

Lower Bounding Minimal Faithful Sets of Verbose Persistence Diagrams

Brittany Terese Fasy^{1,2}, David L. Millman², and Anna Schenfisch³

1 Montana State University
brittany.fasy@montana.edu

2 Blocky Inc.
david@blocky.rocks

3 TU Eindhoven, Eindhoven
a.k.schenfisch@tue.nl

Abstract

An important aim of inverse problems in topological data analysis is to better understand sets of directional topological descriptors that uniquely correspond to an underlying shape; such sets are called *faithful* for the shape. Here, we specifically focus on sets of verbose persistence diagrams that arise from lower-star filtrations of geometric simplicial complexes. While explicit constructions of finite faithful sets in this setting exist in the literature, they do not come with any guarantees of optimality in terms of cardinality. To better understand faithful sets with low cardinality, we first establish a tight lower bound on the size of any faithful set. Then, we construct a family of simplicial complexes for which faithful sets must have size at least linear in the number of vertices.

1 Introduction

The persistent homology transform of a shape in Euclidean space was first explored in [15], and is the set of persistence diagrams corresponding to lower-star filtrations in every possible direction. Importantly, [15] establishes that this uncountably infinite set of persistence diagrams uniquely represents the underlying shape, i.e., it is *faithful* for the underlying shape. Since then, related theoretical work has focused on finding finite sets of persistence diagrams or other topological descriptors (such as Euler Characteristic functions or Betti functions) that are faithful [1, 3, 7, 14]. A key parameter in such work is whether or not the descriptors are assumed to be *verbose* or *concise*, i.e., if they contain information with a trivial lifespan. The relevance of verbose or concise descriptors was explored in [5, 13, 18], although with slightly different language.

In the full version of this work [8], we develop a framework for comparing the relative strengths of different topological descriptor types through the cardinality of faithful sets. Roughly speaking, if faithful sets of a particular topological descriptor type are always larger than faithful sets of another type, it is *weaker* than that other type. Thus, in such quantitative comparisons, it can be vital to understand minimum faithful sets. While explicitly identifying a minimum faithful set is a difficult problem in general, we are able to provide lower bounds on the cardinality of minimum faithful sets, both in general and in a worst-case construction. In what follows, we focus on the specific topological descriptor type of *verbose persistence diagrams*; we refer the reader to our full version to see how this and similar constructions apply to other common descriptor types.

40th European Workshop on Computational Geometry, Ioannina, Greece, March 13–15, 2024.
This is an extended abstract of a presentation given at EuroCG’24. It has been made public for the benefit of the community and should be considered a preprint rather than a formally reviewed paper. Thus, this work is expected to appear eventually in more final form at a conference with formal proceedings and/or in a journal.

2 Preliminary Considerations

We assume the reader has familiarity with ideas from topology, including homology and simplicial complexes. See [4, 9] for further details. We always take $\mathbb{N}$ to include zero. For a simplicial complex K , we use the notation K_i for its i -skeleton and n_i as the number of i -simplices. We always assume our simplicial complexes are geometric and finite. A *filter* of K is a map $f: K \rightarrow \mathbb{R}$ such that, for $\tau, \sigma \in K$, whenever τ is a face of σ , then we have $f(\tau) \leq f(\sigma)$. Then, letting $F(t) := f^{-1}(-\infty, t]$, the sequence $\{F(t)\}_{t \in \mathbb{R}}$ is the *filtration* associated to f ; in particular, the filtration is a sequence of nested subcomplexes along with inclusion maps $F(s) \hookrightarrow F(t)$ for every $s \leq t$. Moreover, for each $k \in \mathbb{N}$, the inclusion $F(s) \hookrightarrow F(t)$ induces a linear map on homology, $H_k(F(s)) \rightarrow H_k(F(t))$. We write $\beta_k^{s,t}(K, f)$ to mean rank of this map, or simply $\beta_k^{s,t}$ if K and f are clear from context.

In particular, 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 $f_s: K \rightarrow \mathbb{R}$ defined by $f_s(\sigma) = \max\{s \cdot v \mid v \in K_0 \cap \sigma\}$. Note that s defines a preorder on K_0 , $v_0 \leq v_1 \leq \dots \leq v_{n_0-1}$. Then, the *lower-star filtration* of K with respect to s is

$$\emptyset \subset f_s^{-1}(-\infty, s \cdot v_0] \subseteq f_s^{-1}(-\infty, s \cdot v_1] \subseteq \dots \subseteq f_s^{-1}(-\infty, s \cdot v_{n_0-1}] = K.$$

Any filter function has *compatible index filters*, which are functions $f': K \rightarrow \mathbb{R}$ such that f' orders all the simplices of K uniquely and if $f(\tau) \leq f(\sigma)$, then $f'(\tau) \leq f'(\sigma)$. We say their corresponding filtrations are *compatible index filtrations*.

Our principal objects of study are *verbose persistence diagrams*. As they are closely related to the more familiar *concise persistence*, we begin with the following definition.

► **Definition 2.1** (Concise Persistence Diagram, ρ). Let $f: K \rightarrow \mathbb{R}$ be a filter function. We define the *kth-dimensional persistence diagram* as the following multiset:

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

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

Since simplices can appear at the same parameter value in a general filtration, not all cycles are represented in the persistence diagram. However, having every simplex “appear” in the persistence diagram is helpful, in addition to being natural. Thus, we introduce *verbose persistence diagrams*, which contain this information.

► **Definition 2.2** (Verbose Persistence Diagram, $\hat{\rho}$). Let $f: K \rightarrow \mathbb{R}$ be a filter and let f' be a compatible index filter. For $k \in \mathbb{N}$, the *k-dimensional verbose persistence diagram* is the following multiset:

$$\hat{\rho}_k^f := \left\{ (f(\sigma_i), f(\sigma_j)) \mid (i, j) \in \rho_k^{f'} \right\}. \quad (1)$$

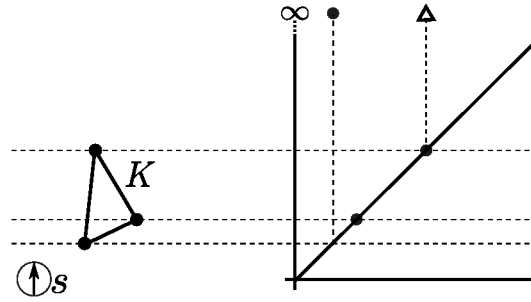
The *verbose persistence diagram* of f , denoted $\hat{\rho}^f$, is the union of all $\hat{\rho}_k^f$.

While concise persistence diagrams may feel more familiar, the idea of verbose persistence diagrams is not new. Indeed, many typical algorithms for computing persistence (e.g., [4, Chapter VII]), explicitly compute topological events with trivial lifespan but then discard them from the output. In [11], persistence diagrams are the same as our definition of $\hat{\rho}$. In [16], we see filtered chain complexes as a source of verbose persistence; [2, 12, 13, 17] also take this view.

If a verbose persistence diagram corresponds to a direction s in addition to a simplicial complex K , we use the notation $\hat{\rho}(K, s)$, or $\hat{\rho}(s)$ when K is clear from context. Furthermore, we refer to the set of verbose persistence diagrams parameterized by S as $\hat{\rho}(K, S)$. We denote the i th standard basis vector by e_i , so e_1 is the unit vector in the x -direction, etc. We say a (parameterized) set of verbose diagrams is *faithful* for a simplicial complex K if only K could have generated that same set of diagrams.

► **Definition 2.3 (Faithful Set).** Let K be simplicial complex in $\mathbb{R}^d$ and let $S \subseteq \mathbb{S}^{d-1}$. We say that $\hat{\rho}(K, S)$ is *faithful* if, for any K' we have the equality $\hat{\rho}(K', S) = \hat{\rho}(K, S)$ if and only if $K' = K$.

See Figure 1 for an example of a verbose persistence diagram and a non-faithful set.



■ **Figure 1** The verbose persistence diagram shown above, $\hat{\rho}(K, s)$, is identical to the corresponding concise persistence diagram, except for the on-diagonal points. Notice that the singleton set $\hat{\rho}(K, s)$ is not faithful; any cycle with vertices at the same heights as the vertices of K produces the same verbose persistence diagram.

3 Bounds on Faithful Sets

This section provides lower bounds on the size of faithful sets of verbose persistence diagrams. We begin with a tight lower bound.

► **Lemma 3.1 (Tight Lower Bound).** Let K be a simplicial complex in $\mathbb{R}^d$. Suppose that $\hat{\rho}(K, S)$ is faithful. Then $|S| \geq d$, and this bound is tight.

Proof. No vertex in K can be described using fewer than d coordinates. Thus, a set of verbose persistence diagrams with cardinality less than d cannot be faithful for K_0 , let alone K . To see that this bound is tight, consider the case where K is a single vertex; verbose descriptors generated by any d pairwise linearly independent directions form a faithful set for the vertex (e.g., $\hat{\rho}(K, \{e_1, e_2, \dots, e_d\})$). ◀

In many examples, we find that minimum faithful sets of verbose descriptors for simple simplicial complexes in $\mathbb{R}^d$ often have cardinality $d + 1$. Precise statements characterizing simplicial complexes with faithful sets of size $d + 1$ remain as ongoing work. However, frequently encountering faithful sets with a cardinality independent from the size of the simplicial complex is not at all surprising, since (unlike concise descriptors), verbose descriptors always have events corresponding to each simplex, regardless of how the complex is imbedded or immersed. We did not expect to find a simplicial complex whose minimum faithful set depended on the size of the complex; in fact, it was through trying to *disprove* the existence of such a complex that we came across the construction in this section.

In particular, in this section, we identify a family of simplicial complexes for which minimum faithful sets of verbose persistence diagrams are 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. We assume angles take value in $[0, 2\pi)$. We establish a preliminary observation, a specific instance of the general phenomenon that a simplicial complex stratifies the sphere of directions based on vertex order [3, 10].

► **Observation 1.** *Suppose that a simplicial complex K in $\mathbb{R}^2$ contains an edge $[v_1, v_2]$ such that v_1 and v_2 have degree one. Then a birth event occurs at the height of v_1 in $\hat{\rho}(K, s)$ for all s in the half of $\mathbb{S}^1$ defined by the open interval $H = (\alpha_{1,2} - \pi, \alpha_{1,2} + \pi)$ (i.e., all s so that $s \cdot v_1 > s \cdot v_2$) and as an instantaneous event for $s \in H^C$ (i.e., all s so that $s \cdot v_1 \leq s \cdot v_2$).*

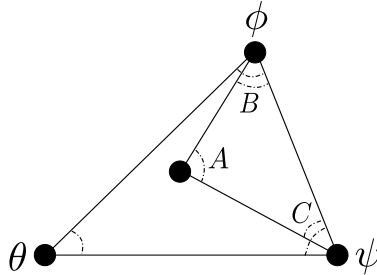
We also establish the following elementary lemma.

► **Lemma 3.2.** *Consider a pair of nested triangles as in Figure 2. 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,

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

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



■ **Figure 2** Nested triangles as discussed in Lemma 3.2

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 $\{v_1, v_2, v_3, v_4\}$. 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 image in Figure 3.*

Construction 1 was built specifically for the following necessary condition for faithful sets of verbose descriptors. We state this condition in terms of $\hat{\rho}$'s in the following lemma.

► **Lemma 3.3 (Clothespin Representability).** *Let K be a clothespin motif, as in Construction 1, and suppose that $\hat{\rho}(K, S)$ is faithful. Then we have at least one direction $s \in S$ such that the angle formed between s and $e_1 = (1, 0)$ lies in the region $[\alpha_{3,2} - \pi, \alpha_{3,4} - \pi] \cup [\alpha_{3,2} + \pi, \alpha_{3,4} + \pi]$.*

Proof. Let K' be a simplicial complex in $\mathbb{R}^2$ with the same vertex set as K , but with edges $[v_1, v_4]$ and $[v_2, v_3]$ (see the left side of Figure 3). Recall that, since $\hat{\rho}(K, S)$ is faithful, the set S must contain some s so that $\hat{\rho}(K, s) \neq \hat{\rho}(K', s)$.



■ **Figure 3** The two simplicial complexes considered in the proof of Lemma 3.3.

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 4 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 $B = (\alpha_{1,2} - \pi, \alpha_{1,2} + \pi)$ and $v_1 \in K'$ 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, identify 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' . Similar arguments for $i \in [2, 3, 4]$ give us the complete list;

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

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

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 3.2; the angles swept out by each regions correspond to the angles formed by pairs of edges in K and K' ; in particular, the angle $\angle v_2 v_3 v_4$ 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 . Observing that each of these four angles appears as an included endpoint for some R_i , we see $R_1, R_2, R_4 \subseteq \overline{R}_3 = W$ (see Figure 4), as desired. ◀

To get a deeper intuition for this result, observe that the verbose diagrams corresponding to K and K' of Figure 3 are identical when we filter in direction e_1 , but when we filter in direction e_2 , they are distinct. We refer to the wedge shaped region of directions for which the corresponding verbose diagrams have this distinction as a clothespin's *region of observability* (similar to observability for χ 's discussed in [6, 3]). We notate the region as $W = [\alpha_{3,4} - \pi, \alpha_{2,3} - \pi] \cup [\alpha_{3,4} + \pi, \alpha_{2,3} + \pi]$. Crucially, W is defined by the angle $v_2 v_3 v_4$, so a different embedding of K could result in a smaller region.

► **Remark (W Can be Arbitrarily Small).** As the angle $\angle v_2 v_3 v_4$ approaches zero, the region of observability, W , described in the proof of Lemma 3.3 also approaches zero.

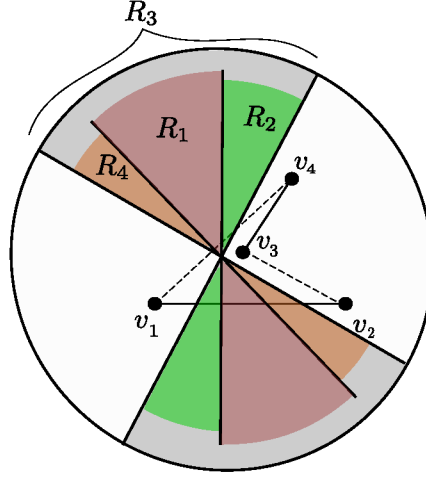


Figure 4 The regions described in the proof of Lemma 3.3. K is shown as solid black edges and K' as dashed edges. For any lower-star filtration in a direction contained in R_i , the event at vertex v_i differs when considering K or K' , thus, such directions are able to distinguish K from K' . Note that any direction outside the regions of observability (i.e., the non-shaded portions of the circle) is not able to distinguish K from K' .

To construct a family of simplicial complexes, each of which must have at least $\Theta(n_0)$ verbose descriptors to form a faithful set, we use the preceding remark to knit together clothespin motifs (Construction 1) in the following way.

► **Construction 2** (Clothespins on a Semicircular Clothesline). *Let K_m be a simplicial complex in $\mathbb{R}^2$ formed by m clothespins (Construction 1) such that the regions of observability for each clothespin do not overlap. Note this is possible for any m by the remark above.*

See Figure 5 for an example of K_m for $m = 4$. This construction implies a lower bound on the number of $\hat{\rho}$'s needed to fully represent a simplicial complex.

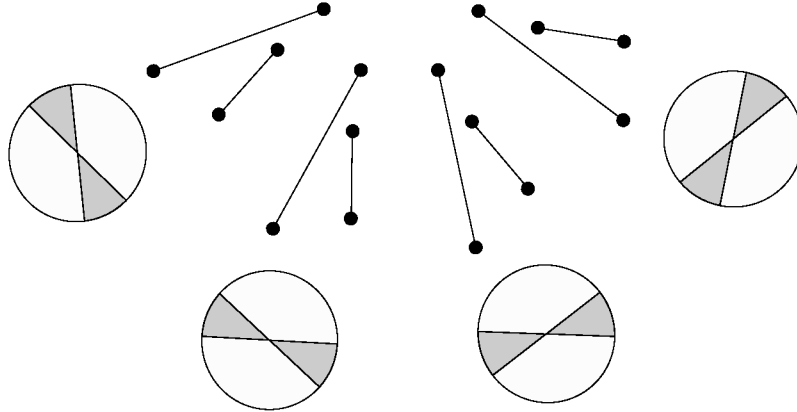


Figure 5 An example of Construction 2 for $m = 4$. The regions of observability are shown below each clothespin. By construction, each of these four double wedges define disjoint regions of $\mathbb{S}^2$.

► **Theorem 3.4** (Lower Bound for Worst-Case $\hat{\rho}$'s Complexity). *Let K_m be as in Construction 2 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 $\hat{p}$'s for K_m is $\Omega(n_0)$.

References

- 1 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*, 90, October 2020.
- 2 Wojciech Chachólski, Barbara Giunti, Alvin Jin, and Claudia Landi. Decomposing filtered chain complexes: Geometry behind barcoding algorithms. *Computational Geometry*, 109:101938, 2023.
- 3 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.
- 4 Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2010.
- 5 Brittany Terese Fasy, Samuel Micka, David L. Millman, Anna Schenfisch, and Lucia Williams. Challenges in reconstructing shapes from Euler characteristic curves, 2018. arXiv:1811.11337.
- 6 Brittany Terese Fasy, Samuel Micka, David L. Millman, Anna Schenfisch, and Lucia Williams. Challenges in reconstructing shapes from Euler characteristic curves. *Fall Workshop Computational Geometry*, 2018.
- 7 Brittany Terese Fasy, Samuel Micka, David L. Millman, Anna Schenfisch, and Lucia Williams. Efficient graph reconstruction and representation using augmented persistence diagrams. *Proceedings of the 34th Annual Canadian Conference on Computational Geometry*, 2022.
- 8 Brittany Terese Fasy, David L. Millman, and Anna Schenfisch. Ordering topological descriptors, 2024. arXiv:2402.13632.
- 9 Allen Hatcher. Algebraic topology, Cambridge Univ. Press, Cambridge, 2002.
- 10 Jacob Leygonie, Steve Oudot, and Ulrike Tillmann. A framework for differential calculus on persistence barcodes. *Foundations of Computational Mathematics*, pages 1–63, 2021.
- 11 Alexander McCleary and Amit Patel. Edit distance and persistence diagrams over lattices. *SIAM Journal on Applied Algebra and Geometry*, 6(2):134–155, 2022.
- 12 Facundo Mémoli and Ling Zhou. Stability of filtered chain complexes, 2022. arXiv:2208.11770.
- 13 Facundo Mémoli and Ling 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.
- 14 Samuel Micka. *Algorithms to Find Topological Descriptors for Shape Reconstruction and How to Search Them*. PhD thesis, Montana State University, 2020.
- 15 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.
- 16 Michael Usher and Jun Zhang. Persistent homology and Floer–Novikov theory. *Geometry & Topology*, 20(6):3333–3430, 2016.
- 17 Ling Zhou. *Beyond Persistent Homology: More Discriminative Persistent Invariants*. PhD thesis, The Ohio State University, 2023.
- 18 Zhen Zhou, Yongzhen Huang, Liang Wang, and Tieniu Tan. Exploring generalized shape analysis by topological representations. *Pattern Recognition Letters*, 87:177–185, 2017.