How Small Can Faithful Sets Be? Ordering Topological Descriptors

Abstract

Recent developments in shape reconstruction and comparison call for the use of many different (topological) descriptor types, such as persistence diagrams and Euler characteristic functions. We establish a framework to quantitatively compare the strength of different descriptor types, setting up a theory that allows for future comparisons and analysis of descriptor types and that can inform choices made in applications. We use this framework to partially order a set of six common descriptor types. We then give lower bounds on the size of sets of descriptors that uniquely correspond to simplicial complexes, giving insight into the advantages of using verbose rather than concise topological descriptors.

1 Introduction

The persistent homology transform and Euler characteristic transform were first explored in [38], which shows the uncountable set of persistence diagrams (or Euler characteristic functions, respectively) corresponding to lower-star filtrations in every possible direction uniquely represents the shape being filtered. That is, the uncountable set of topological descriptors is *faithful* for the shape. Of course, applications can only use finite sets, which are not guaranteed to be faithful. This motivates theoretical work on finding finite faithful sets of descriptors [3, 7, 13, 30], and such theoretical results support the use of topological descriptors in shape comparison applications. A wide range of descriptor types are used in applications, such as versions of persistence diagrams [4,20,21,35,39,42], Euler characteristic functions [1, 6, 19, 24, 29, 31, 34], Betti functions [15, 23, 33, 36, 41, 43], and other descriptor types [2, 16, 37].

Faithfully representing a shape with a small number of descriptors is desirable for computational and storage reasons. How, then, should investigators choose the particular topological descriptor type to use in applications? While computational complexities of computing each topological descriptor type are well-studied, it is not yet known how the use of particular descriptor types impacts the minimum size of faithful sets. This uncertainty motivates our main questions: *how can we rigorously compare descriptor types in terms of their ability to uniquely correspond to shapes*, and *how do popular descriptor types compare?*

We prove the parital order of Figure 1. Additionally, we provide lower bounds on the cardinality of faithful sets for both concise and verbose descriptors, and identify properties that indicate concise descriptors are generally much weaker than verbose descriptors. This suggests applications research may benefit from the use of verbose descriptors instead of the more widely adopted concise descriptors.

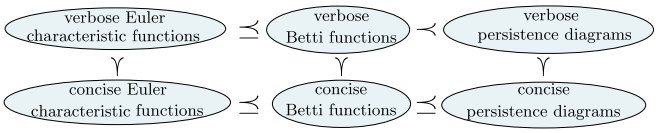


Figure 1: Summary of the relations between common descriptor types. For example, concise persistence diagrams are always at least as efficient as concise Betti functions at forming faithful sets, where efficiency is measured by cardinality of faithful sets.

2 Preliminary Considerations

In this section, we provide background and definitions used throughout.

Simplicial Complexes and Filtrations We assume the reader has familiarity with foundational ideas from topology; as presented in, e.g., [9, 17]. We establish a few notational conventions here. For a simplicial complex K and $i \in \mathbb{N}$, we use the notation K_i for its i-simplices and n_i as the number of i-simplices. Furthermore, we assume our simplicial complexes are *abstract* simplicial complexes immersed in Euclidean space such that each simplex is embedded and the vertices are in general position; see Assumption 1 in Appendix C. A *filter* of K is a map $f\colon K \to \mathbb{R}$ such that, each sublevel set $f^{-1}(-\infty, t]$ is either empty or a simplicial complex. Letting $F(t) := f^{-1}(-\infty, t]$, the sequence $\{F(t)\}_{t\in\mathbb{R}}$ is the *filtration* associated to f. For each $k \in \mathbb{N}$, the inclusion $F(i) \hookrightarrow F(j)$ induces a linear map on homology, $H_k(F(i)) \to H_k(F(j))$. We write $\beta_k^{i,j}(K, f)$ to mean rank of this map, or simply $\beta_k^{i,j}$ if K and f are clear from context. Any filter function has *compatible index filters*, which are bijections $f' : K \to \{1, 2, \ldots, \#K\}$ such that f' is a filter and for all $\tau, \sigma \in K$ with $f(\tau) \leq f(\sigma)$, then $f'(\tau) \leq f'(\sigma)$.

The *lower-star filter* of a simplicial complex K immersed in $\mathbb{R}^d$ with respect to some direction $s \in \mathbb{S}^{d-1}$, is the map $f_s : K \to \mathbb{R}$ that takes a simplex σ to the maximum height of its vertices with respect to direction s, i.e., $f_s(\sigma) := \max\{s \cdot v \mid v \in K_0 \cap \sigma\}$, where $s \cdot v$ denotes the dot product.

Faithfully Representing a Simplicial Complex Since we define relations based on the ability of descriptor types to represent particular filtrations of simplicial complexes, we take the following definition.

Definition 1 (Topological Descriptors) *A* (topological) descriptor type *is a map whose domain is the collection of filtered simplicial complexes. Given such map,* D*, a* (topological) descriptor of type D *is the image of a specific filtered simplicial complex under* D*.*

When considering many filtrations of the same simplicial complex, we may index the filtrations by some parameter set, P. If a descriptor of type D corresponds to a filtration of a simplicial complex K where the filtration is parameterized by $p \in P$, we use the notation $\mathrm{D}(K, p)$, or $\mathrm{D}(p)$ when K is clear from context. We refer to the parameterized set of descriptors as $\mathrm{D}(K, P) := \{(p, \mathrm{D}(K, p))\}_{p \in P}$.

We compare descriptor types by quantifying the number of descriptors needed to uniquely identify a shape. Specifically:

Definition 2 (Faithful) *Let K be a simplicial complex, P parameterize a set of filtrations of K, and* D *be a topological descriptor type. We say that* $\mathrm{D}(K, P)$ is faithful *if, for any simplicial complex L we have the equality* $\mathrm{D}(L, P) = \mathrm{D}(K, P)$ *if and only if* $L = K$*.*

We note here that $|\mathrm{D}(K, P)| = |P|$. See Appendix B for a reformulation of Definition 2 in terms of set intersections, as well as a lemma that can be gained from this perspective.

3 Six Common Descriptor Types

The set we partially order is the strength equivalence classes of six popular descriptor types, which we define here. We begin with concise persistence diagrams.

Definition 3 (Concise Persistence Diagram, ρ) *Let $f : K \to \mathbb{R}$ be a filter function. For $k \in \mathbb{N}$, the* k-dimensional persistence diagram *is:*

$$\rho_k^f := \Big\{(i,j)^{\mu^{(i,j)}} \ \textit{s.t.}\ (i,j) \in \overline{\mathbb{R}}^2 \\ \textit{and}\ \mu^{(i,j)} = \beta_k^{i,j-1} - \beta_k^{i,j} - \beta_k^{i-1,j-1} + \beta_k^{i-1,j}\Big\},$$

where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ and $(i,j)^m$ denotes m copies of the point (i,j). The concise persistence diagram *of f, denoted ρ^f, is the indexed union of all k-dimensional concise persistence diagrams $\rho^f := \cup_{k \in \mathbb{N}} \rho_k^f$.*

Since simplices can appear at the same parameter value in a filtration, not all cycles are represented in the persistence diagram. However, having every simplex "appear" in a topological descriptor is helpful, in addition to being natural. Thus, we introduce *verbose* descriptors, which contain this information. We define verbose descriptors via compatible index filtrations; by Lemma 52 and Corollaries 54-55 of [12], the resulting descriptors are well-defined and independent of our choice of compatible index filtration. We begin with *verbose persistence diagrams*:

Definition 4 (Verbose Persistence Diagram) *Let $f : K \to \mathbb{R}$ be a filter for K, and let f' be a compatible index filter. For $k \in \mathbb{N}$, the* k-dimensional verbose persistence diagram *is the following multiset:*

$$\tilde{\rho}_k^f := \left\{ (f(\sigma_i), f(\sigma_j)) \ \textit{s.t.}\ (i,j) \in \rho_k^{f'} \right\}.$$

The verbose persistence diagram *of f, denoted $\tilde{\rho}^f$, is the indexed union of all $\tilde{\rho}_k^f$.*

Recording invariants other than homology leads to other topological descriptor types; recording Betti numbers gives us *Betti functions*.

Definition 5 (Betti Functions, β and $\tilde{\beta}$) *Let $f \colon K \to \mathbb{R}$ be a filter function. The* kth Betti function, *$\beta_k^f : \mathbb{R} \to \mathbb{Z}$, is defined by*

$$\beta_k^f(t) := \beta_k\left(f^{-1}(-\infty, t]\right).$$

The indexed collection of Betti functions for all dimensions, $\beta^f := \{\beta_k^f \mid k \in \mathbb{N}\}$, is the Betti function*.*

Let f' be an index filter compatible with filter function f. We call $\sigma \in K$ positive *(respectively,* negative*)* for *β_k if the inclusion of σ into the index filtration of f' increases (resp., decreases) β_k. We denote the positive (resp., negative) simplices by $K_k^+ \subseteq K_k$ (and $K_{k+1}^- \subseteq K_{k+1}$). Then the* kth verbose Betti function, *$\tilde{\beta}_k^f : \mathbb{R} \to \mathbb{Z}^2$, is defined by*

$$\tilde{\beta}_k^f(p) := \Big(\left|\{\sigma \in K_k^+ \ \textit{s.t.}\ f(\sigma) \le p\}\right|, \\ \left|\{\sigma \in K_{k+1}^- \ \textit{s.t.}\ f(\sigma) \le p\}\right| \Big).$$

The collection of verbose Betti number functions for each dimension is known as the verbose Betti function *and is denoted $\tilde{\beta}^f$.*

If we record *Euler characteristic* in a filtration, we obtain *Euler characteristic functions.*[1]

Definition 6 (Euler Characteristic Functions, χ and $\tilde{\chi}$) *Let $f \colon K \to \mathbb{R}$ be a filter function. The* Euler characteristic function, *$\chi^f : \mathbb{R} \to \mathbb{Z}$, is defined by:*

$$\chi^f(p) := \chi\left(\{f^{-1}(-\infty, p]\}\right).$$

Let f' be an index filter compatible with f. We call $\sigma \in K$ even *(respectively,* odd*) if the dimension of σ is even (resp., odd). Denoting the set of even (resp., odd) simplices by E (and O), the* verbose Euler characteristic function, *$\tilde{\chi}^f : \mathbb{R} \to \mathbb{Z}^2$, is defined by*

$$\tilde{\chi}^f(p) := \left(|\sigma \in E \ \textit{s.t.}\ f(\sigma) \le p|, |\sigma \in O \ \textit{s.t.}\ f(\sigma) \le p|\right).$$

In other words, $\tilde{\chi}^f$ represents χ^f as a parameterized count of even- and odd-dimensional simplices.

In each of the descriptor types above, we drop the superscript f when it is clear from context. See Appendix A for examples of these descriptor types.

[1] Euler characteristic functions and Betti functions are sometimes called Euler (characteristic) curves or Betti curves, respectively.

While concise descriptors may feel more familiar, verbose descriptors are not new. Many algorithms for computing persistence (e.g., [9, Chapter VII]), explicitly compute events with trivial lifespan. In [25], the definition of persistence diagrams agrees with our Definition 4. Verbose descriptors are closely connected to the charge-preserving morphisms of [14, 26]. In [40], verbose persistence is defined via filtered chain complexes; [5, 27, 28, 44] also take this view as a foundational definition. The behavior of verbose versus concise descriptors is explored in [10, 28, 45].

Verbose (concise) descriptors are sometimes called augmented (non-augmented, respectively) in the literature. We refer to points on a verbose diagram with zero-lifespan as *instantaneous*. Such points correspond to length-zero barcodes in a verbose barcode, which are sometimes referred to as *ephemeral*.

While we chose the six descriptor types above due to their relevance in applications, we emphasize that Definition 1 is very general. We explore a few pathological descriptor types in Appendix C.

4 Relating Descriptor Types

We now develop tools to compare descriptor types, by comparing the sizes of faithful sets. Given a topological descriptor type D and a simplicial complex K immersed in $\mathbb{R}^d$, let the infimum size of faithful sets for K be denoted

$$\Gamma(K, D) := \inf_{D(K,P) \text{ faithful}} \{|P|\}.$$

Intuitively, the stronger D is, the smaller $\Gamma(K, D)$. Often, we find $\Gamma(K, D)$ is finite. For some descriptors and K, we find $\Gamma(K, D) = \aleph_0$ (the cardinality of $\mathbb{N}$) or $\Gamma(K, D) = \aleph_1$ (the cardinality of $\mathbb{R}$); see Appendix C for examples. If no faithful set of type D exists for K, we write $\inf_{x \in \emptyset}\{x\} = \aleph_\top$, and we think of this as "the highest" cardinality. By the axiom of choice, $\aleph_0 < \aleph_1$; see e.g., [18, Ch. 2]. Thus, we have a total order on possible values of $\Gamma(K, D)$:

$$c < \aleph_0 < \aleph_1 < \aleph_\top,$$

where $c \in \mathbb{N}$.

Definition 7 (Strength Relation) *Let A and B be two topological descriptor types. If, for every simplicial complex K immersed in $\mathbb{R}^d$, we have $\Gamma(K, A) \geq \Gamma(K, B)$, then we say that A is* weaker *than B (and B is* stronger *than A) denoted $[A] \preceq [B]$. If $[A] \preceq [B]$ and $[B] \preceq [A]$, then we say that A and B have equal strength, denoted $[A] = [B]$.*

The relations $=$ are $\preceq$ are well-defined on strength equivalence classes. See Lemma 16 in Appendix D.1. Also see Example 1 of Appendix C for two different descriptor types in the same equivalence class.

We write $[A] \prec [B]$ if $[A] \preceq [B]$ and $[A] \neq [B]$. That is, if $[A] \preceq [B]$ and there exists a simplicial complex for which the minimum faithful set of type B is strictly smaller than that of type A, or for which there exists a faithful set of type B but not of type A. Descriptor types need not be comparable; see Lemma 14 of Appendix C.

We conclude this section by defining reduction of one descriptor to another, and show this is a valid strategy for determining equivalence class order.

Definition 8 (Reduction) *Let A and B be two topological descriptor types. We say B is* reducible *to A if, for all simplicial complexes K and any filtration f of K, we can compute $A(f)$ from $B(f)$ alone.*

Intuitively, B is at least as informative as A. More formally, we have the following lemma, whose proof is in Appendix D.1:

Lemma 1 *Let A and B be two topological descriptor types. If B is reducible to A, we have $[A] \preceq [B]$.*

5 A Proof of Partial Order

In this section, we provide a partial order on the six topological descriptors of Section 3. Omitted proofs are in Appendix D.2. While the results and definitions of previous sections were general, we now focus on descriptors corresponding to lower-star filtrations.

By simple reduction arguments, we immediately have the following lemma.

Lemma 2 *$[\chi] \preceq [\beta] \preceq [\rho]$ and $[\tilde{\chi}] \preceq [\tilde{\beta}] \preceq [\tilde{\rho}]$.*

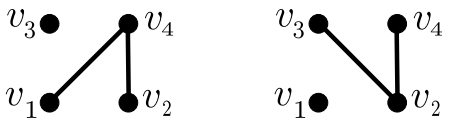
We also use reduction to order a class of a concise descriptor type and its verbose counterpart.

Lemma 3 *$[\chi] \preceq [\tilde{\chi}]$, $[\beta] \preceq [\tilde{\beta}]$, and $[\rho] \preceq [\tilde{\rho}]$.*

Next, we see that no concise class is equal to a verbose class.

Lemma 4 *For $\mathrm{D} \in \{\chi, \beta, \rho\}$ and $\tilde{\mathrm{D}} \in \{\tilde{\chi}, \tilde{\beta}, \tilde{\rho}\}$, we have $[\mathrm{D}] \neq [\tilde{\mathrm{D}}]$.*

The proof is given by considering a single edge in $\mathbb{R}^2$, and showing a minimal faithful set of type $\tilde{\mathrm{D}}$ has cardinality two; whereas, a minimal faithful set of type D has cardinality at least three; see Appendix D.2 for full details.



(a) K (b) K'

Figure 2: The two simplicial complexes of Lemma 5.

In [13], faithfulness is shown via knowing the dimension of event rather than any birth-death pairings, so the resulting faithful sets of augmented persistence diagrams and augmented Betti functions have equal cardinality for every simplicial complex. One might wonder, then, if $[\tilde{\rho}]$ equals $[\tilde{\beta}]$. However, this is an incorrect leap; faithful sets of [13] are almost certainly not minimal. The next lemma gives an instance where birth-death pairings matter, rather than just an event's existence or dimension.

Lemma 5 $[\tilde{\beta}] \prec [\tilde{\rho}]$.

Proof. We know by Lemma 2 that $[\tilde{\beta}] \leq [\tilde{\rho}]$. We must show $[\tilde{\beta}] \neq [\tilde{\rho}]$, that is, there exists a simplicial complex for which the cardinality of minimal faithful sets of augmented Betti functions and augmented persistence diagrams differ. Consider the simplicial complex K in $\mathbb{R}^2$ 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 Figure 2(a).

Let $e_1 = (1,0)$ and $e_2 = (0,1)$. We first claim that $\tilde{\rho}(K, \{e_1, e_2\})$ is faithful, meaning $\Gamma(K, \tilde{\rho}) \leq 2$. Such diagrams uniquely identify the vertex set of K by [3, Lemma 4]; we provide further details here. From $\tilde{\rho}(e_1)$, we know K has four vertices, two with x-coordinate 0, and two with x-coordinate 1. Similarly, from $\tilde{\rho}(e_2)$, we know two of the four vertices have y-coordinate 0 and two have y-coordinate 1. There are exactly four ways to pair our x- and y-coordinates, so we know the locations of each vertex. See Figure 3(b).

Next, we argue that such diagrams uniquely identify the edges of K. From $\tilde{\rho}(e_2)$ in degree zero we see an instantaneous birth/death at height one as well as a connected component born at height zero that dies at height one, so we know K has exactly two edges that merge connected components at height one. Namely, we know we have either the edges $[v_1, v_3]$ and $[v_2, v_3]$, or $[v_1, v_4]$ and $[v_2, v_4]$, i.e., we have either of the two simplicial complexes shown in Figure 2. Because $\tilde{\rho}(e_1)$ sees two zero-dimensional births at height zero with an infinite lifespan, we know there is no edge from v_1 to v_3. Finally, since higher homology is trivial, we know there are no other simplices and have determined K exactly; thus, we have a faithful set of size two.

We next show that $\Gamma(K, \tilde{\beta}) > 2$. Suppose, by way of contradiction, that s_1 and s_2 are two directions such that $\tilde{\beta}(K, \{s_1, s_2\})$ is a faithful set. We first show, without loss of generality, $s_1 \in \{e_1, -e_1\}$ and $s_2 \in \{e_2, -e_2\}$. Suppose this is not the case. Since $s_1 = -s_2$ clearly does not correspond to a faithful set, further suppose $s_1 \neq -s_2$. Then at least one of s_1 or s_2 sees the vertices of K at more than two distinct heights; see Figure 3(a). In order to know the precise coordinates of each vertex, we need to correctly pair heights in directions s_1 and s_2. However, since at least one of s_1 or s_2 reports more than two distinct heights, we have more than four possible pairings (see also [3, Lemma 4]). We claim it is not possible to find the four correct pairings. The degree-zero information $\tilde{\beta}_0(s_1)$ and $\tilde{\beta}_0(s_2)$ alone is insufficient, as it only tells us the heights of vertices. From $\tilde{\beta}_1$, we know the height of edges, which only confirms the height of the top vertex and that there is some vertex below, information we already had from $\tilde{\beta}_0$.

Thus, in order to know K_0 precisely, we must have $s_1 \in \{e_1, -e_1\}$ and $s_2 \in \{e_2, -e_2\}$. However, for each of these four directions, the associated verbose Betti function is not able to distinguish K from K' as in Figure 2(b). For instance, both of $\tilde{\beta}(K, e_1)$ and $\tilde{\beta}(K', e_1)$ see two vertices at height zero, and two vertices and two edges at height one, meaning $\tilde{\beta}(K, e_1) = \tilde{\beta}(K', e_1)$. The other possible values of s_1 and s_2 are similar. Thus, we have found a faithful set of verbose persistence diagrams with cardinality two, but have shown any faithful set of verbose Betti functions must have cardinality greater than two, and the claim follows. □

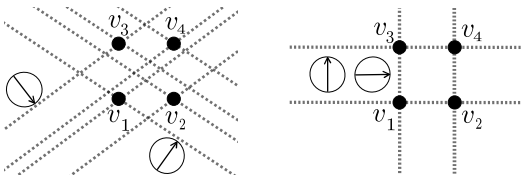


Figure 3: For four vertices on an axis-aligned grid, the heights of events from two filtrations in the indicated directions are shown as dashed grey lines. Although we know the number of vertices on each line, for two directions not both in $\{\pm e_1, \pm e_2\}$, as in (a), we can not identify vertex locations. Only when choosing one each of $\pm e_1$ and $\pm e_2$, as in (b), is the set of vertices satisfying these constraints unique.

Combining results, we arrive at our main theorem.

Theorem 6 (Partial Ordering) *The partial order of strength classes of topological descriptor types shown in Figure 1 is correct.*

6 Bounds on Faithful Sets

Here, we provide lower bounds on the size of faithful sets of the six descriptor types of Section 3.

6.1 Concise Descriptor Bounds

A defining feature of concise descriptors is that there are not generally events at every vertex height in a filtration. The closer a feature is to coplanar, the smaller the range of directions that can detect it becomes ([11, Sec. 4] explores this specifically for Euler characteristic functions). Difficulty detecting the presence or absence of structures near to the same affine subspace puts greater restrictions on the ability of concise descriptors to form faithful sets. We use the following definition to help this claim precise.

Definition 9 (Simplex Envelope) *Let K be a simplicial complex in $\mathbb{R}^d$, let $\sigma \in K$, and let $S \subseteq \mathbb{S}^{d-1}$. Then, we define the* envelope of σ, *denoted $\mathcal{E}_\sigma^S$, as the intersection of (closed) supporting halfspaces*

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

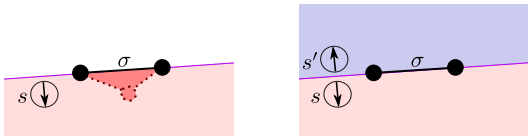
If S is clear from context, we write $\mathcal{E}_\sigma$. By the dimension *of $\mathcal{E}_\sigma$, we mean the largest dimension of ball that can be contained entirely in $\mathcal{E}_\sigma$.*

Since $\mathcal{E}_\sigma^S$ is an intersection of convex regions, it is itself convex. Furthermore, with respect to each $s \in S$, the height of each point of σ is greater than or equal to its minimum vertex, so $\mathcal{E}_\sigma^S$ contains σ.

Remark 1 *The simplex envelopes of Definition 9 have connections to well-studied topics such as convex cones, support functions, etc. See [8, 32]. In particular, [32, Thm 3.1.1, Cor 3.1.2] establish that a simplex envelope corresponding to the entire sphere of directions is the simplex itself.*

We use simplex envelopes to define a necessary condition for concise descriptors to form a faithful set.

Lemma 7 (Envelopes for Faithful Concise Sets) *Let K be a simplicial complex immersed in $\mathbb{R}^d$, let $D \in \{\chi, \beta, \rho\}$, and let $S \subseteq \mathbb{S}^{d-1}$ so that $\mathrm{D}(K, S)$ is faithful. Then, for any maximal simplex σ in K, the dimension of $\mathcal{E}_\sigma$ equals the dimension of σ.*



(a) (b)

Figure 4: With only the single direction s perpendicular to maximal edge σ in $\mathbb{R}^2$, the envelope $\mathcal{E}_\sigma^s$ is two-dimensional. Then, we could place an adversarial two-simplex contained in $\mathcal{E}_\sigma^s$ that is undetectable by $D(s)$, for $D \in \{\chi, \beta, \rho\}$, as in (a). In (b), the inclusion of s' reduces $\mathcal{E}_\sigma^{\{s,s'\}}$ to a linear subspace (purple intersection of pink and blue halfspaces) and the adversarial two-simplex would be detected by $D(s')$.

A proof of Lemma 7 appears in Appendix D.3. See Figure 4 for an example of what might go wrong if a simplex envelope does not satisfy the conditions of Lemma 7. Since we require the envelopes of a k-simplex to be k-dimensional, and since envelopes are the intersections of closed half spaces, standard arguments from manifold theory give us the following.

Corollary 1 (Concise Descriptors Per Maximal Simplex) *Let K be a simplicial complex immersed in $\mathbb{R}^d$, let $\mathrm{D} \in \{\chi, \beta, \rho\}$, and let $S \subseteq \mathbb{S}^{d-1}$. If $\mathrm{D}(K, S)$ is faithful, then for each maximal simplex $\sigma \in K$ of dimension $k < d$, the set S has at least $d - k + 1$ directions perpendicular to σ. If $k < d - 1$, these directions are pairwise linearly independent.*

Lemma 7 and Corollary 1 each give us the following.

Corollary 2 (Tight Lower Bound) *Let K be a simplicial complex in $\mathbb{R}^d$, $\mathrm{D} \in \{\chi, \beta, \rho\}$, and $S \subseteq \mathbb{S}^{d-1}$. Suppose that $\mathrm{D}(K, S)$ is faithful. Then $|S| \geq d + 1$, and this bound is tight.*

This bound is met whenever K is a single vertex. However, minimal faithful sets of concise descriptors are generally much larger. Counteracting the need for perpendicular directions is the fact that, as d increases, more simplices span common hyperplanes, so perpendicular directions can increasingly be shared. We use these observations to lower bound bound the worst-case size of faithful set of consice descriptors.

Theorem 8 (Lower Bound for Worst-Case Concise Descriptor Complexity) *Let $\mathrm{D} \in \{\chi, \beta, \rho\}$ and let K be a simplicial complex in $\mathbb{R}^d$ with n_1 edges. Then the worst-case cardinality of a minimal descriptor set of type D is $\Omega(d + n_1)$.*

Proof. We construct a simplicial complex, K, and bound the minimum cardinality of a faithful set for K. Suppose that, for $d > 2$, that K is a graph in $\mathbb{R}^d$ with $n_1 < d - 1$ edges, and for some $S \subseteq \mathbb{S}^{d-1}$, the set $\mathrm{D}(K, S)$ is faithful. Then by Lemma 7, the envelope of each maximal edge σ must be one-dimensional. Then by Corollary 1, for every such σ, S contains $d - 1 + 1 = d$ pairwise linearly independent directions perpendicular to σ. Let S^* be a minimal subset of directions in S satisfying the conditions of perpendicularity and one-dimensional envelopes.

To build S^*, first note all edges of K are contained in a common n_1-plane, so there is a $(d - n_1 - 1)$-sphere's worth of directions perpendicular to *all* edges simultaneously. Such directions are maximally efficient in the sense that each can "count" for all edges at once. We choose any $d - 1$ pairwise linearly independent directions from this sphere to be included in S^*. Now we need an additional perpendicular direction for each edge to bring the total for each edge to d. 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^* has cardinality $d - 1 + n_1$. Since $|S^*|$ lower bounds $|S|$, we find $|S| \in \Omega(d + n_1)$. □

6.2 Verbose Descriptor Bounds

We now shift to verbose descriptors, and begin with the tight lower bound.

Lemma 9 (Tight Lower Bound) *Let K be a simplicial complex in $\mathbb{R}^d$ and $\tilde{\mathrm{D}} \in \{\tilde{\chi}, \tilde{\beta}, \tilde{\rho}\}$. Suppose for some $S \subseteq \mathbb{S}^{d-1}$ the set $\tilde{\mathrm{D}}(K, S)$ is faithful. Then $|S| \geq d$, and this bound is tight.*

Proof. No vertex in K can be described using fewer than d coordinates, so a set of descriptors of type $\tilde{\mathrm{D}}$ with cardinality less than d can never be faithful. To see that this bound is tight, when K is a single vertex, verbose descriptors generated by any d pairwise linearly independent directions form a faithful set. □

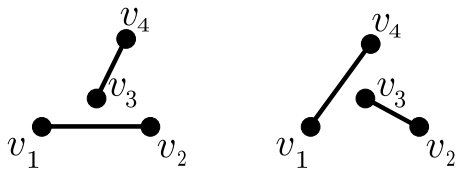
Next, we identify a family of simplicial complexes for which minimal faithful sets of verbose descriptors are at least linear in the number of vertices. We use $\alpha_{i,j}$ to denote the angle that vector $v_j - v_i$ makes with the x-axis, taking value in $[0, 2\pi)$. We first observe a consequence of a specific instance of the general phenomenon that a simplicial complex stratifies the sphere of directions based on vertex order [7, 22].

Observation 1 *Suppose a simplicial complex in $\mathbb{R}^2$ contains an isolated edge $[v_1, v_2]$. Then a birth event occurs height $s \cdot v_1$ in $\tilde{\rho}(K, s)$ for all s the interval $I =$*

$(\alpha_{1,2}-\pi/2, \alpha_{1,2}+\pi/2)$ *of* $\mathbb{S}^1$ *(all* s *so that* $s\cdot v_1 > s\cdot v_2$*) and as an instantaneous event for all* $s \in I^C$*.*

We now construct the building block that forms the complexes used in our bound.

Construction 1 (Clothespin Motif) *Let* K *be a simplicial complex in* $\mathbb{R}^2$ *with a vertex set* $\{v_1, v_2, v_3, v_4\}$ *such 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 Figure 5a.*



(a) K (b) K'

Figure 5: The two simplicial complexes of Lemma 10.

Construction 1 was built specifically for the following necessary condition, the proof of which may be found in Appendix D.3.

Lemma 10 (Clothespin Representability) *Let* K *be as in Construction 1, and suppose that* $\tilde{\rho}(K, S)$ *is faithful. Then, we have at least one direction* $s \in S$ *such that the angle formed between* s *and the* x*-axis lies in the region*

$$W = [\alpha_{3,2} - \pi/2, \alpha_{3,4} - \pi/2] \cup [\alpha_{3,2} + \pi/2, \alpha_{3,4} + \pi/2].$$

We call W, the intervals of directions in $\mathbb{S}^1$ for which corresponding verbose descriptors can distinguish K from K' a clothespin's *region of observability* (similar to observability for concise Euler characteristic functions in [7, 11]). Crucially, W is defined by $\measuredangle v_2 v_3 v_4$, so we have the following.

Remark 2 (W **Can be Arbitrarily Small)** *As the angle* $\measuredangle v_2 v_3 v_4$ *approaches zero, the region of observability from Lemma 10 also approaches zero.*

We use Remark 2 to piece together clothespins so their regions of observability do not overlap.

Construction 2 (Clothespins on a Clothesline) *Let* $K^{(m)}$ *be a simplicial complex in* $\mathbb{R}^2$ *formed by* m *copies of Construction 1 (*m *clothespin motifs) such that the regions of observability for each clothespin do not intersect. This is possible for any* m *by Remark 2.*

See Figure 6. Construction 2 implies a lower bound on the worst-case size of faithful sets of verbose persistence diagrams.

Lemma 11 (Augmented Persistence Diagram Complexity) *Let* $K^{(m)}$ *be as in Construction 2 and suppose* $\tilde{\rho}(K^{(m)}, S)$ *is a faithful set. Then* S *contains at least one direction in each of the* m *regions of observability, so* $|S| \geq m = n_0/4$*. Thus,* $|S|$ *is* $\Omega(n_0)$*.*

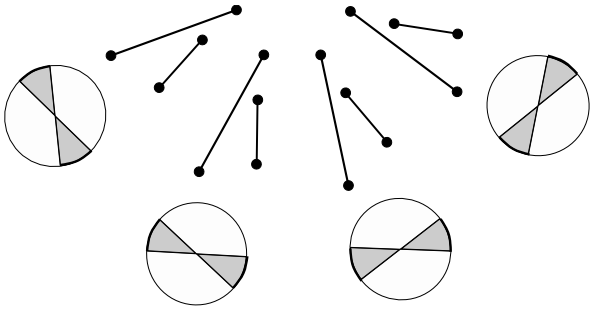


Figure 6: An example of $K^{(m)}$ for $m = 4$. Regions of observability are shown below each clothespin. By construction, each of these regions of $\mathbb{S}^1$ are disjoint.

By Theorem 6, Lemma 11 implies the following:

Theorem 12 (Lower Bound for Worst-Case Verbose Descriptor Complexity) *Let* $\tilde{\mathrm{D}} \in \{\tilde{\chi}, \tilde{\beta}, \tilde{\rho}\}$*. Then the worst-case cardinality of a minimal descriptor set of type* $\tilde{\mathrm{D}}$ *is* $\Omega(n_0)$*.*

7 Discussion

We provide a framework for comparing general topological descriptor types by their ability to efficiently represent simplicial complexes. The tools developed here are a first step towards more theoretical justifications of choosing a particular descriptor type for use in applications.

We focus on the descriptors that are particularly relevant to applications and related work, and give a partial order on this set of six descriptors, including the strict inequality, $[\tilde{\beta}] \prec [\tilde{\rho}]$. We also identify tight lower bounds for descriptor types, as well as bounds for worst-case complexity of sizes of faithful sets. Because faithful sets of concise descriptors require many perpendicular directions to each maximal simplex, a huge hiderance in practice, we believe applications research may benefit from the use of verbose descriptors rather than the current standard of concise descriptors. We are investigating other descriptor types in this framework as well, including merge trees.

Perhaps the strength classes $[\chi], [\beta]$, and $[\rho]$ intuitively feel as though they should be related by strict inequalities. However, this issue is nuanced. Lemma 15 (Appendix C) shows the impact that general position assumptions have on relations in this set. But even with general position, the seemingly advantageous "extra" information of homology compared to, e.g., Euler characteristic may no longer be so useful when we require tight envelopes around each maximal simplex. That is, once we have all the (many) required directions, we have already carved out the space filled by the complex, and already know quite a lot simply from the presence of events. Non-equality/equality of concise descriptors remains an area active of research.

References

[1] E. J. Amézquita, M. Y. Quigley, T. Ophelders, J. B. Landis, D. Koenig, E. Munch, and D. H. Chitwood.

Measuring hidden phenotype: Quantifying the shape of barley seeds using the Euler characteristic transform. *in silico Plants*, 4(1):diab033, 2022.

[2] L. Batakci, A. Branson, B. Castillo, C. Todd, E. W. Chambers, and E. Munch. Comparing embedded graphs using average branching distance. *Involve, a Journal of Mathematics*, 16(3):365–388, 2023.

[3] R. L. Belton, B. T. Fasy, R. Mertz, S. Micka, D. L. Millman, D. Salinas, A. Schenfisch, J. Schupbach, and L. Williams. Reconstructing embedded graphs from persistence diagrams. *Computational Geometry: Theory and Applications*, 90, October 2020.

[4] P. Bendich, J. S. Marron, E. Miller, A. Pieloch, and S. Skwerer. Persistent homology analysis of brain artery trees. *The Annals of Applied Statistics*, 10(1):198, 2016.

[5] W. Chachólski, B. Giunti, A. Jin, and C. Landi. Decomposing filtered chain complexes: Geometry behind barcoding algorithms. *Computational Geometry*, 109:101938, 2023.

[6] L. Crawford, A. Monod, A. X. Chen, S. Mukherjee, and R. Rabadán. Predicting clinical outcomes in glioblastoma: an application of topological and functional data analysis. *Journal of the American Statistical Association*, 115(531):1139–1150, 2020.

[7] J. Curry, S. Mukherjee, and K. 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.

[8] J. Darrotto. Convex optimization and Euclidean distance geometry. *Palo Alto: Meboo*, 2013.

[9] H. Edelsbrunner and J. Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2010.

[10] B. T. Fasy, S. Micka, D. L. Millman, A. Schenfisch, and L. Williams. Challenges in reconstructing shapes from Euler characteristic curves, 2018. arXiv:1811.11337.

[11] B. T. Fasy, S. Micka, D. L. Millman, A. Schenfisch, and L. Williams. Challenges in reconstructing shapes from Euler characteristic curves. *Fall Workshop Computational Geometry*, 2018.

[12] B. T. Fasy, S. Micka, D. L. Millman, A. Schenfisch, and L. Williams. A faithful discretization of the verbose persistent homology transform, 2019. arXiv:1912.12759.

[13] B. T. Fasy, S. Micka, D. L. Millman, A. Schenfisch, and L. Williams. Efficient graph reconstruction and representation using augmented persistence diagrams. *Proceedings of the 34th Annual Canadian Conference on Computational Geometry*, 2022.

[14] B. T. Fasy and A. Patel. Persistent homology transform cosheaf, 2022.

[15] P. Frosini and C. Landi. Persistent Betti numbers for a noise tolerant shape-based approach to image retrieval. *Pattern Recognition Letters*, 34(8):863–872, 2013.

[16] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov. Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences*, 112(44):13455–13460, 2015.

[17] A. Hatcher. Algebraic topology, Cambridge Univ. *Press, Cambridge*, 2002.

[18] T. Jech. Set theory. *Journal of Symbolic Logic*, pages 876–77, 1981.

[19] Q. Jiang, S. Kurtek, and T. Needham. The weighted Euler curve transform for shape and image analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 844–845, 2020.

[20] P. Lawson, J. Schupbach, B. T. Fasy, and J. W. Sheppard. Persistent homology for the automatic classification of prostate cancer aggressiveness in histopathology images. In *Medical Imaging 2019: Digital Pathology*, volume 10956, page 109560G. International Society for Optics and Photonics, 2019.

[21] Y. Lee, S. D. Barthel, P. Dłotko, S. M. Moosavi, K. Hess, and B. Smit. Quantifying similarity of pore-geometry in nanoporous materials. *Nature Communications*, 8:15396, 2017.

[22] J. Leygonie, S. Oudot, and U. Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021.

[23] J. Li, L. Yang, Y. He, and O. Fukuda. Classification of hand movements based on EMG signals using topological features. *International Journal of Advanced Computer Science and Applications*, 14(4), 2023.

[24] L. Marsh and D. Beers. Stability and inference of the Euler characteristic transform, 2023. arXiv:2303.13200.

[25] A. McCleary and A. Patel. Edit distance and persistence diagrams over lattices. *SIAM Journal on Applied Algebra and Geometry*, 6(2):134–155, 2022.

[26] A. McCleary and A. Patel. Edit distance and persistence diagrams over lattices. *SIAM Journal on Applied Algebra and Geometry*, 6(2):134–155, 2022.

[27] F. Mémoli and L. Zhou. Stability of filtered chain complexes, 2022. arXiv:2208.11770.

[28] F. Mémoli and L. Zhou. Ephemeral persistence features and the stability of filtered chain complexes. In *39th International Symposium on Computational Geometry (SoCG 2023)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2023.

[29] K. Meng, J. Wang, L. Crawford, and A. Eloyan. Randomness and statistical inference of shapes via the smooth Euler characteristic transform, 2022. arXiv:2204.12699.

[30] S. Micka. *Algorithms to Find Topological Descriptors for Shape Reconstruction and How to Search Them*. PhD thesis, Montana State University, 2020.

[31] K. V. Nadimpalli, A. Chattopadhyay, and B. Rieck. Euler characteristic transform based topological loss for reconstructing 3D images from single 2D slices. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 571–579, 2023.

[32] M. J. Panik. *Fundamentals of convex analysis: duality, separation, representation, and resolution*, volume 24. Springer Science & Business Media, 2013.

[33] P. Pranav, H. Edelsbrunner, R. Van de Weygaert, G. Vegter, M. Kerber, B. J. Jones, and M. Wintraecken. The topology of the cosmic web in terms

of persistent Betti numbers. *Monthly Notices of the Royal Astronomical Society*, 465(4):4281–4310, 2017.

[34] E. Richardson and M. Werman. Efficient classification using the euler characteristic. *Pattern Recognition Letters*, 49:99–106, 2014.

[35] A. H. Rizvi, P. G. Camara, E. K. Kandror, T. J. Roberts, I. Schieren, T. Maniatis, and R. Rabadan. Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nature Biotechnology*, 35(6):551, 2017.

[36] A. Saadat-Yazdi, R. Andreeva, and R. Sarkar. Topological detection of alzheimer's disease using Betti curves. In *Interpretability of Machine Intelligence in Medical Image Computing, and TDA and Its Applications for Medical Data.*, pages 119–128. Springer, 2021.

[37] G. Singh, F. Mémoli, and G. E. Carlsson. Topological methods for the analysis of high dimensional data sets and 3D object recognition. *SPBG*, 91:100, 2007.

[38] K. Turner, S. Mukherjee, and D. M. Boyer. Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA*, 3(4):310–344, 2014.

[39] S. Tymochko, E. Munch, J. Dunion, K. Corbosiero, and R. Torn. Using persistent homology to quantify a diurnal cycle in hurricanes. *Pattern Recognition Letters*, 2020.

[40] M. Usher and J. Zhang. Persistent homology and Floer–Novikov theory. *Geometry & Topology*, 20(6):3333–3430, 2016.

[41] R. Van De Weygaert, E. Platen, G. Vegter, B. Eldering, and N. Kruithof. Alpha shape topology of the cosmic web. In *2010 International Symposium on Voronoi Diagrams in Science and Engineering*, pages 224–234. IEEE, 2010.

[42] Y. Wang, H. Ombao, and M. K. Chung. Statistical persistent homology of brain signals. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1125–1129, 2019.

[43] G. Wilding, K. Nevenzeel, R. van de Weygaert, G. Vegter, P. Pranav, B. J. Jones, K. Efstathiou, and J. Feldbrugge. Persistent homology of the cosmic web–I. hierarchical topology in λCDM cosmologies. *Monthly Notices of the Royal Astronomical Society*, 507(2):2968–2990, 2021.

[44] L. Zhou. *Beyond Persistent Homology: More Discriminative Persistent Invariants*. PhD thesis, The Ohio State University, 2023.

[45] Z. Zhou, Y. Huang, L. Wang, and T. Tan. Exploring generalized shape analysis by topological representations. *Pattern Recognition Letters*, 87:177–185, 2017.

A Example Filtration with Six Descriptor Types

We consider the simplicial complex on four vertices given in Figure 7(a). In the e_1 direction we see four distinct heights of vertices, a, b, c, and d.

First, we describe what happens in $\rho(K)$ and $\tilde{\rho}(K)$. At height a, and then again at height b, we see connected components born. At height c, the homology of the sublevel set does not change, so no change is recorded in $\rho(K)$. However, a corresponding index filtration sees the connected component corresponding to first adding the vertex at c, which then immediately dies as we include the edge at height c. Thus, in $\tilde{\rho}(K)$, we have the point (c, c). For similar reasons, we see the point (d, d) in $\tilde{\rho}(K)$. Also at height d, our two connected components merge into a single connected component. It is a standard convention to choose the eldest component to survive, so we have the point (b, d) in both diagrams. We also see a cycle appear at height d, giving us the point (d, ∞) in both diagrams. Finally, since the connected component born at height a did not die, we have the point (a, ∞) in both diagrams.

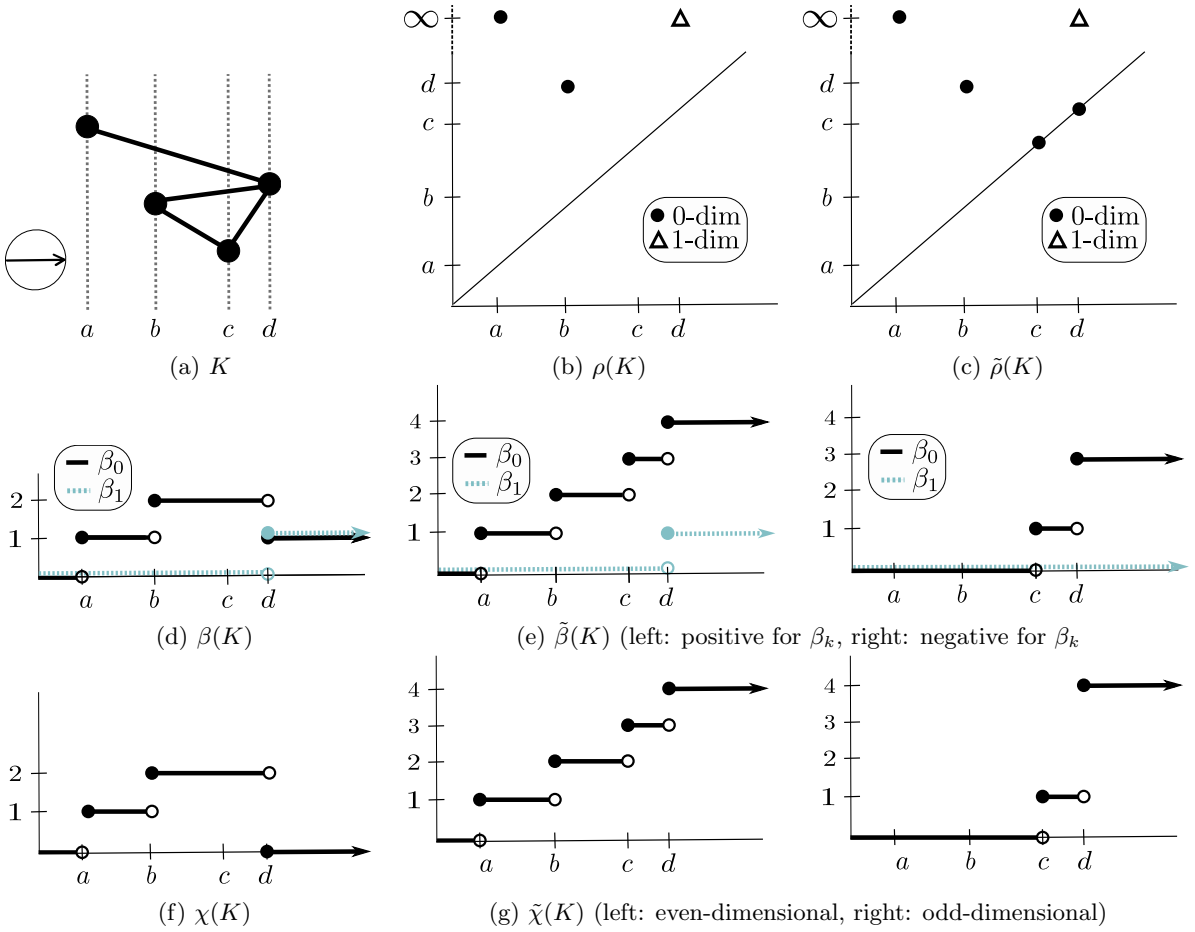
Next, we describe what happens in $\beta(K)$ and $\tilde{\beta}(K)$. At the height a, only the Betti number in dimension zero changes, going from zero to one. Since the inclusion of this vertex increased Betti zero, we count the vertex as positive in $\tilde{\beta}_0(K)$. At the height b, again, only Betti zero changes, going from one to two, and we also count the corresponding vertex as positive for $\tilde{\beta}_0(K)$. At the height of c, no Betti number changes, and thus there is no event in $\beta(K)$. However, in a correpsonding index filtration, the inclusion of the vertex at height c increases β_0 by one, and the inclusion of the edge then reduces β_0 by one. This is recorded in $\tilde{\beta}(K)$ at c as an additional positive simplex (going from two total to three total), and our first negative simplex. We see similar behavior in dimension zero at the height of d. At height d we also see β_1 go from zero to one, which is recorded in the concise Betti function. In a corresponding index filtration, the inclusion of the second edge at height d increases β_1, and is thus recorded as positive.

Finally, we describe what happens in $\chi(K)$ and $\tilde{\chi}(K)$. At both heights a and b, the Euler characteristic of the sublevel sets increases by one. Since this is due to inclusions of vertices, which are even-dimensional simplices, both of these increases are recorded as even-dimensional in $\tilde{\chi}(K)$. At c, the Euler characteristic remains the same, so no change occurs in $\chi(K)$. In a corresponding index filtration, we see the vertex (an even-dimensional simplex) and an edge (odd-dimensional simplices), which are recorded in $\tilde{\chi}(K)$. Finally, the Euler characteristic at d changes from two to zero, which is recorded directly in $\chi(K)$. An index filtration witnesses the appearance of one even and two odd simplices at height d, and this is recorded in $\tilde{\chi}(K)$.

B Faithfulness and Set Intersections

In this appendix, we explore reframing the definition of faithful in terms of set intersection. In Definition 2, we identify $\mathrm{D}(K, P)$ as faithful topological transform if, for any simplicial complex L, we have $\mathrm{D}(L, P) = \mathrm{D}(K, P)$ if and only if $K = L$. To unpack the equality $\mathrm{D}(L, P) = \mathrm{D}(K, P)$, recall

$$\mathrm{D}(K, P) := \{(p, \mathrm{D}(K, p))\}_{p \in P}.$$



(a) K (b) $\rho(K)$ (c) $\tilde{\rho}(K)$

(d) $\beta(K)$ (e) $\tilde{\beta}(K)$ (left: positive for β_k, right: negative for β_k

(f) $\chi(K)$ (g) $\tilde{\chi}(K)$ (left: even-dimensional, right: odd-dimensional)

Figure 7: Six descriptors corresponding to the lower-star filtration in the direction indicated by the arrow of the simplicial complex in (a).

Thus, $\mathrm{D}(L,P) = \mathrm{D}(K,P)$ if and only if for all $p \in P$, we have $\mathrm{D}(K,p) = \mathrm{D}(L,p)$. Then, $\mathrm{D}(K,P)$ is faithful if and only if

$$\bigcap_{p\in P} \{K' \subset \mathbb{R}^d \mid \mathrm{D}(K',p) = \mathrm{D}(K,p)\} = \{K\}. \quad (1)$$

From this perspective, we prove the following lemma providing a sufficient condition for finite faithful sets.

Lemma 13 (Sufficient Conditions for Finite Faithful Set) *Let K be a simplicial complex immersed in $\mathbb{R}^d$ and let D be a type of topological descriptor that can faithfully represent K. Suppose there exists a finite set of descriptors of type D that is faithful for K_0. Then, there exists a finite faithful set of descriptors of type D for K.*

Proof. Let P be a parameter set such that $\mathrm{D}(K,P)$ that is faithful for K, and let P_0 be a finite parameter set such that $\mathrm{D}(K,P_0)$ that is faithful for K_0.

Now, let B be the set of simplicial complexes that are indistinguishable from K using only parameter set P_0; that is,

$$B := \bigcap_{p\in P_0} \{K' \subset \mathbb{R}^d \mid \mathrm{D}(K',p) = \mathrm{D}(K,p)\} \quad (2)$$

Since P_0 is faithful for K_0, we know 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 $|K_0|$ is finite, there are a finite number of simplicial complexes we can build over this set of vertices.

If $B = \{K\}$, we are done. Otherwise, since $D(K,P)$ is faithful for K, for each $L \neq K$ in B, there exists a direction $p_L \in P$ such that $\mathrm{D}(L,p_L) \neq \mathrm{D}(K,p_L)$. Let $P^* = P \cup \{p_L \mid L \in B\}$. Then, $\mathrm{D}(K,P^*)$ faithfully represents K. Furthermore, since P and B are finite, we also know that P^* is finite. □

C Zoo of Other Descriptor Types

In Section 2, we adopt a general definition of topological descriptor (Definition 1). In this appendix, we explore non-standard topological descriptors, and corresponding scenarios that may arise as a result of this generality. The descriptors presented here are not intended to be taken as anything that would necessarily make sense to use in practice, but rather, as a sort of zoo of examples to get a quick glance at the mathematical extremes and properties of the space of strength classes of topological descriptors.

First, we give an example of two distinct descriptor types that are in the same equivalence class of strength.

Example 1 *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 generally $\rho(s) \neq \rho(-s)$ (as multisets), 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, then it has no faithful set of type $-\rho$. Thus, $[\rho] = [-\rho]$.

Next, we observe that many examples of topological descriptors are not capable of faithfully representing most simplicial complexes, such as the following.

Descriptor Type 1 (First Vertex) *Consider a descriptor D_V that returns (1) the coordinates of the first vertex (or vertices) encountered and (2) the cardinality of the vertex set, but no other information.*

If the filtrations are directional filtrations, then this descriptor is only faithful for convex point clouds. Any set of vertices that defines the corners of a convex region can be faithfully represented by this Descriptor Type 1. However, since no vertex interior to the convex hull nor any higher dimensional simplices are witnessed by any direction, 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.

Descriptor Type 2 (Trivial) *Consider the trivial descriptor type D_0 that returns zero for all sublevel sets in a lower-star filtration.*

Although this trivial descriptor type is an invariant of any filtration, it can not faithfully represent any simplicial complex. Thus, $\Gamma(K, D_0) = \aleph_\top$ for all K. And so, in the space of all topological descriptors, Descriptor Type 3 is in the minimum strength class. We can (also trivially) construct a descriptor type that is in the maximum strength class.

Descriptor Type 3 (Filtration-Returning) *Consider the descriptor type D_{Filt} that returns the input filtration.*

Thus, a single descriptor of this type is always faithful for a simplicial complex.

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

Descriptor Type 4 (Indicator) *Let K be a simplicial complex immersed in $\mathbb{R}^d$ let $D_{\mathbb{R}}$ be a descriptor type parameterized by $\mathbb{R}^d$ that is constant over a filtration and is defined by*

$$D(K, x) = \begin{cases} 1 & \text{if } x \in |K| \\ 0 & \text{else.} \end{cases}$$

Note that then, $D(K, \mathbb{R}^d)$ is the (only) minimum faithful set for K, and so $\Gamma(K, D_{\mathbb{R}}) = \aleph_1$ for all K. Thus, the (minimal) strength class of Descriptor Type 4 is greater than the strength class of the trivial descriptor in Descriptor Type 3, and there are no strength equivalence classes between them.

We now know the space of strength classes of topological descriptors has a minimum and maximum, and we have identified a second smallest descriptor type; is it a total order? The following example shows that it is not; there do indeed exist incomparable descriptor types.

Lemma 14 (Incomparable Strength Classes) *There exist incomparable strength classes of topological descriptor types.*

Proof. Let D_V denote Descriptor Type 1. That is, given a direction s, D returns: (1) the coordinates of the lowest vertex (or vertices) in direction s, and (2) the cardinality of the vertex set. We compare D_V with verbose persistence diagrams. Let $v_1 = (0, 0)$ and $v_2 = (0, 1)$.

First, consider the simplicial complex $K = \{v_1\}$. Then, regardless of direction, a single descriptor of is faithful for K. However, since K is in $\mathbb{R}^2$, any faithful set of augmented persistence diagrams 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_V 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 augmented persistence diagrams suffice to form a faithful set for K'; for example, using the standard basis vectors $\{e_1, e_2\}$ as the set. Thus, if $\mathrm{D}_V(K, S_{\mathrm{D}_V})$ and $\tilde{\rho}(K, S_{\tilde{\rho}})$ are both minimal faithful sets, we see that $|S_{\mathrm{D}_V}| < |S_{\tilde{\rho}}|$ but $|S_{\mathrm{D}_V}| > |S_{\tilde{\rho}}|$. Thus, although we have shown $[\mathrm{D}_V] \neq [\tilde{\rho}]$, they are incomparable. □

Finally, we give a lemma that shows an impact of not adopting the general position assumption for the vertices. Specifically, the general position assumption we use throughout this paper is:

General Position 1 *A simplicial complex K immersed in $\mathbb{R}^d$ is in* general position *if, for all subsets $V \subseteq K_0$ with $|V| \leq d + 1$ is affinely independent*

Removing this assumption, we obtain:

Lemma 15 (Concise Equality) *Without Assumption 1, the strength equivalence classes of the three concise topological descriptor types from Section 3 are all equal.*

Proof. Let $D \in \{\chi, \beta, \rho\}$. We must consider faithful sets of such descriptors for an arbitrary simplicial complex K immersed in $\mathbb{R}^d$ (that may not be in general posistion). The argument differs depending on if K is a vertex set, or contains at least one edge; we consider each case.

First, suppose $n_1 = 0$. Then, K has no edges and is a vertex set, meaning each vertex is a maximal simplex of K. Then, by Corollary 1, faithful sets of type D must include descriptors from at least $d + 1$ directions. By Lemma 7, the envelopes of each vertex must

be zero-dimensional. Since the only zero-dimensional convex sets are singleton points, the envelope of each vertex contains that vertex and nothing else.

Let S be a set of $d+1$ directions such that the envelope of each vertex is zero-dimensional (note that such a set exists, for example, the standard basis directions (e_i), and the negative diagonal direction, $-1/\sqrt{d}(1,1,\ldots,1)$). Consider a single direction, $s \in S$ and $a \leq b \in \mathbb{R}^d$. If no event occurs in $\mathrm{D}(s)$ between heights a and b, we know no connected component of K has its lowest vertex (or vertices) with respect to s in the range from a to b. Then from $\mathrm{D}(s)$, we identify that K has n_0 connected components, and we know their starting heights with respect to the s direction. In other words, we know these connected components are contained in the (closed) upper half-spaces defined by s.

Each additional direction in S gives us more information about these n_0 connected components, and, just like s, each additional direction provides an additional upper half-space in which we know the connected component is contained; each connected component must lie in the intersection of these half-spaces. Since we use a total of $d+1$ directions, chosen so that the envelope of each vertex, $\mathcal{E}_v^S$ (the intersection of half-spaces), is zero-dimensional, we conclude each connected component is zero-dimensional.

That is, we know the exact location of each vertex by identifying its envelope. Thus, minimal faithful sets of type D have cardinality of exactly $d+1$, meaning the infimums considered in Definition 7 are $\Gamma(K,D) = d+1$ for such K.

Next, suppose $n_1 > 1$. We show no set of descriptors of type D can faithfully represent K. Let τ be an edge in K and construct another complex L by starting with K and taking the barycentric subdivision of τ and all simplices containing τ. Then, since $|K| = |L|$ the Euler characteristics/Betti numbers/homology throughout the filtrations of K or L agree. That is, for every direction s, $\mathrm{D}(K,s) = \mathrm{D}(L,s)$. Since this is true for every s, the descriptor type D is incapable of forming a faithful set for K, meaning the infimums considered in Definition 7 are $\Gamma(K,D) = \aleph_\top$ for such K.

We have shown that, for each K without a general position assumption, we have $\Gamma(K,\chi) = \Gamma(K,\beta) = \Gamma(K,\rho)$. Thus, when removing the general position assumption, we find $[\chi] = [\beta] = [\rho]$. □

Fortunately, as shown in Section 5, the relations among concise descriptors becomes more interesting when we assume general position. This also has the benefit of reflecting the general position assumptions that are often taken in practical applications.

D Omitted Proofs and Lemmas

In this appendix, we provide proofs that were omitted from the main text.

D.1 Proof from Section 4

The following lemma is stated, but not proven, in Section 4.

Lemma 1 *Let A and B be two topological descriptor types. If B is reducible to A, we have $[A] \preceq [B]$.*

Proof. Let K be a simplicial complex. Define the sets $W_A := \{P \text{ s.t. } A(K,P) \text{ is faithful}\}$ and $W_B := \{P \text{ s.t. } B(K,P) \text{ is faithful}\}$. For each $P \in W_A$, by definition, $A(K,P)$ is faithful. Since B is reducible to A, this also means $B(K,P)$ is faithful, and so P is also in W_B. Hence, $W_A \subseteq W_B$. Hence, $\inf_{P\in W_A} |P| \geq \inf_{P\in W_B} |P|$. Note that these are exactly the infimums in Definition 7, and so, we have $[A] \preceq [B]$. □

The following lemmas are referenced, but not explicitly stated, in Section 4.

Lemma 16 *The relation $=$ is an equivalence relation, and the relation $\preceq$ is well-defined on sets of strength equivalence classes.*

Proof. When we compare infimums in Definition 7, we compare values in $\mathbb{N} \cup \{\aleph_0, \aleph_1, \aleph_\top\}$. The relation $\leq$ on values in this set is reflexive, antisymmetric, and transitive. The relation $=$ on this set is reflexive, symmetric, and transitive. The result follows. □

D.2 Proofs from Section 5

We prove two lemmas that were originally stated in Section 5.

Lemma 2 $[\chi] \preceq [\beta] \preceq [\rho]$ *and* $[\tilde{\chi}] \preceq [\tilde{\beta}] \preceq [\tilde{\rho}]$.

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 from $\beta(s)$. A nearly identical argument shows the relationship between verbose versions of these descriptors. □

The reductions described above are well-known, and are observed in other work; for example [7, Prop 4.13] points out the reduction from an Euler characteristic function to a persistence diagram.

Lemma 3 $[\chi] \preceq [\tilde{\chi}]$, $[\beta] \preceq [\tilde{\beta}]$, *and* $[\rho] \preceq [\tilde{\rho}]$.

Proof. Each verbose descriptor has a clear reduction to its concise counterpart. A verbose persistence diagram becomes a concise persistence diagram by removing all on-diagonal points. Verbose Betti functions and verbose Euler characteristic functions become concise if we subtract their second coordinates from their first coordinates. Then by Lemma 1, we have the desired relations. □

Lemma 4 *For* $\mathrm{D} \in \{\chi, \beta, \rho\}$ *and* $\tilde{\mathrm{D}} \in \{\tilde{\chi}, \tilde{\beta}, \tilde{\rho}\}$, *we have* $[\mathrm{D}] \neq [\tilde{\mathrm{D}}]$.

Proof. To show inequality of strength classes, we find a simplicial complex for which minimum faithful sets of type D and $\tilde{\mathrm{D}}$ have different cardinalities. Let K be the simplicial complex that is a single edge in $\mathbb{R}^2$ with vertex coordinates $(1,1)$ and $(1,2)$. See Figure 8 for this complex, and an illustration for the specific case $\tilde{\mathrm{D}} = \tilde{\rho}$. In the direction $e_1 = (1,0)$, if $\tilde{\mathrm{D}} = \tilde{\rho}$, we see an instantaneous birth/death and an infinite birth in degree zero. If $\tilde{\mathrm{D}} = \tilde{\beta}$, we see two positive simplices and one negative simplex for Betti zero. If $\tilde{\mathrm{D}} = \tilde{\chi}$, we see two even simplices and one odd simplex. This all occurs at height 1, and there are no other events, which, is only explainable by the presence of a single edge. From $\tilde{\mathrm{D}}(e_2)$, we see a non-instantaneous and instantaneous event at heights 1 and 2, respectively, which give us the y-coordinates of our two vertices. Then K is the only complex that could have generated both $\tilde{\mathrm{D}}(e_1)$ and $\tilde{\mathrm{D}}(e_2)$, i.e., the set $\tilde{\mathrm{D}}(K, \{e_1, e_2\})$ is faithful.

Next, consider the descriptor type D. For any $s \in \mathbb{S}^1$, $\mathrm{D}(s)$ contains exactly one event; if the lowest vertex of K with respect to s has height a in direction s, then $\mathrm{D}(s)$ records a change in homology/Betti number/Euler characteristic at height a and records no other changes. Thus, $\mathrm{D}(s)$ can only give us information about one coordinate of the vertex set of K at a time, corresponding to whichever vertex is lowest in direction s. However, K has three relevant coordinates; namely, $x = 1$, $y = 1$, and $y = 2$ meaning it is not possible for any faithful set of type D to have size less than three. Thus, since $3 \neq 2$, we have shown $[\mathrm{D}] \neq [\tilde{\mathrm{D}}]$, as desired. □

The specific inequality $[\chi] \neq [\tilde{\rho}]$ is also implied by [30, Thm. 10].

D.3 Proofs from Section 6

We provide proofs of Lemmas 7 and 10.

Lemma 7 (Envelopes for Faithful Concise Sets) *Let K be a simplicial complex immersed in $\mathbb{R}^d$, let $D \in \{\chi, \beta, \rho\}$, and let $S \subseteq \mathbb{S}^{d-1}$ so that $\mathrm{D}(K, S)$ is faithful. Then, for any maximal simplex σ in K, the dimension of $\mathcal{E}_\sigma$ equals the dimension of σ.*

Proof. Let k be the dimension of σ, and let c be the dimension of $\mathcal{E}_\sigma$. First, we observe that since σ is contained in $\mathcal{E}_\sigma$, we must have $k \leq c$. The claim is trivial when $k = d$, so we proceed with the case $k < d$ and assume, by way of contradiction, that $k < c$.

We claim that in this case, 1) at every interior point of σ, there is a vector normal to σ that ends in the interior of $\mathcal{E}_\sigma$, and 2) letting p denote the endpoint of such a vector, 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, $\mathcal{E}_\sigma$ would be k-dimensional. 2) is true since each halfspace defining $\mathcal{E}_\sigma$ contains the lowest vertex (or vertices) of σ with respect to the corresponding direction. Thus, since p is in the interior of $\mathcal{E}_\sigma$, it must be higher than this lowest vertex (or vertices) with respect to the corresponding direction.

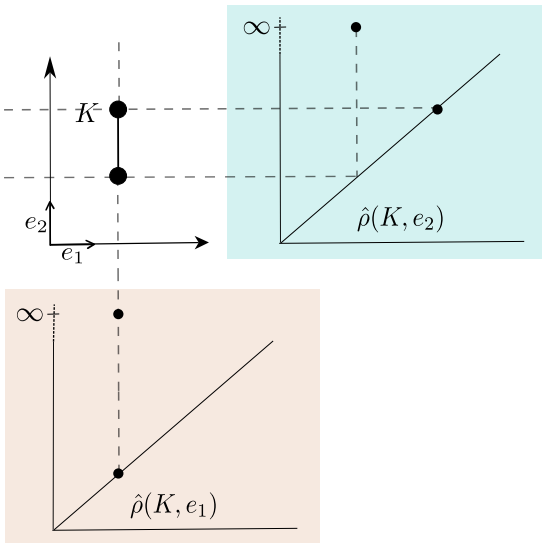


Figure 8: The simplicial complex considered in the proof of Lemma 4 as well as the verbose diagrams in directions e_1 and e_2. Note that $\tilde{\rho}(K, \{e_1, e_2\})$ is indeed a faithful set, since we can recover the coordinates of both vertices ($\tilde{\rho}(K, e_1)$ tells us the x-coordinates and $\tilde{\rho}(K, e_2)$ tells us the y-coordinates), as well as determine there is only a single edge present (there is only one instantaneous zero-dimensional point in each verbose diagram). The concise versions of these diagrams do not have on-diagonal points, and each only contain a single point at ∞. This is true for concise diagrams corresponding to any direction.

Now consider the simplicial complex K', defined as having all the simplices of K in addition to the simplex formed by taking the geometric join of p and 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 higher than the lowest vertex of σ with respect to any direction $s \in S$, 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, we have shown $\mathrm{D}(K', S) = \mathrm{D}(K, S)$. This contradicts the assumption that $\mathrm{D}(K, S)$ is faithful, so we must have $k = c$. □

The next statement we wish to prove is Lemma 10. We first establish the following lemma.

Lemma 17 *Consider a pair of nested triangles as in Figure 9. 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,

$$\begin{aligned}\theta + (\phi - B) + B + (\psi - C) + C &= \pi \\ A + \theta + (\phi - B) + B + (\psi - C) + C &= A + \pi \\ (A + B + C) + \theta + (\phi - B) + (\psi - C) &= A + \pi \\ \pi + \theta + (\phi - B) + (\psi - C) &= A + \pi \\ \theta + (\phi - B) + (\psi - C) &= A.\end{aligned}$$

All the terms in the last line are positive, meaning A is larger than θ, $\phi - B$, and $\psi - C$. □

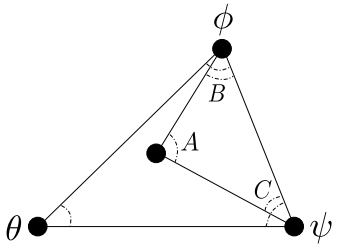


Figure 9: Nested triangles as discussed in Lemma 17

Now, we are armed with the tools needed to prove the following lemma.

Lemma 10 (Clothespin Representability) *Let K be as in Construction 1, and suppose that $\tilde{\rho}(K, S)$ is faithful. Then, we have at least one direction $s \in S$ such that the angle formed between s and the x-axis lies in the region*

$$W = [\alpha_{3,2} - \pi/2, \alpha_{3,4} - \pi/2] \cup [\alpha_{3,2} + \pi/2, \alpha_{3,4} + \pi/2].$$

Proof. Let K' be a simplicial complex immersed in $\mathbb{R}^2$ with the same vertex set as K, but with edges $[v_1, v_4]$ and $[v_2, v_3]$ (see Figure 5b). Recall that, since $\tilde{\rho}(K, S)$ is faithful, by definition, the set S must contain some direction s so that $\tilde{\rho}(K, s) \neq \tilde{\rho}(K', s)$.

Each vertex corresponds to either a birth event or an instantaneous event depending on the direction of filtration. 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 10 shows these regions, and in what follows, we define them precisely.

First, consider v_1. By Observation 1, $v_1 \in K$ corresponds to a birth event for all directions in the interval $B = (\alpha_{1,2} - \pi/2, \alpha_{1,2} + \pi/2)$ and $v_1 \in K'$ corresponds to a birth event for all directions in the interval $B' = (\alpha_{1,4} - \pi/2, \alpha_{1,4} + \pi/2)$. 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 $\tilde{\rho}(K, s) \neq \tilde{\rho}(K', s)$.

Using this same notation, identify the wedge shaped region R_i for vertex $i \in [2, 3, 4]$ such that any direction from R_i generates verbose persistence diagrams that have different event types at the height of vertex v_i when filtering over K versus K'. Similar arguments for $i \in [2, 3, 4]$ give us the complete list;

$$\begin{aligned}R_1 &= (\alpha_{1,2} - \pi/2, \alpha_{1,4} - \pi/2] \cup [\alpha_{1,2} + \pi/2, \alpha_{1,4} + \pi/2) \\ R_2 &= (\alpha_{2,3} - \pi/2, \alpha_{2,1} - \pi/2] \cup [\alpha_{2,3} + \pi/2, \alpha_{2,1} + \pi/2) \\ R_3 &= (\alpha_{3,2} - \pi/2, \alpha_{3,4} - \pi/2] \cup [\alpha_{3,2} + \pi/2, \alpha_{3,4} + \pi/2) \\ R_4 &= (\alpha_{1,4} - \pi/2, \alpha_{3,4} - \pi/2] \cup [\alpha_{1,4} + \pi/2, \alpha_{3,4} + \pi/2)\end{aligned}$$

Let $W = \cup_{i=1}^{4} R_i$. Then, for any $s \in W$, we have $\tilde{\rho}(K, s) \neq \tilde{\rho}(K', s)$, and for any $s_* \in W^C$, we have $\tilde{\rho}(K, s_*) = \tilde{\rho}(K', s_*)$.

Finally, we claim that W is the closure of R_3, denoted $\overline{R_3}$, i.e., exactly the region described in the lemma statement. This is a direct corollary to Lemma 17; 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 $\alpha_{2,3} \pm \pi/2$ and $\alpha_{3,4} \pm \pi/2$, the defining angles of R_3. Each of these four angles appears as an included endpoint for some R_i, so $R_1, R_2, R_4 \subseteq \overline{R_3} = W$ (Figure 10) and we have shown our claim. □

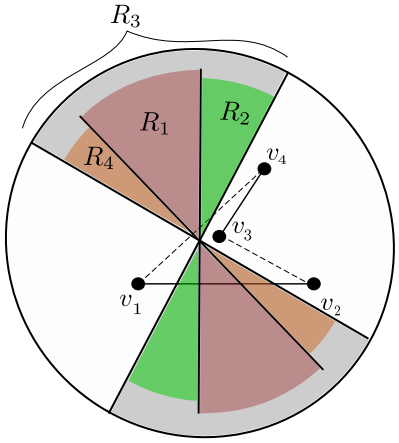


Figure 10: The regions described in the proof of Lemma 10, with additional shading in the interior of the sphere of directions to aid in visibility. K is shown as solid black edges and K' as dashed edges. For any lower-star filtration in a direction contained in R_i, the event at vertex v_i differs when 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'.