective maps of posets. So from above, the category of multi-parameter zig-zag modules is the category of pairs $(\mathcal{P}, \rho)$ with $\mathcal{P} \in \mathsf{Poset}_I$ and $\rho\colon \mathcal{P} \to \mathsf{Vect}$ a representation of $\mathcal{P}$.

Definition 5.38. *Define the poset $\mathcal{ZZ}_{\mathbb{N}}$ to have objects $\frac{1}{2}\mathbb{N}$ with non-identity morphisms*

$$\frac{a}{2} \leq \frac{a+1}{2} \quad \textit{and} \quad \frac{a}{2} \leq \frac{a-1}{2} \quad \textit{for all } a \in \mathbb{N}, \textit{ with } a \textit{ odd.}$$

Then, we may write $\mathcal{ZZ}_{\mathbb{N}}^{\mathrm{d}} = \prod_1^d \mathcal{ZZ}_{\mathbb{N}}$, and we equip $\mathcal{ZZ}_{\mathbb{N}}^{\mathrm{d}}$ with the product order.

Notice that $\mathcal{ZZ}_{\mathbb{N}}^{\mathrm{d}}$ is the stratifying poset of $(\mathbb{R}^d; \mathbb{N}^d)$. (See Remark 8).

Lemma 5.39. *There is a canonical isomorphism of categories*

$$\mathcal{ZZ}_{\mathbb{N}}^{\mathrm{d}} \cong \mathsf{Ent}_{\Delta}(\mathbb{R}^d; \mathbb{N}^d).$$

We wish to "mark" our posets by passing to the *under category* of $\mathcal{ZZ}_{\mathbb{N}}^{\mathrm{d}}$, $\mathsf{Poset}_I^{\mathcal{ZZ}_{\mathbb{N}}^{\mathrm{d}}/}$,

i.e., we will consider posets equipped a map from $\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}}$. Passing to marked objects/the under category has the effect of replacing a given poset by all possible labelings of that poset by the d-fold product of natural numbers. (The notion of marking persistence modules is not at all unusual. For instance, in most applications, the passage from one-parameter persistence modules to barcodes or diagrams depends on an explicit marking, e.g., the event times/parameters.) Morphisms in the under category are commutative triangles. As we will use later, passage to the under category introduces an initial object: $\mathrm{Id} : \mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \to \mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}}$. The under category is an example of a comma category and are also known as *coslice* categories.

Definition 5.40. *Define the category of* marked zig-zag grid modules*,* ZZmod^d*, to be the category of pairs* $(\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P}, \rho)$ *with* $\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P} \in \mathsf{Poset}_I^{\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}}/}$ *and* $\rho\colon \mathcal{P} \to \mathsf{Vect}$ *a representation. A morphism is a pair* $(f, \varphi)\colon (\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P}, \rho) \to (\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{Q}, \eta)$ *with* $f\colon \mathcal{P} \twoheadrightarrow \mathcal{Q}$ *defining a morphism in the under category and* $\varphi\colon \rho \Rightarrow f^*\eta$ *a natural transformation.*

It turns out that isomorphism in ZZmod^d is too strong to capture our preferred notion of "sameness," so we introduce a notion of weak equivalence. An example of an operation that creates a weakly equivalent module is "subdividing" a row of vertices in a product poset (corresponding to restricting one factor of the product to a single vertex) into several rows of vertices where all of the new maps in the corresponding representation are isomorphisms.

Definition 5.41 (Weak Equivalences)**.** *A morphism* $(f, \varphi)\colon (\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P}, \rho) \to (\mathcal{ZZ}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{Q}, \eta)$ *in* ZZmod^d *is a* weak equivalence *if* $\varphi\colon \rho \Rightarrow f^*\eta$ *is a natural isomorphism. Let* $\mathcal{W}$ *denote the collection of weak equivalences.*

We caution the data-analytically oriented reader here; notice that weakly equivalent objects of ZZmod^d do not generally have the same indices of "events," i.e., vertices at which the corresponding image of the representation changes. For example, in the one-parameter case, the standard map from ZZmod^d to persistence diagrams (as described in [14]) does not factor through $\mathsf{ZZmod}^d[\mathcal{W}^{-1}]$. However, the order and number of events is preserved.

Theorem 5.42. *The category of (marked) multi-parameter zig-zag modules localized at weak equivalences is equivalent to the category of constructible cosheaves on $\mathbb{R}^d$ stratified by the natural numbers. That is, we have an equivalence of categories*

$$\mathsf{ZZmod}[\mathcal{W}^{-1}] \cong \mathsf{Fun}(\mathsf{Ent}_\Delta(\mathbb{R}^d;\mathbb{N}^d), \mathsf{Vect}) \simeq \mathsf{cShv}^{\mathsf{Vect}}_{\mathsf{cbl}}((\mathbb{R}^{\mathsf{d}};\mathbb{N}^{\mathsf{d}})).$$

The second equivalence is just an example of the exodromy equivalence. The first equivalence, which is actually an isomorphism of categories, is proved in the next section (Section 5.42.1). There are some technicalities in proving the previous theorem, but the main idea of the equivalence is as follows. Let $\varphi\colon \mathcal{P} \to \mathsf{Vect}$ be a representation of $\mathcal{P}$. Pullback φ along the map $\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \to \mathcal{P}$ to obtain a representation of $\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}$ so that, via Lemma 5.39, we have a representation of the corresponding entrance path category, i.e., a constructible cosheaf.

5.42.1 Proof of Theorem 5.42

The key ideas we use in the proof of the theorem go back to Grothendieck (and Verdier), specifically SGA4 Exposé VI [1]. The key observation—which we make precise—is that the category of zig–zag persistence modules is a localization of the *Grothendieck construction* on the (psuedo)functor that sends a poset to its category of representations:

$\mathcal{R}\colon \mathsf{Poset}^{op} \to \mathsf{Cat}$. As we will explain, the Grothendieck construction is the lax colimit of $\mathcal{R}$. Our domain category has an initial object, $\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}$, hence, the colimit of $\mathcal{R}$ is isomorphic to the evaluation $\mathcal{R}(\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}})$. Finally, in Lemma 5.39 we recognized $\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}$ as the poset underlying the (combinatorial) entrance path category of $(\mathbb{R}^d; I)$.

Throughout this section we will work with bicategories. Any category can be considered as a bicategory with the only 2-morphisms being identities. The bicategory of categories, Cat, consists of (small) categories, functors, and natural transformations. A reader who finds this section terse may find the recent book of Johnson and Yau [40] of great use.

5.42.2 Psuedo and lax (co)limits

When going from 1-categories to 2-categories there is more flexibility in definitions. This is already apparent when considering that the notion of 2-functor extends to limits and colimits as well. Many details of (co)limits in 2-categories were explicated in the 1980's by Ross Street and collaborators, for instance [10]. As an orienting exercise, let us recall the definition of lax and pseudo functors.

Definition 5.43. *Let $\mathcal{A}$ and $\mathcal{B}$ be bicategories. A* lax functor $P\colon \mathcal{A} \to \mathcal{B}$ *consists of*

- *A function $P\colon \mathrm{Obj}(\mathcal{A}) \to \mathrm{Obj}(\mathcal{B})$;*
- *For each hom-category $\mathcal{A}(X,Y)$ in $\mathcal{A}$, a functor*

$$P_{X,Y}\colon \mathcal{A}(X,Y) \to \mathcal{B}(P_X, P_Y);$$

- *For each object $X \in \mathcal{A}$ a 2-cell $P_{\mathsf{id}_X}\colon \mathsf{id}_{P_X} \Rightarrow P_{X,X}(1_X)$;*

- *For each triple of objects and morphisms $f\colon X \to Y$ and $g\colon Y \to Z$, a natural (in f and g) transformation*

$$P_{f,g}\colon P_{Y,Z}(g) \circ P_{X,Y}(f) \Rightarrow P_{X,Z}(g \circ f).$$

A lax functor is a *psuedofunctor* if the 2-cells/natural transformations in the definition above are invertible. So a pseudofunctor is more strict than a lax functor, but not yet a strict 2-functor, which would require all higher morphisms to be identities. Correspondingly we have variable notions of colimit. For details see Chapter 5 of [40] and/or [10].

Definition 5.44. *Let $\Phi\colon \mathcal{A} \to \mathcal{B}$ be a lax functor.*

- *A* lax colimit *of Φ is an initial object in the category of lax cocones under Φ;*
- *A* psuedocolimit *of Φ is an initial object in the category of psuedococones under Φ.*

The lax colimit of Φ is unique up to equivalence, while the psuedocolimit is unique up to isomorphism. We will use the notation $\mathsf{colim}\,\Phi$ for "the" psuedocolimit of Φ.

Lemma 5.45. *Let $\Phi\colon \mathcal{A} \to \mathcal{B}$ be a lax functor, $\mathcal{A}$ an honest 1-category and $\mathfrak{T} \in \mathcal{A}$ a terminal object. Then, we have an isomorphism $\mathsf{colim}\,\Phi \cong \Phi(\mathfrak{T})$. Correspondingly, if $\Psi\colon \mathcal{A}^{op} \to \mathcal{B}$ is a lax functor, $\mathcal{A}$ an honest 1-category and $\mathfrak{I} \in \mathcal{A}$ is initial, then $\mathsf{colim}\,\Psi \cong \Psi(\mathfrak{I})$.*

5.45.1 The Grothendieck Construction

Definition 5.46. *Let C a category and $\Phi\colon \mathsf{C}^{op} \to \mathsf{Cat}$ be a lax functor. The* Grothendieck construction, *$\int \Phi$, is the following category:*

- *An object of* $\int \Phi$ *is a pair,* (A, X)*, with* $A \in \mathsf{C}$ *and* $X \in \Phi(A)$*;*

- *A morphism* $(f, p)\colon (A, X) \to (B, Y)$ *consists of*

  - *A morphism* $f\colon A \to B$ *in the category* C*; and*

  - *A morphism* $p\colon X \to \Phi(f)(Y)$ *in* $\Phi(A)$*.*

There are (reasonably) clear composition and identities in $\int \Phi$ and it is standard to verify that $\int \Phi$ actually defines a category. The symbol "$\int$" is meant to convey that the Grothendieck construction is amalgamating the data of Φ "over" the domain category C. Indeed, projection defines a functor $\int \Phi \to \mathsf{C}$ that is a fibration.

Proposition 2 (Theorem 10.2.3 of [40])**.** *Let* C *a category and* $\Phi\colon \mathsf{C}^{op} \to \mathsf{Cat}$ *be a lax functor. The Grothendieck construction,* $\int \Phi$*, is a lax colimit of* Φ*.*

Let $U\colon \mathsf{E} \to \mathsf{C}$ be a functor and $\varphi\colon e \to e'$ a morphism in E. Recall that φ is *cartesian* if every commutative triangle in C involving $U(\varphi)$ with a chosen lift of a 2-horn has a unique filler. (This recollection is a bit colloquial, see Section 9.1 of [40].)

Corollary 5.47. *After localizing* $\int \Phi$ *at the collection of cartesian morphisms (with respect to the projection* $\int \Phi \to \mathsf{C}$*) we obtain a pseudocolimit of* Φ*, i.e., if* Cart *denotes the class of cartesian morphisms in* $\int \Phi$*, then* $\int \Phi[\mathrm{Cart}^{-1}] \cong \mathsf{colim}\, \Phi$*.*

5.47.1 Proving the Theorem

Let $\mathcal{R}\colon \mathsf{Poset}_I^{\mathcal{Z}\mathcal{Z}_{\mathbb{N}}^{\mathrm{d}}/} \to \mathsf{Cat}$ be the psuedofunctor of linear representations, i.e.,

$$\mathcal{R}(\mathcal{Z}\mathcal{Z}_{\mathbb{N}}^{\mathrm{d}} \twoheadrightarrow \mathcal{P}) := \mathsf{Fun}(\mathcal{P}, \mathsf{Vect}).$$

By design, the Grothendieck construction of $\mathcal{R}$ recovers the category of marked zig-zag grid persistence modules.

Lemma 5.48. *For $\mathcal{R}$ defined above, $\int \mathcal{R} \cong \mathsf{ZZmod}^d$.*

Lemma 5.49. *A morphism $(f, \varphi)\colon (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P}, \rho) \to (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{Q}, \eta)$ in ZZmod^d is Cartesian if and only if φ is a natural isomorphism.*

Proof. Let $(f, \varphi)\colon (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P}, \rho) \to (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{Q}, \eta)$ be a morphism and $(g, \psi)\colon (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{R}, \alpha) \to (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{Q}, \eta)$ a morphism such that $h\colon (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{R}) \to (\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P})$ defines a commutative triangle in $\mathsf{Poset}_I^{\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}/}$. We need to find a (unique) natural transformation $\chi\colon \alpha \Rightarrow h^*\rho$ such that (h, χ) fills the 2-horn upstairs in ZZmod^d. This is possible precisely when $\varphi\colon \rho \Rightarrow \eta$ is an isomorphism. Indeed, $g^* = h^* \circ f^*$, and $\psi\colon \alpha \Rightarrow g^*\eta$, so if φ is an isomorphism we define

$$\chi := \psi\colon \alpha \Rightarrow g^*\eta \cong h^*(\phi^*\eta) \cong h^*\rho.$$

□

Lemma 5.50. *The object $(\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \xrightarrow{\mathrm{Id}} \mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}})$ is initial in $\mathsf{Poset}_I^{\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}/}$.*

Proof. Let $(\varphi\colon\ \mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \twoheadrightarrow \mathcal{P}) \in \mathsf{Poset}_I^{\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}/}$. A map in the under category from $(\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \xrightarrow{\mathrm{Id}} \mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}})$ is a commutative diagram in Poset_I

$$\begin{tikzcd}
 & \mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \arrow[dl, "\mathrm{Id}"'] \arrow[dr, "\varphi"] & \\
\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \arrow[rr, "\psi"'] & & \mathcal{P}.
\end{tikzcd}$$

By commutativity of the triangle, the map $\psi = \varphi$, so there is indeed a unique map in the under category. □

The preceding lemmas assemble to a proof of Theorem 5.42. That is, we have shown

$$\mathsf{ZZmod}^d[\mathcal{W}^{-1}] \cong \int \mathcal{R}[\mathrm{Cart}^{-1}] \cong \mathsf{colim}\,\mathcal{R} \cong \mathcal{R}(\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}} \xrightarrow{\mathrm{Id}} \mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}) \cong \mathsf{Fun}(\mathcal{Z}\mathcal{Z}^{\mathrm{d}}_{\mathbb{N}}, \mathsf{Vect}).$$

5.51 K-theory of Multi-Parameter Modules

We now shift gears and compute the K-theory of the category of multi-parameter persistence modules. The category in which our modules take values plays a central role and we consider two different constructions: one for modules valued in vector spaces and one for set-valued modules. To begin, we work with an arbitrary cubical manifold as parameter space and only later specialize to the case where it is d-dimensional. Motivated by the Exodromy Theorem and our equivalence result above, we make the following definition.

Definition 5.52. *Let* $(X \xrightarrow{\phi} \mathcal{P})$ *be a cubical manifold with its native stratification,* $\mathsf{Ent}_\Delta(X, \mathcal{P})$ *its combinatorial entrance path category, and* Ω *any category. The* category of Ω-valued persistence modules parameterized by X, $\mathsf{pMod}^\Omega(X)$, *is given by*

$$\mathsf{pMod}^\Omega(X) := \mathsf{Fun}(\mathsf{Ent}_\Delta(X, \mathcal{P}), V).$$

Hence, the K-theory of Ω-valued persistence modules (parametrized by X) *is the* K*-theory spectrum (whenever it exists) of the category above:* $\mathbb{K}(\mathsf{pMod}^\Omega(X))$.

5.52.1 K-Theory of Vect-Valued coSheaves

The category of finitely generated modules for a commutative ring is an Abelian category, so we define/compute K-theory using the work of Quillen and Waldhausen. (If our ring is a field, we recover our old friend Vect).

Lemma 5.53. *Let R be a commutative ring, $\mathcal{M}$ the associated Waldhausen category of finitely generated modules. Furthermore, let X be a cubical manifold and let A denote a closed sub-stratified space of X. Then the following sequence is split short exact sequence of Waldhausen categories*

$$\begin{tikzcd} \mathsf{Fun}(\mathsf{Ent}_\Delta(X \setminus A), \mathcal{M}) \arrow[r, "j_*"] & \mathsf{Fun}(\mathsf{Ent}_\Delta(X), \mathcal{M}) \arrow[r, "i^*"] \arrow[l, bend left, "j^*"] & \mathsf{Fun}(\mathsf{Ent}_\Delta(A), \mathcal{M}), \arrow[l, bend left, "i_*"] \end{tikzcd}$$

where $i\colon A \hookrightarrow X$ and $j\colon X \setminus A \hookrightarrow X$ are the inclusion maps.

Proof. We first show that the sequence is a split short exact sequence. That is, it satisfies the conditions of Definition 5.25.

We begin with exactness of the sequence. Observe that the inverse and direct image functors (in this setting) are compatible with the equivalences and cofibrations, so indeed we have a sequence of exact functors. Furthermore, the composition $i^* \circ j_*$ is manifestly the zero functor. Next, since j_* is the extension by zero map, it is fully faithful. Finally, i^* presents $\mathsf{Fun}(\mathsf{Ent}_\Delta(A), \mathcal{M})$ as the cokernel of j_*.

Next, we show that the sequence splits. It is standard that i_* is right adjoint to i^* and in this case, the counit of the adjunction is a natural isomorphism. Because our domain categories are discrete (finite even), j_* is indeed left adjoint to j^* and the unit is a natural isomorphism; recall that j_* is the extension by zero map.

Finally, by Lemma 5.24, all categories in the sequence above are Waldhausen. Thus, the sequence is a split short exact sequence of Waldhausen categories, as desired. $\square$

Lemma 5.54. *The split short exact sequence of Lemma 5.53 is standard.*

Proof. We will show that the sequence of Lemma 5.53 satisfies Definition 5.26. Let $\mathcal{F} \in \mathsf{Fun}(\mathsf{Ent}_\Delta(X), \mathcal{M})$. Each component of the natural transformation $(j_* \circ j^*)(\mathcal{F}) \to \mathcal{F}$ is an isomorphism, except for the components corresponding to strata of A. The components corresponding to strata of A are inclusions of zero, which are cofibrations. In either case, compnents of $(j_* \circ j^*)(\mathcal{F}) \to \mathcal{F}$ are cofibrations, meaning $(j_* \circ j^*)(\mathcal{F}) \to \mathcal{F}$ is a cofibration in the functor category. Thus, Condition (1) of Definition 5.26 is satisfied.

Next, let $\varphi\colon \mathcal{F} \to \mathcal{F}'$ be a cofibration in $\mathsf{Fun}(\mathsf{Ent}_\Delta(X), \mathcal{M})$. We need to check that unique map

$$\psi\colon \mathcal{F} \coprod_{(j_*\circ j^*)(\mathcal{F})} (j_* \circ j^*)(\mathcal{F}') \to \mathcal{F}'$$

is a cofibration. That is, we must check that ψ is a monomorphism for every component. For a strata $T \subseteq A$, the pushout is identified with $\mathcal{F}(T)$ and $\psi_T = \varphi_T$, so by hypothesis it is a monomorphism. For a strata $S \not\subseteq A$, we may consider the following commutative diagram, where the square is a pushout.

$$\begin{tikzcd}
\mathcal{F}(S) \arrow[r, "\varphi_S"] \arrow[d, "\mathrm{Id}"'] & \mathcal{F}'(S) \arrow[d] \arrow[ddr, bend left, "\mathrm{Id}"] & \\
\mathcal{F}(S) \arrow[r] \arrow[drr, bend right, "\varphi_S"'] & \mathcal{F}(S) \coprod_\varphi \mathcal{F}'(S) \arrow[dr, "\psi_S"] & \\
& & \mathcal{F}'(S).
\end{tikzcd}$$

Since φ_S is a monomorphism, so is ψ_S. Thus, we have shown ψ is a cofibration, so Condition (2) of Definition 5.26 is satisfied.

Finally, Condition (3) of Definition 5.26 holds for categories of modules (see Remark 2.18 of [32]) and by the same reasoning, our category of functors valued in $\mathcal{M}$. $\square$

Since the sequence in Lemma 5.53 is a split short exact sequence of Waldhausen categories, Waldhausen Additivity (Theorem 5.27) immediately gives us the following key corollary, which will be the main tool we use to in our argument to"break apart and glue together" modules from submodules.

Corollary 5.55 (K-Theory is Additive Over Sub-Modules)**.** *Let R be a commutative ring, $\mathcal{M}$ the associated Waldhausen category of finitely generated modules. Furthermore, let X be a cubical manifold and let A denote a closed sub-stratified space of X. Then we have an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X \setminus A)) \vee \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(A)).$$

The following is a main result of our work with one-parameter modules ([37, Lemma 4.1.5]), and serves as a base case for our induction into multi-parameter modules. We present the proof here as well, since the techniques extend to our induction.

Theorem 5.56 (K-Theory of One-Parameter Modules: Base Case)**.** *Let X be a cubical one-manifold with a finite set of strata. There is an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \bigvee_{x_0 \in X_0} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_0)) \vee \bigvee_{x_1 \in X_1} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_1)).$$

where X_i is the set of i-strata of X.

Proof. We proceed by induction over the number of zero-strata. As our base case, note that when there are no zero-strata, we have $X_0 = \emptyset$ and $x_1 = X_1 = X$, so the claim holds. Now suppose that the claim holds whenever X contains $n-1$ zero-strata, for all $n-1 \geq 0$. Then consider the case that X contains n zero-strata. For an arbitrary zero-stratum $x_0^* \in X_0$, we know by Corollary 5.55 that there is an equivalence of spectra

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X \setminus x_0^*)) \vee \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_0^*))$$

Since $X \setminus x_0^*$ is itself a one-dimensional cubical manifold with $n-1$ zero-strata, our inductive hypothesis allows us to write

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \left(\bigvee_{x_0 \neq x_0^* \in X_0} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_0)) \vee \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X_1)) \right) \vee \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_0^*)).$$

We reindex by absorbing the last term into the first and we have the desired result. □

Theorem 5.56 deals with cubical one-manifolds. For the multi-parameter setting, we restrict our attention to a specific type of cubical d-manifold that arises when considering zig-zag grid persistence modules.

Definition 5.57 (Cubical Grid Manifold)**.** *Let X be a cubical manifold, possibly with boundary. If X is a substratified space of $(\mathbb{R}^d; I)$ for some $I = I_1 \times I_2 \times \ldots \times I_d$ where each $I_n \subset \mathbb{R}$ is discrete, we say X is a* cubical grid d-manifold.

The above definition serves two purposes; firstly, it allows us to consider manifolds with boundary, which will be important in our inductive argument. Secondly, it restricts our

attention to subsets of $\mathbb{R}^d$ rather than to general d-manifolds.

In the base case (Theorem 5.56), we inducted on the number of zero-strata of X. We formally generalize this notion to the multi-parameter setting in the following definition.

Definition 5.58 (Height)**.** *Let $I = I_1 \times I_2 \times \ldots I_d$, where each $I_n \subset \mathbb{R}$ is a one-parameter zig-zag poset. We say the* height *of I is the maximum number of objects in any $I_n \subset I$. Given a cubical grid d-manifold X, we say the* height *of X is the minimum height over all posets I such that $X \subseteq (\mathbb{R}^d; I)$. If the height of X equals the number of objects in I_n, we say that the nth parameter* realizes the height of X.

Note that in the one-parameter case, the height of X is just the number of zero-strata of X. In the general d-parameter case, height is a measure of the longest axis-aligned "slice," although the height of X may be realized in more than one parameter.

We are now ready to state our main result, the K-theory of multi-parameter zig-zag grid modules. The proof will use a double induction on the number of parameters and on the height of the module.

Theorem 5.59 (K-Theory of Multi-Parameter Zig-Zag Modules)**.** *Let X be a cubical grid d-manifold with a finite number of strata. There is an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \bigvee_{x_0 \in X_0} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_0)) \vee \bigvee_{x_1 \in X_1} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_1)) \vee \ldots \vee \bigvee_{x_d \in X_d} \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(x_d))$$

where X_i is the set of i-strata of X.

Proof. We proceed by double induction; first on h, the height of the persistence module, and then on d, the number of parameters. As a base case, we observe that Theorem 5.56 asserts the statment holds for $d = 1$ and all h. Suppose that there exist $d_*, h_* \in \mathbb{N}$ so that the claim holds for all $1 \leq d \leq d_*$ and $1 < h \leq h_*$. [3]

Induction on Height: First, we induct on the height of X. Suppose that X is d_*-dimensional and has height $h_* + 1$. For simplicity, we first consider the case that only one parameter realizes the height $h_* + 1$. and suppose this parameter is *not* the ith parameter for some $i \in [1, d]$. Then, for some $m \in I_i$, consider the closed $(d_* - 1)$-parameter sub-stratified space A corresponding to the poset $I_1 \times \ldots \times I_{i-1} \times m \times I_{i+1} \times \ldots I_d$, so that, conceptually, A corresponds to a level-set of X at height m, slicing through the parameter that realizes the maximum height. Note that $X \setminus A$ is generally two disconnected d_*-dimensional spaces (or one d_*-dimensional space in the specific case that A is on the boundary of X), each with height no more than h_*. Furthermore, note that A has height no more than h_*. By the inductive hypothesis, the claim holds for the connected components of $X \setminus A$ as well as for A, so by Corollary 5.55, we see that the claim holds for all of X.

We describe the process of dividing X into connected components each with height no more than h_* algorithmically, using a stack data structure. Recall that, just like a stack of plates, a stack utilizes a "first on, first off" organization, where the element that was most recently pushed onto the stack is the element availible to be popped off (see [18] for further details). Specifically, we use a stack, T, of "tall" connected components of this division that

[3] Here, we make the inequality $h > 1$ strict to avoid tautologies; a cubical manifold with height one is simply a point.

have height $h_* + 1$, initialized to $T = X$. We keep track of a list S of "short" connected components that have height less than $h_* + 1$, initialized as empty. The procedure is as follows. Pop $X_i \in T$. Then we choose a closed $(\dim(X_i) - 1)$-parameter sub-stratified space, $A_i \subset X_i$, that is perpendicular to some parameter of X_i that realizes height $h_* + 1$. Note then that the connected components of $X_i \setminus A_i$ and A_i may still have height $h_* + 1$, but in one fewer parameter than X_i had height $h_* + 1$. We push the connected components of $X_i \setminus A_i$ and A_i that have height $h_* + 1$ back on the stack T and move any connected components of this division with height less than $h_* + 1$ to our list S. Since each processed element of T has height $h_* + 1$ in one fewer parameter direction than before it was processed, we eventually have T empty and S a division of X into cubical grid manifolds for modules each with height less than $h_* + 1$. Noting that each time a sub-stratified space was removed, it was a closed subspace of the cubical manifold containing it, we "glue" back the pieces using Corollary 5.55, eventually showing the K-theory for persistence modules over all of X is as claimed.

Induction on Number of Parameters: Next, we induct on the number of parameters. Proceeding with our first induction argument, we may consider parameters of any height and we let X' be a cubical grid $(d_* + 1)$-manifold with height two. This means X' is a $(d_* + 1)$-cube. Let $\mathfrak{c}$ denote the $(d_* + 1)$-stratum of this cube (i.e., the interior). Since $X' \setminus \mathfrak{c}$ is a closed substratified space of X', we know by Corollary 5.55 that we have an

equivalence of spectra

$$\mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X)) \cong \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(\mathfrak{c})) \vee \mathbb{K}(\mathsf{pMod}^{\mathcal{M}}(X \setminus \mathfrak{c})). \tag{5.6}$$

(Note the slight flip of roles and notation from Corollary 5.55; our closed subspace is $A = X' \setminus \mathfrak{c}$ and thus, $X' \setminus A = X' \setminus (X' \setminus \mathfrak{c}) = \mathfrak{c}$).

It remains, then, to compute the K-theory of persistence modules over $X' \setminus \mathfrak{c}$. The space $X' \setminus \mathfrak{c}$ has the geometric structure of the boundary of a $(d_* + 1)$-cube, or, equivalently, a cube with no interior. By [23, Theorem 3], every unfolding of a cube with empty interior along a connected collection of codimension-two faces will not self-overlap and will lie on grid points of space in one dimension lower (i.e., every *ridge unfolding* of a finite cube will produce a *net*)[4].

These unfoldings correspond to trees in the cube's dual. A path in this dual corresponds to an unfolding where codimension-two faces are only pairwise adjacent to one other. This pairwise adjacency is along a common codimension-three face, so that the collection of codimension-two faces corresponds to a submanifold (with boundary) of the cube. Choose some such unfolding of the $(d_* + 1)$-cube, denoting the corresponding closed and connected collection of codimension-two faces (i.e., $(d_* - 1)$-faces) by U. Since the unfolding forms a net, it embeds into and stratifies $\mathbb{R}^{d_*}$, so $(X' \setminus \mathfrak{c}) \setminus U$ may be no more than a cubical grid d_*-manifold. Since U corresponds to a path in the dual graph of the cube, U is a cubical

[4]In [23], the word "cube" is taken in the computational geometric sense, and is assumed to have empty interior.

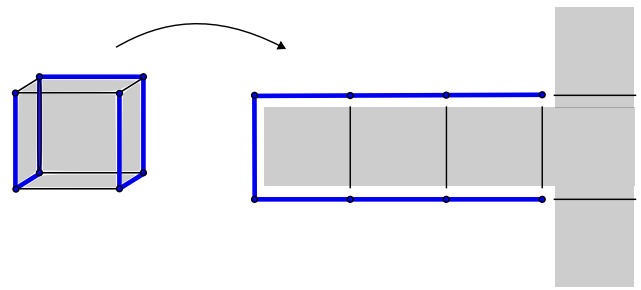


Figure 5.6: The unfolding of a three-dimensional cube C with height two in each parameter (an instance of the cube $X' \setminus \mathfrak{c}$ discussed in the induction on height in the proof of Theorem 5.59). The thick blue submodule U is a closed connected collection of codimension-two faces of the cube along which we can unfold (right). Since the unfolding (left) is a net, connected components of both the $C \setminus U$ and U cannot be more than two-parameter modules.

grid $(d_* - 2)$-manifold. (see Figure 5.6). Thus, by the inductive hypothesis, the claim holds for $(X' \setminus \mathfrak{c}) \setminus U$ as well as for U. Then by Corollary 5.55, we see that the claim holds for all of $X' \setminus \mathfrak{c}$. Combining this result with the equivalence in Equation (5.6), we have shown the desired result holds for all of X.

Finally, we observe that since we have shown that the claim holding for a cubical grid d-manifold implies the claim holds for a cubical $(d+1)$-manifold with height two, and since we have furthermore shown the claim holds for cubical grid manifolds of any height, we have shown the desired result in generality. $\square$

We end this section by discussing how Theorem 5.59 translates to the specific case of Vect-valued multi-parameter persistence modules.

Theorem 5.60 (K-theory of Vect-Valued Multi-Parameter Modules)**.** *Let X be a cubical grid d-manifold with a finite number of strata. There is an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathsf{Vect}_{\mathbb{F}}}(X)) \cong \bigvee_{X_0} \mathbb{K}(\mathbb{F}) \vee \bigvee_{X_1} \mathbb{K}(\mathbb{F}) \vee \ldots \vee \bigvee_{X_d} \mathbb{K}(\mathbb{F})$$

where X_i is the set of i-strata of X and $\mathbb{K}(\mathbb{F})$ denotes the K-theory spectrum of the field $\mathbb{F}$.

Proof. First, we identify the K-theory of components of the stratification, i.e., we identify $\mathbb{K}(\mathsf{pMod}^{\mathsf{Vect}}(x_i))$ for $x_i \in X_i$ and $i \in \{1, \ldots, d\}$.

By Definition 5.52, we have $\mathbb{K}(\mathsf{pMod}^{\mathsf{Vect}_{\mathbb{F}}}(x_i)) = \mathbb{K}(\mathsf{Fun}(\mathsf{Ent}_{\Delta}(x_i), \mathsf{Vect}_{\mathbb{F}}))$. Since $\mathsf{Ent}_{\Delta}(x_i)$ is the terminal category (a single object and an identity morphism), $\mathsf{Fun}(\mathsf{Ent}_{\Delta}(x_i), \mathsf{Vect}_{\mathbb{F}})$ is isomorphic to the category of $\mathsf{Vect}_{\mathbb{F}}$ itself. Thus, $\mathbb{K}(\mathsf{pMod}^{\mathsf{Vect}}_{\mathbb{F}}(x_i)) = \mathbb{K}(\mathsf{Vect})$. Now, the category of finite dimensional vector spaces over $\mathbb{F}$ is exactly the category of finitely generated projective modules over $\mathbb{F}$ (considered as a ring). Hence, $\mathbb{K}(\mathsf{Vect}_{\mathbb{F}})$ is just the algebraic K-theory of $\mathbb{F}$.

Thus, we have shown the K-theory of each strata is a copy of $K(\mathbb{F})$. We know by Theorem 5.59 that $\mathbb{K}(\mathsf{pMod}^{\mathsf{Vect}_{\mathbb{F}}}(X))$ is additive over strata, so the result follows. □

Remark 11. *An alternative approach to proving the preceding theorem could be to use a dévissage-type theorem, as in the "multiple type dévissage" of [51, Thm 6.11]. However,* Vect*-valued persistence modules do not satisfy the conditions on certain maps factoring through cofibrations that is required in [51]. We could correct this by shifting to modules valued in chain complexes (replacing a vector space V with the chain complex that has V in degree zero), but we leave this adaptation and proof technique to future work.*

The K-groups of an arbitrary field can be generally stated in the first two degrees. The following corollary combines this fact with Theorem 5.60.

Corollary 5.61. *For X, a cubical grid d-manifold with a finite number of strata, we have*

$$K_0(\mathsf{pMod}^{\mathsf{Vect}_{\mathbb{F}}}(X)) \cong \bigoplus_{X_0} \mathbb{Z} \oplus \bigoplus_{X_1} \mathbb{Z} \oplus \ldots \oplus \bigoplus_{X_d} \mathbb{Z}$$

and

$$K_1(\mathsf{pMod}^{\mathsf{Vect}_{\mathbb{F}}}(X)) \cong \bigoplus_{X_0} \mathbb{F}^\times \oplus \bigoplus_{X_1} \mathbb{F}^\times \oplus \ldots \oplus \bigoplus_{X_d} \mathbb{F}^\times$$

where X_i is the set of i-strata of X and $\mathbb{F}^\times$ is the group of units of $\mathbb{F}$.

The higher K-theory of fields contains interesting torsion and other phenomena. We refer the reader to Chapter IV of [63] for an in-depth description.

5.61.1 Pointed Set Valued coSheaves

While persistence modules are most often assumed to take values in vector spaces, there are interesting modules/cosheaves that take values in other categories. Of particular interest to us is the *Leray–Reeb* cosheaf, $\mathcal{L}_f$, associated to a map $f\colon Y \to X$, see [21]. Let us consider a simple situation: let $f\colon Y \to \mathbb{R}$ be a Morse function on a closed manifold Y. Now, given $U \subseteq \mathbb{R}$, let $\mathcal{L}_f(U) := \pi_0 f^{-1}(U)$. It is standard that the critical values of f stratify $\mathbb{R}$ and that $\mathcal{L}_f$ is constructible with respect to this stratification. So the Leray–Reeb cosheaf defines a persistence module taking values in the category of finite sets Set.

We may also consider *higher Morse functions*, $g : Y \to \mathbb{R}^d$, which also have an associated cosheaf. Just as in the one-dimensional case, given $U \subseteq \mathbb{R}^d$, let $\mathcal{J}_f(U) := \pi_0 f^{-1}(U)$. Given certain general position constraints (see [26]) the critical values of g stratify $\mathbb{R}^d$ and that $\mathcal{J}_f$ is constructible with respect to this stratification. The stratifications arising from higher-

dimensional Morse functions is a main object of study in [36]. Again, the cosheaf defines a persistence module taking values in the category Set.

For technical convenience we prefer our sets to be pointed/based (see Remark 12). Let us consider Set_*, the category of finite pointed sets and base point preserving functions. The category Set_* is Waldhausen (cofibrations are injections and weak equivalences are bijections), so given a cubical manifold $(X \xrightarrow{\phi} \mathcal{P})$ we can compute the K-theory of the associated (Waldhausen) category of persistence modules $\mathsf{pMod}^{\mathsf{Set}_*}(X)$.

Note that the proof Lemma 5.53 goes through for Set_* valued functors *mutatis mutandis*. The following version of Lemma 5.54 requires only slightly more care.

Lemma 5.62. *Let X be a cubical manifold and $A \subseteq X$ a closed substratified space. The following split short exact sequence of Waldhausen categories is standard*

$$\begin{tikzcd} \mathsf{Fun}(\mathsf{Ent}_\Delta(X \setminus A), \mathsf{Set}_*) \arrow[r, "j_*"] & \mathsf{Fun}(\mathsf{Ent}_\Delta(X), \mathsf{Set}_*) \arrow[r, "i^*"] \arrow[l, bend left, "j^*"] & \mathsf{Fun}(\mathsf{Ent}_\Delta(A), \mathsf{Set}_*), \arrow[l, bend left, "i_*"] \end{tikzcd}$$

where $i\colon A \hookrightarrow X$ and $j\colon X \setminus A \hookrightarrow X$ are the inclusion maps.

Proof. Condition (3) of Definition 5.26 is inherited from Set_* where a cofibration is an injection and a cofiber sequence of finite pointed sets $S \hookrightarrow T \to *$ requires a bijection $S \cong T$.

Let $\mathcal{F} \in \mathsf{Fun}(\mathsf{Ent}_\Delta(X), \mathsf{Set}_*)$. As before, each component of the natural transformation $(j_* \circ j^*)(\mathcal{F}) \to \mathcal{F}$ is an isomorphism, except for the components corresponding to the strata of A. The component corresponding to the components of A are the inclusions of zero (the singleton set $*$), which is a cofibration. Therefore, $(j_* \circ j^*)(\mathcal{F}) \to \mathcal{F}$ is a cofibration in the functor category.

Finally, let $\varphi\colon \mathcal{F} \to \mathcal{F}'$ be a cofibration in $\mathsf{Fun}(\mathsf{Ent}_\Delta(X), \mathsf{Set}_*)$. We need to check that unique map

$$\psi\colon \mathcal{F} \coprod_{j_* \circ j^* \mathcal{F}} j_* \circ j^* \mathcal{F}' \to \mathcal{F}'$$

is a cofibration. The same (componentwise) argument works as before. That is, for a strata $T \subseteq A$ the pushout is identified with $\mathcal{F}(T)$ and $\psi_T = \varphi_T$. Let $T \neq S \in \mathsf{Ent}_\Delta(X)$, we are left to consider the commutative diagram below, where the square is a pushout,

$$\begin{tikzcd}
\mathcal{F}(S) \arrow[r, "\varphi_S"] \arrow[d, "\mathrm{Id}"'] & \mathcal{F}'(S) \arrow[d] \arrow[ddr, bend left, "\mathrm{Id}"] & \\
\mathcal{F}(S) \arrow[r] \arrow[drr, bend right, "\varphi_S"'] & \mathcal{F}(S) \coprod_\varphi \mathcal{F}'(S) \arrow[dr, "\psi_S"] & \\
& & \mathcal{F}'(S).
\end{tikzcd}$$

Hence, as φ_S is injective, so is ψ_S. $\square$

Arguing as in the preceding subsection, we deduce the following.

Lemma 5.63. *Let X be a cubical grid d-manifold with a finite number of strata. There is an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_*}(X)) \cong \bigvee_{x_0 \in X_0} \mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_*}(x_0)) \vee \bigvee_{x_1 \in X_1} \mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_*}(x_1)) \vee \ldots \vee \bigvee_{x_d \in X_d} \mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_*}(x_d))$$

where X_i is the set of i-strata of X.

The Barratt–Priddy–Quillen–Segal Theorem (see Chapter 4 of [63]) proves that there is an equivalence of spectra

$$\mathbb{S} \cong \mathbb{K}(\mathsf{Set}_*) \cong \mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_*}(x)),$$

for $x \in X$ a connected contractible stratum in a cubical manifold X, and where $\mathbb{S}$ is the sphere spectrum. Recall that the homotopy groups of $\mathbb{S}$ are the stable homotopy groups of spheres. Consequently, by assembling our work to this point, we have proven the following.

Theorem 5.64. *Let X be a cubical grid d-manifold with a finite number of strata. There is an equivalence of spectra*

$$\mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_*}(X)) \cong \bigvee_{X_0} \mathbb{S} \vee \bigvee_{X_1} \mathbb{S} \ldots \vee \bigvee_{X_d} \mathbb{S}$$

where X_i is the set of i-strata of X and where $\mathbb{S}$ denotes the sphere spectrum. In particular,

$$K_0(\mathsf{pMod}^{\mathsf{Set}_*}(X)) \cong \bigoplus_{X_0} \mathbb{Z} \oplus \bigoplus_{X_1} \mathbb{Z} \oplus \ldots \oplus \bigoplus_{X_d} \mathbb{Z},$$

and

$$K_1(\mathsf{pMod}^{\mathsf{Set}_*}(X)) \cong \bigoplus_{X_0} \mathbb{Z}/2 \oplus \bigoplus_{X_1} \mathbb{Z}/2 \oplus \ldots \oplus \bigoplus_{X_d} \mathbb{Z}/2.$$

As it is the central object in homotopy theory, much is known about $\mathbb{S}$, though mysteries remain. A remarkable theorem of Serre implies that $\pi_n(\mathbb{S})$ is finite for $n > 0$ and these groups are known up to around $n = 100$.

Remark 12. *If one wants to avoid using pointed sets/basepoints, one can consider the plain old category of sets* Set *and functions. This category does not have a zero object as the initial object is the empty set, while a final object is a singleton set. Hence,* Set *does not define a Waldhausen category in a straightforward manner. If one considers the subcategory* Set_i *consisting of the same objects, but where a morphism must be injective, one can define* $\mathbb{K}(\mathsf{pMod}^{\mathsf{Set}_i}(X))$*. Indeed,* Set_i *and the resulting functor category can be equipped with the structure of an* assembler *and Zakharevich defines* K*-theory for assemblers in [66] and [67]. It is again a consequence of the Barratt–Priddy–Quillen–Segal Theorem that for each* n *we have an isomorphism*

$$K_n(\mathsf{pMod}^{\mathsf{Set}_i}(X)) \cong K_n(\mathsf{pMod}^{\mathsf{Set}_*}(X)).$$

5.65 Euler Characteristic Surfaces and Virtual Diagrams

In this section, we give two applications of our K-theoretic work.

5.65.1 Euler Characteristic Surfaces and K_0

Let V be a d-parameter persistence module of vector spaces, with indexing set I. We choose an embedding $I \hookrightarrow \mathbb{N}^d$ and—in what follows—identify I with its image in the d-fold product of natural numbers. Then we can consider constructible cosheaves, F_V, on $\mathbb{R}^d$ stratified by the cubulation defined by $\mathbb{N}^d$ (i.e., constructible on the stratified parameter space $(\mathbb{R}^d; \mathbb{N}^d)$), for instance the propogated persistence cosheaf (Construction 4).

Definition 5.66 (Euler Manifold)**.** *The (scaled)* Euler manifold *of* V *is the constructible function* $\chi_V \colon \mathbb{R}^d \to \mathbb{Z}$ *given by* $\chi(x) = rank(F_V)_x$*, the rank of the costalk at* $x \in \mathbb{R}^d$*.*

Although similar in name to the Euler characteristic curves of Chapters 3 and 4, the Euler manifolds of Definition 5.66 are more generally defined.

Remark 13 (Connections to Chapters 3 and 4)**.** *Our choice of module V and cosheaf F allows us to recover d-dimensional generalizations of both (augmented) Euler characteristic curves and (augmented) Betti curves. For instance, when $d = 1$ and when V is a* Vect*-valued module reporting the Euler characteristic of a filtration, the Euler manifold of Definition 5.66 is equivalent to a scaled version of the (augmented) Euler characteristic curves χ (or $\hat{\chi}$) of Chapters 3 and 4. We might also recover a filtration-independent notion of Euler characteristic curve/manifold when V is* grVect*-valued and the cosheaf reports the alternating sum of vector space dimensions. If V is a* Vect*-valued module reporting homology in degree n, the Euler manifold instead corresponds to the nth (augmented) Betti curve, β_n (or $\hat{\beta}_n$).*

Note that any constructible $\mathbb{Z}$-valued function on $\mathbb{R}^d$ naturally defines a class in K_0. As noted in the proof of Theorem 5.60 (see also [41]), the class in K_0 of a Vect-valued cosheaf only depends on the dimension of its costalks, so we have the following.

Proposition 3. *Let V be a standard, finite persistence module of vector spaces and $(\mathbb{R}^d; \mathbb{N}^d)$ the stratified parameter space. Then,*

$$[F_V] = [\chi_V] \in K_0(\mathsf{pMod}^{\mathsf{Vect}}(\mathbb{R}^d)).$$

This proposition demonstrates one of the classical motivations for simplicial homology; fixing the functoriality of the Euler characteristic. In general, a map of complexes $f\colon X_\bullet \to Y_\bullet$

does not induce a map between $\chi(X_\bullet)$ and $\chi(Y_\bullet)$; only if f is covering map is there a multiplicative relationship between Euler characteristics. The *categorification* of the Euler characteristic to homology fixes this functoriality issue. Given any (co)homological setting there is an analogue of Euler class (in topology, this can be achieved by considering orientations for cohomology theories). Property 3 witnesses a K-theoretic Euler class.

Definition 5.67. *Let X be a cubical manifold, Ω a category, and $\mathcal{F} \in \mathsf{pMod}^{\Omega}(X)$ a persistence module. The* Euler class*, $\chi(\mathcal{F})$, of $\mathcal{F}$ is the K-class*

$$\chi(\mathcal{F}) := [\mathcal{F}] \in K_0(\mathsf{pMod}^{\Omega}(X)).$$

5.67.1 Virtual Diagrams

In [12], Bubenik and Elchesen describe the group completion of a monoid of persistence diagrams. The resulting equivalence classes are called *virtual persistence diagrams* and can be realized by extending the diagrams to include arbitrary points in the (extended) first quadrant, i.e., not just points above the diagonal. We will denote Bubenik and Elchesen's Abelian group of virtual persistence diagrams by $K_0(\mathsf{Diag})$. We now describe a homomorphism (and its image)

$$\delta\colon K_0(\mathsf{pMod}^{\mathsf{Vect}}_{\mathsf{fin}}(\mathbb{R}) \to K_0(\mathsf{Diag}),$$

where $\mathbb{R}$ is stratified by its subset $\mathbb{N}$, i.e., the parameter space is $(\mathbb{R};\mathbb{N})$.

To begin, let $\mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R})$ denote the category of Vect-valued constructible cosheaves on our stratified $\mathbb{R}$ that are eventually constant, i.e., there exists $N \in \mathbb{N}$ such that beyond N the cosheaf is constant. This category has a monoidal structure induced by $\oplus$ in Vect, so the objects in the category form a (commutative, unital) monoid.

We require a small tweak to the category Diag from [12]. As we allow features to persist for all future time, our persistence diagrams are built from the extended real line $\mathbb{R} \cup \{\infty\}$; this is a minor point and we suppress it from notation.

Now, as noted above, we have an identification of $\mathsf{Ent}_{\Delta}(\mathbb{R}; \mathbb{N})$ with the poset $\mathcal{Z}\mathcal{Z}_{\mathbb{N}}$. Hence, an object $\mathcal{F} \in \mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R}))$ is simply a representation of $\mathcal{Z}\mathcal{Z}_{\mathbb{N}}$ (which is eventually finite). Following [14], we use indecomposables of the associated representation of $\mathcal{Z}\mathcal{Z}_{\mathbb{N}}$ to associate a diagram to $\mathcal{F}$. More explicitly, we have the following assignment of a multi-set of points to a cosheaf

$$\check{\delta} \colon \mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R}) \to \mathsf{Diag} \subset \mathsf{K}_0(\mathsf{Diag}), \quad \mathcal{F} \mapsto \{(\mathsf{b_i}, \mathsf{d_i})\},$$

where each b_i and d_i correspond to the left and right indices (respectively) of an indecomposable element of the associated representation of $\mathcal{Z}\mathcal{Z}_{\mathbb{N}}$.

Note that $\check{\delta}$ can easily be adapted to be a map into barcodes, where, instead of a point (b, d), we draw a bar between b and d. This map $\check{\delta}$ (and the adaptation to signed barcodes) is nearly identical to the one described in Definition 2.6 of [14] with two notable differences. Firstly, the diagrams of [14] have points only on the integer lattice, whereas our diagrams have points on the 1/2-integer lattice. This is a consequence of us additionally considering

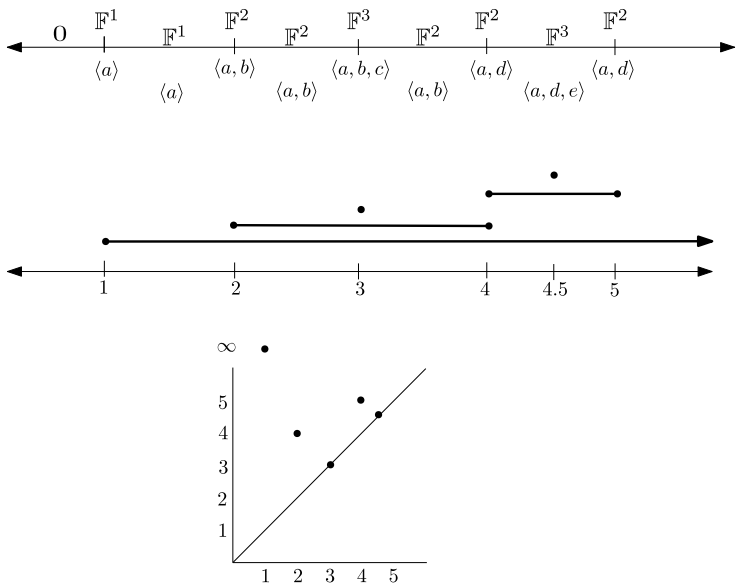


Figure 5.7: The result of applying δ to the cosheaf shown on the top of the figure is the persistence diagram shown on the bottom. In the middle, we have drawn the associated barcode. In the spirit of [14], we have shown all bars as closed intervals to emphasize that they do not necessarily arise from a monotone filtration. Note the presence of length-zero barcodes and on-diagonal points, corresponding to indecomposable elements supported at a single point.

edges of the stratification rather than only vertices, and of our convention to then index vertices by non-integers. Furthermore, the diagrams of [14] contain on-diagonal points only when the maps to a particular vertex both have a nontrivial kernel. Our diagrams allow for these type of on-diagonal points, but additionally allow for on-diagonal points when the maps from an edge to its endpoints both have a nontrivial kernel.

In general, these differences may be attributed to begining with persistence modules (the starting place for the map in [14]) or begining with persistence module cosheaves (the starting place for our map $\check{\delta}$). When one begins with persistence modules, the resulting cosheaf is a specific type – importantly, each edge has an identity morphism to at least one of its endpoints (see Construction 4). This allows [14] to only consider a poset on vertices,

which may be otained from our poset of vertices and edges by collapsing these identity morphisms. The language used is therefore more compatible with an explicit connection to filtrations, whereas our setting is generalized.

Returning to our primary goal, we note that our diagram map δ takes direct sums to sums of multisets.

Lemma 5.68. *The map*

$$\check{\delta}\colon \mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R}) \to K_0(\mathsf{Diag})$$

is a monoid homomorphism.

By the universal property of K_0 we obtain our desired homomorphism.

Corollary 5.69. *There is a homomorphism of Abelian groups*

$$\delta\colon K_0(\mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R})) \to K_0(\mathsf{Diag})$$

such that the following diagram commutes

$$\begin{tikzcd}
\mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R}) \arrow[r, "\check{\delta}"] \arrow[d] & K_0(\mathsf{Diag}) \\
K_0(\mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R})) \arrow[ur, "\delta"'] &
\end{tikzcd}$$

Because we have scaled/standardized our modules—as reflected by the parameter space $(\mathbb{R};\mathbb{N})$—the map δ has zero chance of being surjective (let alone an isomorphism). Following [12], let $\mathsf{Diag}\left(\left(\frac{1}{2}\mathbb{Z}\right)^2, \left(\frac{1}{2}\mathbb{Z}_{\geq 0}\right)^2\right)$ be monoid of (classical) persistence diagrams with half-integer (or infinite) coefficients. This monoid is a submonoid of the monoid of all (classical)

persistence diagrams, Diag, and we let $\mathsf{G} < K_0(\mathsf{Diag})$ denote the subgroup it generates. The following is clear from construction.

Proposition 4. *The subgroup* $\mathsf{G} < K_0(\mathsf{Diag})$ *is isomorphic to the image of the homomorphism*

$$\delta\colon K_0(\mathsf{pMod}_{\mathsf{fin}}^{\mathsf{Vect}}(\mathbb{R})) \to K_0(\mathsf{Diag}).$$

5.70 Discussion

In this chapter, we define a way to view multi-parameter zig-zag persistence modules as constructible cosheaves. This viewpoint allows for the structure of augmented persistence modules to be reflected in cosheaves as well. We then show an equivalence between a localized category of persistence modules and constructible cosheaves over $\mathbb{R}^d$ stratified by the d-fold product of natural numbers. We are able to extend the results of [37] and compute the algebraic K-theory of multi-parameter zig-zag persistence modules, showing that it is additive over strata. We end with a discussion of how our work relates back to other topics in TDA, including showing how the (augmented) Euler characteristic curves and (augmented) Betti curves of Chapters 3 and 4 can be seen as classes of K_0.

In ongoing work, we are particularly curious to explore viewing topological transforms as types of persistence module cosheaves themselves (valued over the parameterized sphere of directions). For instance, given a simplicial complex $K \in \mathbb{R}^d$, suppose that the topological transform associated to a descriptor type D is known to be faithful. Then, given a set of directions $S \in \mathbb{S}^{d-1}$ one way of determining whether or not a set of descriptors $\mathrm{D}(K, S)$ is faithful is to determine if the functor out of S that gives the costalks over each $s \in \mathbb{S}^{d-1}$ has

unique natural extension to the transform cosheaf over all of $\mathbb{S}^{d-1}$.

We also hope to identify and understand examples of persistence modules with nontrivial K_1. One way to address this is to use the mirror sequences of Sherman [58] to identify classes of K_1, producing mirror sequences by finding automorphisms of persistence modules. Another key component to this exploration is identifying "real-life" situations that arise in TDA where automorphisms of persistence modules occur.

REFERENCES CITED

[1] *Théorie des topos et cohomologie Étale des schémas. Tome 2.* Lecture Notes in Mathematics, Vol. 270. Springer-Verlag, Berlin-New York, 1972. Séminaire de Géométrie Algébrique du Bois-Marie 1963–1964 (SGA 4), Dirigé par M. Artin, A. Grothendieck et J. L. Verdier. Avec la collaboration de N. Bourbaki, P. Deligne et B. Saint-Donat.

[2] Erik J Amézquita, Michelle Y Quigley, Tim Ophelders, Jacob B Landis, Daniel Koenig, Elizabeth Munch, and Daniel H Chitwood. Measuring hidden phenotype: Quantifying the shape of barley seeds using the Euler characteristic transform. *in silico Plants*, 4(1):diab033, 2022.

[3] Shreya Arya, Justin Curry, and Sayan Mukherjee. A sheaf-theoretic construction of shape space. *arXiv:2204.09020*, 2022.

[4] David Ayala, John Francis, and Hiro Lee Tanaka. Local structures on stratified spaces. *Advances in Mathematics*, 307:903–1028, 2017.

[5] Werner Ballmann and Jacek Swiatkowski. *On groups acting on nonpositively curved cubical complexes.* Sonderforschungsbereich 256, 1998.

[6] Clark Barwick, Saul Glasman, and Peter Haine. Exodromy. *arXiv:1807.03281*, 2018.

[7] Robin Lynne Belton, Brittany Terese Fasy, Rostik Mertz, Samuel Micka, David L. Millman, Daniel Salinas, Anna Schenfisch, Jordan Schupbach, and Lucia Williams. Learning simplicial complexes from persistence diagrams. *Canadian Conference on Computational Geometry*, 2018.

[8] Robin Lynne Belton, Brittany Terese Fasy, Rostik Mertz, Samuel Micka, David L. Millman, Daniel Salinas, Anna Schenfisch, Jordan Schupbach, and Lucia Williams. Reconstructing embedded graphs from persistence diagrams. *Computational Geometry: Theory and Applications*, 2020.

[9] Nicolas Berkouk, Grégory Ginot, and Steve Oudot. Level-sets persistence and sheaf theory. *arXiv:1907.09759*, 2019.

[10] G. J. Bird, G. M. Kelly, A. J. Power, and R. H. Street. Flexible limits for 2-categories. *J. Pure Appl. Algebra*, 61(1):1–27, 1989.

[11] Andrew J. Blumberg, David Gepner, and Gonçalo Tabuada. A universal characterization of higher algebraic K-theory. *Geom. Topol.*, 17(2):733–838, 2013.

[12] Peter Bubenik and Alex Elchesen. Virtual persistence diagrams, signed measures, wasserstein distances, and banach spaces. *Journal of Applied and Computational Topology*, 6(4):429–474, 2022.

[13] Peter Bubenik and Nikola Milićević. Homological algebra for persistence modules. *Foundations of Computational Mathematics*, pages 1–46, 2021.

[14] Gunnar Carlsson and Vin de Silva. Zigzag persistence. *Found. Comput. Math.*, 10(4):367–405, 2010.

[15] Wojciech Chachólski, Barbara Giunti, Alvin Jin, and Claudia Landi. Decomposing filtered chain complexes: Geometry behind barcoding algorithms. *Computational Geometry*, 109:101938, 2023.

[16] Frédéric Chazal, Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Aarti Singh, and Larry Wasserman. On the bootstrap for persistence diagrams and landscapes. *arXiv:1311.0376*, 2013.

[17] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Symposium on Computational Geometry*, pages 119–126, 2006.

[18] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2022.

[19] Justin Curry. Abstract existence of cosheafification. *Abstract at math.upen.edu*, 2013.

[20] Justin Curry, Sayan Mukherjee, and Katharine Turner. How many directions determine a shape and other sufficiency results for two topological transforms. *Transactions of the American Mathematical Society, Series B*, 9(32):1006–1043, 2022.

[21] Justin Curry and Amit Patel. Classification of constructible cosheaves. *Theory Appl. Categ.*, 35:Paper No. 27, 1012–1047, 2020.

[22] J Darrotto. Convex optimization and Euclidean distance geometry. *Palo Alto: Meboo*, 2013.

[23] Kristin DeSplinter, Satyan L Devadoss, Jordan Readyhough, and Bryce Wimberly. Nets of higher-dimensional cubes. In *Canadian Conference on Computational Geometry*, pages 114–120, 2020.

[24] Luis Diaz-Garcia, Giovanny Covarrubias-Pazaran, Brandon Schlautman, Edward Grygleski, and Juan Zalapa. Image-based phenotyping for identification of qtl determining fruit shape and size in american cranberry (vaccinium macrocarpon l.). *PeerJ*, 6:e5461, 2018.

[25] Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2010.

[26] Herbert Edelsbrunner, John Harer, and Amit K Patel. Reeb spaces of piecewise linear mappings. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 242–250, 2008.

[27] Brittany Terese Fasy, Samuel Micka, David L Millman, Anna Schenfisch, and Lucia Williams. Challenges in reconstructing shapes from Euler characteristic curves. *arXiv:1811.11337*, 2018.

[28] Brittany Terese Fasy, Samuel Micka, David L Millman, Anna Schenfisch, and Lucia Williams. Challenges in reconstructing shapes from Euler characteristic curves. *arXiv:1811.11337*, 2018.

[29] Brittany Terese Fasy, Samuel Micka, David L Millman, Anna Schenfisch, and Lucia Williams. A faithful discretization of the augmented persistent homology transform. *arXiv:1912.12759*, 2019.

[30] Brittany Terese Fasy, Samuel Micka, David L Millman, Anna Schenfisch, and Lucia Williams. Efficient graph reconstruction and representation using augmented persistence diagrams. *arXiv:2212.13206.*, 2022.

[31] Brittany Terese Fasy and Amit Patel. Persistent homology transform cosheaf. *arXiv:2208.05243*, 2022.

[32] Thomas M. Fiore and Malte Pieper. Waldhausen additivity: classical and quasicategorical. *J. Homotopy Relat. Struct.*, 14(1):109–197, 2019.

[33] Robert Ghrist and Yasuaki Hiraoka. Applications of sheaf cohomology and exact sequences to. *Notes on the Institute of Mathematical Analysis*, 1752:31–40, 2011.

[34] Robert Ghrist, Rachel Levanger, and Huy Mai. Persistent homology and euler integral transforms. *Journal of Applied and Computational Topology*, 2:55–60, 2018.

[35] Robert Ghrist, H Owen, and Michael Robinson. Discrete topological imaging for multipath environments. 2010.

[36] Ryan E Grady and Anna Schenfisch. Natural stratifications of Reeb spaces and higher Morse functions. *arXiv:2011.08404*, 2020.

[37] Ryan E Grady and Anna Schenfisch. Zig-zag modules: Cosheaves and K-theory. *arXiv:2110.04591*, 2021.

[38] Allen Hatcher. Algebraic topology, Cambridge Univ. *Press, Cambridge*, 2002.

[39] Abigail Hickok. Persistence diagram bundles: A multidimensional generalization of vineyards. *arXiv:2210.05124*, 2022.

[40] Niles Johnson and Donald Yau. *2-dimensional categories*. Oxford University Press, Oxford, 2021.

[41] Masaki Kashiwara and Pierre Schapira. *Sheaves on manifolds*, volume 292 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1994. With a chapter in French by Christian Houzel, Corrected reprint of the 1990 original.

[42] Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021.

[43] Mao Li, Hong An, Ruthie Angelovici, Clement Bagaza, Albert Batushansky, Lynn Clark, Viktoriya Coneva, Michael J Donoghue, Erika Edwards, Diego Fajardo, et al. Topological data analysis as a morphometric method: using persistent homology to demarcate a leaf morphospace. *Frontiers in plant science*, 9:553, 2018.

[44] J. Lurie. Higher algebra. available at Author's Homepage, 2017.

[45] Alexander McCleary and Amit Patel. Edit distance and persistence diagrams over lattices. *arXiv 2010.07337*, 2020.

[46] Facundo Mémoli and Ling Zhou. Stability of filtered chain complexes. *arXiv:2208.11770*, 2022.

[47] David L. Millman and Vishal Verma. A slow algorithm for computing the Gabriel graph with double precision. *CCCG '11: Proceedings of the 23rd Annual Canadian Conference on Computational Geometry*, 2011.

[48] Michael J Panik. *Fundamentals of convex analysis: duality, separation, representation, and resolution*, volume 24. Springer Science & Business Media, 2013.

[49] Amit Patel. Generalized persistence diagrams. *J. Appl. Comput. Topol.*, 1(3-4):397–419, 2018.

[50] Andrei V Prasolov. Cosheafification. *arXiv:1605.01555*, 2016.

[51] George Raptis. D\'{e} vissage for waldhausen k-theory. *arXiv:1811.09564*, 2018.

[52] Eitan Richardson and Michael Werman. Efficient classification using the Euler characteristic. *Pattern Recognition Letters*, 49:99–106, 2014.

[53] Jonathan Rosenberg. *Algebraic K-theory and its applications*, volume 147 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[54] C. P. Rourke and B. J. Sanderson. *Introduction to piecewise-linear topology*. Springer-Verlag, New York-Heidelberg, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 69.

[55] Ameer Saadat-Yazdi, Rayna Andreeva, and Rik Sarkar. Topological detection of alzheimer's disease using Betti curves. In *Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and Its Applications for Medical Data: 4th International Workshop, iMIMIC 2021, and 1st International Workshop, TDA4MedicalData 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 4*, pages 119–128. Springer, 2021.

[56] Michah Sageev and Daniel T Wise. The Tits alternative for CAT (0) cubical complexes. *Bulletin of the London Mathematical Society*, 37(5):706–710, 2005.

[57] Apoorva Bharthur Sanjay, Alice Patania, Xiaoran Yan, Diana Svaldi, and Liana G Apostolova. Characterization of genetic expression patterns in MCI using a multiomics approach and neuroimaging endophenotypes. In *2020 Alzheimer's Association International Conference*. ALZ, 2020.

[58] Clayton Sherman. k_1 of exact categories by mirror image sequences. *Journal of K-Theory*, 11(1):155–181, 2013.

[59] L.N. Trefethen and D. Bau. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, 1997.

[60] David Treumann. Exit paths and constructible stacks. *Compos. Math.*, 145(6):1504–1532, 2009.

[61] Katharine Turner, Sayan Mukherjee, and Doug M. Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[62] Friedhelm Waldhausen. Algebraic K-theory of spaces. In *Algebraic and geometric topology (New Brunswick,N.J., 1983)*, volume 1126 of *Lecture Notes in Math.*, pages 318–419. Springer, Berlin, 1985.

[63] Charles A. Weibel. *The K-book*, volume 145 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2013. An introduction to algebraic K-theory.

[64] Jonathan Woolf. The fundamental category of a stratified space. *arXiv:0811.2580*, 2008.

[65] Xiangdong Xie. Foldable cubical complexes of nonpositive curvature. *Algebraic & Geometric Topology*, 4(1):603–622, 2004.

[66] Inna Zakharevich. The K-theory of assemblers. *Adv. Math.*, 304:1176–1218, 2017.

[67] Inna Zakharevich. On K_1 of an assembler. *J. Pure Appl. Algebra*, 221(7):1867–1898, 2017.

[68] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. *Discrete & Computational Geometry*, 33(2):249–274, 2005.