

Spectral relaxation of the persistence rank invariant

Matthew Piekenbrock^{1*} and Jose Perea^{1,2}

¹ Khoury College of Computer Sciences, Northeastern University .

² Department of Mathematics, Northeastern University .

*Corresponding author(s). E-mail(s): piekenbrock.m@northeastern.edu;
Contributing authors: jose.perea@northeastern.edu;

Abstract

Using a duality result between persistence diagrams and persistence measures, we introduce a framework for constructing families of continuous relaxations of the persistent rank invariant for parametrized families of persistence vector spaces indexed over the real line. Like the rank invariant, these families obey inclusion-exclusion, derive from simplicial boundary operators, and encode all the information needed to construct a persistence diagram. Unlike the rank invariant, these spectrally-derived families enjoy a number of stability and continuity properties, such as smoothness and differentiability over the positive semi-definite cone. Surprisingly, by leveraging a connection between stochastic Lanczos quadrature and implicit trace estimation, our proposed relaxation enables the matrix-free, iterative approximation scheme for all of the persistence invariants—Betti numbers, persistent pairs, and cycle representatives—in linear space and cubic time in the size of the filtration.

Keywords: Topological Data Analysis, Persistent Homology, Matrix functions

MSC Classification: 68Uxx , 55N31 , 55-08

1 Introduction

Persistent homology [1] (PH) is the most widely deployed tool for data analysis and learning applications within the topological data analysis (TDA) community. Persistence-related pipelines often follow a common pattern: given a data set X as input, construct a simplicial complex K and an order-preserving function $f : K \rightarrow \mathbb{R}$ such that useful topological/geometric information may be gleaned from its persistence diagram—a multiset summary of f formed by pairs $(a, b) \in \mathbb{R}^2$ exhibiting non-zero

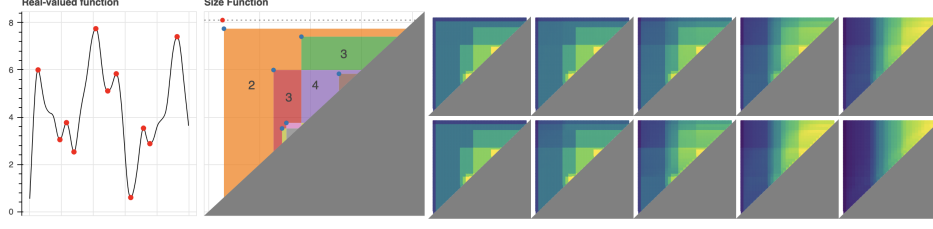


Fig. 1 (left) A function $f : \mathbb{R} \rightarrow \mathbb{R}$ and its corresponding persistence diagram obtained by filtering f by its sublevel sets. (right) Spectral interpolations to various Laplacian norms.

multiplicity $\mu_p^{a,b} \in \mathbb{Z}_+$ [2]:

$$\text{dgm}_p(K, f) = \{(a, b) : \mu_p^{a,b} \neq 0\} \quad (1)$$

$$\mu_p^{a,b} = \min_{\delta > 0} (\beta_p^{a+\delta, b-\delta} - \beta_p^{a+\delta, b+\delta}) - (\beta_p^{a-\delta, b-\delta} - \beta_p^{a-\delta, b+\delta}) \quad (2)$$

The surprising and essential quality of persistence is that these pairings exist, are unique, and are stable under additive perturbations [3]. Whether for shape recognition [4], dimensionality reduction [5], or time series analysis [6], persistence is the de facto connection between homology and the application frontier.

Though theoretically sound, diagrams suffer from many practical issues: they are sensitive to outliers, far from injective, and expensive—both to compute *and* compare. Towards ameliorating these issues, practitioners have equipped diagrams with additional structure by way of maps to function spaces; examples include persistence images [7], persistence landscapes [8], and template functions [9]. Tackling the issue of injectivity, Turner et al. [10] propose an injective shape statistic of directional diagrams associated to a data set $X \subset \mathbb{R}^d$, sparking both an inverse theory for persistence and a mathematical foundation for metric learning. Despite the potential these extensions have in learning applications, scalability issues due to high algorithmic complexity remain. Indeed, this issue is compounded in the parameterized setting, where adaptations of the persistence computation has proven non-trivial [11].

We seek to shift the computational paradigm on persistence while retaining its application potential: rather than following a construct-then-vectorize approach, we devise a spectral method that performs both steps, simultaneously and approximately. Our strategy is motivated both by a technical observation that suggests advantages exist for the rank invariant computation (Section 2.1) and by measure-theoretic results on \mathbb{R} -indexed persistence modules [2, 12], which generalize Equation 2 to rectangles $R = [a, b] \times [c, d] \subset \mathbb{R}^2$:

$$\mu_p^R(K, f) = \text{card}(\text{dgm}_p(K, f)|_R) = \beta_p^{b,c} - \beta_p^{a,c} - \beta_p^{b,d} + \beta_p^{a,d} \quad (3)$$

Notably, our approach not only avoids explicitly constructing diagrams, but is also *matrix-free*, circumventing the reduction algorithm from [13] entirely. Additionally, the relaxation is computable exactly in linear space and quadratic time, requires no complicated data structures or maintenance procedures to implement, and can be

iteratively $(1 \pm \epsilon)$ approximated in effectively linear time in practice for large enough $\epsilon > 0$.

Contributions: Our primary contribution is the introduction of several families of spectral approximations to the rank invariants— μ_p and β_p —all of which are Lipschitz continuous, and differentiable on the positive semi-definite cone (Proposition 2). Leveraging the vast spectral theory that exists for Laplacian operators (Section 2.3), we show these approximations are computable in $O(m)$ memory and $O(mn)$ time, where n (m) are the number of p ($p+1$, respectively) simplices in K (Proposition 3). Moreover, both relaxations recover their corresponding invariants when ϵ is made small enough, and both admit iterative $(1 \pm \epsilon)$ -approximation schemes (Section 4.2).

2 Notation & Background

A *simplicial complex* $K \subseteq \mathcal{P}(V)$ over a finite set $V = \{v_1, v_2, \dots, v_n\}$ is a collection of simplices $\{\sigma : \sigma \in \mathcal{P}(V)\}$ such that $\tau \subseteq \sigma \in K \Rightarrow \tau \in K$. A *p-simplex* $\sigma \subseteq V$ is a set of $p+1$ vertices, the collection of which is denoted as K^p . An *oriented p-simplex* $[\sigma]$ is an ordered set $[\sigma] = (-1)^{|\pi|} [v_{\pi(1)}, v_{\pi(2)}, \dots, v_{\pi(p+1)}]$, where π is a permutation on $[p+1] = \{1, 2, \dots, p+1\}$ and $|\pi|$ the number of its inversions. The *p-boundary* $\partial_p[\sigma]$ of an oriented *p-simplex* $[\sigma] \in K$ is defined as the alternating sum of its oriented co-dimension 1 faces, which collectively for all $\sigma \in K^p$ define the *p-th boundary matrix* ∂_p of K :

$$\partial_p[i, j] \triangleq \begin{cases} (-1)^{s_{ij}} & \sigma_i \in \partial[\sigma_j] \\ 0 & \text{otherwise} \end{cases}, \quad \partial_p[\sigma] \triangleq \sum_{i=1}^{p+1} (-1)^{i-1} [v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{p+1}] \quad (4)$$

where $s_{ij} = \text{sgn}([\sigma_i], \partial[\sigma_j])$ records the orientation. Extending Equation 2 to all simplices in $\sigma \in K$ for all $p \leq \dim(K)$ yields the *full boundary matrix* ∂ . With a small abuse in notation, we use ∂_p to denote both the boundary operator and its ordered matrix representative. When it is not clear from the context, we will clarify which representation is intended.

Generalizing beyond simplices, given a field \mathbb{F} , an *oriented p-chain* is a formal \mathbb{F} -linear combination of oriented *p-simplices* of K whose boundary $\partial_p[c]$ is defined linearly in terms of its constitutive simplices. The collection of *p-chains* under addition yields an \mathbb{F} -vector space $C_p(K)$ whose boundaries $c \in \partial_p[c']$ satisfying $\partial_p[c] = 0$ are called *cycles*. Together, the collection of *p-boundaries* and *p-cycles* forms the groups $B_p(K) = \text{Im } \partial_{p+1}$ and $Z_p(K) = \text{Ker } \partial_p$, respectively. The quotient space $H_p(K) = Z_p(K)/B_p(K)$ is called the *p-th homology group* of K with coefficients in \mathbb{F} and its dimension β_p is the *p-th Betti number* of K .

A *filtration* is a pair (K, f) where $f : K \rightarrow I$ is a *filter function* over an index set I satisfying $f(\tau) \leq f(\sigma)$ whenever $\tau \subseteq \sigma$, for any $\tau, \sigma \in K$. For every pair $(a, b) \in I \times I$ satisfying $a \leq b$, the sequence of inclusions $K_a \subseteq \dots \subseteq K_b$ induce linear transformations $h_p^{a,b} : H_p(K_a) \rightarrow H_p(K_b)$ at the level of homology. When \mathbb{F} is a field, this sequence of homology groups uniquely decompose (K, f) into a pairing (σ_a, σ_b) demarcating the evolution of homology classes [14]: σ_a marks the creation of a

homology class, σ_b marks its destruction, and the difference $|a - b|$ records the lifetime of the class, called its *persistence*. The persistent homology groups are the images of these maps and the persistent Betti numbers are their dimensions:

$$H_p^{a,b} = \begin{cases} H_p(K_a) & a = b \\ \text{Im } h_p^{a,b} & a < b \end{cases}, \quad \beta_p^{a,b} = \begin{cases} \beta_p(K_a) & a = b \\ \dim(H_p^{a,b}) & a < b \end{cases}$$

For a fixed $p \geq 0$, the collection of persistent pairs (a, b) together with unpaired simplices (c, ∞) form a summary representation $\text{dgm}_p(K, f)$ called the *p-th persistence diagram of (K, f)* . Conceptually, $\beta_p^{a,b}$ counts the number of persistent pairs lying inside the box $(-\infty, a] \times (b, \infty)$ —the number of persistent homology groups born at or before a that died sometime after b . When a given quantity depends on fixed parameters that are irrelevant or unknown, we use an asterisk. Thus, $H_p^*(K)$ refers to any homology group of K .

We will at times need to generalize the notation given thus far to the *parameterized* setting. Towards this end, for some $\mathcal{A} \subseteq \mathbb{R}^d$, we define an *\mathcal{A} -parameterized filtration* as a pair (K, f_α) where K is a simplicial complex and $f : K \times \mathcal{A} \rightarrow \mathbb{R}$ an \mathcal{A} -parameterized filter function satisfying:

$$f_\alpha(\tau) \leq f_\alpha(\sigma) \forall \tau \subseteq \sigma \in K \text{ and } f_\alpha(\sigma) \text{ is continuous in } \alpha \in \mathcal{A} \text{ for every } \sigma \in K$$

Intuitively, when $\mathcal{A} = \mathbb{R}$, one can think of α as a *time* parameter and each $f_\alpha(\sigma)$ as tracing a curve in \mathbb{R}^2 parameterized by α . Examples of parameterized filtrations include:

- (Constant filtration) For a filter $f : K \rightarrow \mathbb{R}$, note (K, f_α) generalizes the “static” notion of a filtration (K, f) in the sense one may declare $f_\alpha(\sigma) = f(\sigma)$ for all $\alpha \in \mathcal{A}$ and all $\sigma \in K$.
- (Dynamic Metric Spaces) For a set X , let $\gamma_X = (X, d_X(\cdot))$ denote a dynamic metric space [15], where $d_X(\cdot) : \mathbb{R} \times X \times X$ denotes a time-varying metric. For any fixed $K \subset \mathcal{P}(X)$, the pair (K, f_α) obtained by setting $f_\alpha(\sigma) = \max_{x, x' \in \sigma} d_X(\alpha)(x, x')$ recovers the notion of a *time-varying Rips filtration*.
- (Interpolating filtrations) For $f, g : K \rightarrow \mathbb{R}$ filters over K , a natural family of filtrations (K, h_α) is obtained by choosing a homotopy $h : \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}$ satisfying $h_0 = f$ and $h_1 = g$.
- (Fibred barcode) For a 2-d persistence module M , a common invariant of interest is the *fibred barcode*, which is the collection of barcodes of 1-d affine slices of M . These affine slices are themselves restrictions of a bifiltration parameterized by a 2-parameter family of lines homeomorphic to $[0, 1] \times \mathbb{R}$ [16].

2.1 Technical background

The following results summarize some technical observations motivating this effort, which will be used in several proofs. Though these observations serve as background

material, they contextualize our non-traditional computation of the rank invariant (Corollary ??) and serve as the motivation for this work.

Among the most widely known results for persistence is the structure theorem [14], which shows 1-parameter persistence modules decompose in an *essentially unique* way. Computationally, the corresponding Pairing Uniqueness Lemma [17] asserts that if $R = \partial V$ decomposes the boundary matrix $\partial \in \mathbb{F}^{N \times N}$ to a *reduced* matrix $R \in \mathbb{F}^{N \times N}$ using left-to-right column operations, then:

$$R[i, j] \neq 0 \Leftrightarrow \text{rank}(\partial^{i,j}) - \text{rank}(\partial^{i+1,j}) + \text{rank}(\partial^{i+1,j-1}) - \text{rank}(\partial^{i,j-1}) \neq 0 \quad (5)$$

where $\partial^{i,j}$ denotes the lower-left submatrix defined by the first j columns and the last $m - i + 1$ rows (rows i through m , inclusive). Thus, the existence of non-zero “pivot” entries in R may be inferred entirely from the ranks of certain submatrices of ∂ . Part of the validity of Equation 2.1 can be attributed to the following Lemma:

Lemma 1. *Given filtration (K, f) of size $N = |K|$, let $R = \partial V$ denote the decomposition of the filtered boundary matrix $\partial \in \mathbb{F}^{N \times N}$. Then, for any pair (i, j) satisfying $1 \leq i < j \leq N$, we have:*

$$\text{rank}(R^{i,j}) = \text{rank}(\partial^{i,j}) \quad (6)$$

Equivalently, all lower-left submatrices of ∂ have the same rank as their corresponding submatrices in R .

An explicit proof of both of these facts can be found in [18], though the latter was also noted in passing by Edelsbrunner [1]. Though typically used to prove the correctness of the reduction algorithm, the implications of these two observations are quite general, as recently noted by [19]:

Corollary 1 (Bauer et al. [19]). *Any matrix reduction algorithm which preserves the ranks of the submatrices $\partial^{i,j}(K, f)$ for all $i, j \in [N]$ satisfying $1 \leq i < j \leq N$ is a valid persistence algorithm.*

Indeed, though R is not unique, its non-zero pivots are, and these pivots *define* the persistence diagram. Moreover, due to Equation 6, both β_p^* and μ_p^* may be written as a sum of ranks of submatrices of ∂_p and ∂_{p+1} :

Corollary 2. *Given a fixed $p \geq 0$, a filtration (K, f) with filtration values $\{a_i\}_{i=1}^N$, and a rectangle $R = [a_i, a_j] \times [a_k, a_l] \subset \Delta_+$, the persistent Betti and multiplicity functions may be written as:*

$$\beta_p^{a_i, a_j}(K, f) = \text{rank}(C_p(K_i)) - \text{rank}(\partial_p^{1,i}) - \text{rank}(\partial_{p+1}^{1,j}) - \text{rank}(\partial_{p+1}^{i+1,j}) \quad (7)$$

$$\mu_p^R(K, f) = \text{rank}(\partial_{p+1}^{j+1,k}) - \text{rank}(\partial_{p+1}^{i+1,k}) - \text{rank}(\partial_{p+1}^{j+1,l}) + \text{rank}(\partial_{p+1}^{i+1,l}) \quad (8)$$

Though Equation 8 was pointed out by Cohen-Steiner et al. in [17] and exploited computationally by Chen & Kerber in [20], to the authors knowledge the only explicit derivation and proof of Equation 7 is given by Dey & Wang [18]. For completeness,

we give our own detailed proof of Corollary 2.1 in Section A.1. In practice, neither expressions seem used in any commonly used persistence software.

Two important properties of the expressions from Corollary 2.1 are: (1) they are comprised strictly of *rank* computations, and (2) all terms involve *unfactored* boundary matrices. Coupled with measure-theoretic perspectives on persistence [12], the former suggests variational perspectives of the rank function might yield interesting spectral relaxations of the expressions from Corollary 2.1 useful for e.g. optimization purposes. Both the latter property with Corollary 2.1 suggests a potentially new means of computing persistence information *without* matrix reduction. Moreover, combining these observations suggests advances made in other areas of applied mathematics may be readily exploited, such as the rich theory of matrix functions [21], the tools developed as part of “The Laplacian Paradigm” [22], or the recent connections between rank and trace estimation [23]. The rest of the paper is dedicated to exploring these connections and their implications.

2.2 Spectral relaxation and its implications

Before introducing our proposed relaxation, it is instructive to examine the how traditional expressions of the persistent rank invariants compare to those from Corollary 2.1. Given a filtration (K, f) of size $N = |K|$ with $f : K \rightarrow I$ defined over some index set I , its p -th persistent Betti number $\beta_p^{a,b}$ at index $(a, b) \in I \times I$, is defined as follows:

$$\begin{aligned}\beta_p^{a,b} &= \dim(Z_p(K_a) / B_p(K_b)) \\ &= \dim(Z_p(K_a) / (Z_p(K_a) \cap B_p(K_b))) \\ &= \dim(Z_p(K_a)) - \dim(Z_p(K_a) \cap B_p(K_b))\end{aligned}$$

Computationally, observe that Equation 2.2 reduces to one nullity computation and one subspace intersection computation. While the former is easy to re-cast as a spectral computation, computing the latter typically requires obtaining bases via matrix decomposition. Constructing these bases explicitly using conventional [21] or persistence-based [14, 24] algorithms effectively¹ requires $\Omega(N^3)$ time and $\Omega(N^2)$ space. As the persistence algorithm also exhibits $O(N^3)$ time complexity and completely characterizes $\beta_p^{a,b}$ over *all* values $(a, b) \in I \times I$, there is little incentive to compute $\beta_p^{a,b}$ with such direct methods (and indeed, they are largely unused). Because of this, we will focus on expressions Equation 7 and Equation 8 throughout the rest of the paper.

2.2.1 Parameterized boundary operators

In typical dynamic persistence settings (e.g. [17]), a decomposition $R = \partial V$ of the boundary matrix ∂ must be permuted and modified frequently to maintain a simplex-wise order respecting f_α . In contrast, the rank function is its permutation invariance:

¹An alternative means to perform matrix reduction is through the PLU factorization (see [18], section 5.1.1), which takes $O(n^\omega)$ time where $\omega \in [2, 2.373]$ is the matrix-multiplication constant. However, this approach exploits neither the structure of the boundary matrix nor properties of persistence to accelerate the computation and is not used in practice.

for any $X \in \mathbb{R}^{n \times n}$ and permutation P we have:

$$\text{rank}(X) = \text{rank}(P^T X P)$$

This suggests persistent rank computations like those from Corollary 2.1 need not maintain this ordering—as long as the constitutive boundary matrices the same non-zero pattern as their filtered counterparts, their ranks will be identical. In what follows, we demonstrate how exploiting this permutation invariance significantly simplifies the practical use of Equation 7 and Equation 8 in *parameterized* settings.

Let (K, f_α) denote parameterized family of filtrations of a simplicial complex of size $|K^p| = n$. Fix an arbitrary linear extension (K, \preceq) of the face poset of K . Define the \mathcal{A} -parameterized boundary operator $\hat{\partial}_p(\alpha) \in \mathbb{R}^{n \times n}$ of (K, f_α) as the $n \times n$ matrix ordered by \preceq for all $\alpha \in \mathcal{A}$ whose entries (k, l) satisfy:

$$\partial_p(\alpha)[k, l] = \begin{cases} s_{kl} \cdot f_\alpha(\sigma_k) \cdot f_\alpha(\sigma_l) & \text{if } \sigma_k \in \partial_p(\sigma_l) \\ 0 & \text{otherwise} \end{cases}$$

where $s_{kl} = \text{sgn}([\sigma_k], \partial[\sigma_l])$ is the sign of the oriented face $[\sigma_k]$ in $\partial[\sigma_l]$. Observe that Equation 2.2.1 may be decoupled into a product of diagonal matrices $D_*(f_\alpha)$:

$$\partial_p(\alpha) \triangleq D_p(f_\alpha) \cdot \partial_p(K_{\preceq}) \cdot D_{p+1}(f_\alpha) \quad (9)$$

where $D_p(f_\alpha)$ and $D_{p+1}(f_\alpha)$ are diagonal matrices whose non-zero entries are ordered by restrictions of f_α to K_{\preceq}^p and K_{\preceq}^{p+1} , respectively. Clearly, $\text{rank}(\partial_p(\alpha)) = \text{rank}(\partial_p(K_{\preceq}))$ when the diagonal entries of D_p and D_{p+1} are strictly positive. Moreover, observe we may restrict to those “lower left” matrices from Lemma 1 via post-composing step functions $\bar{S}_a(x) = \mathbf{1}_{x > a}$ and $S_b(x) = \mathbf{1}_{x \leq b}(x)$ to D_p and D_{p+1} , respectively:

$$\hat{\partial}_p^{a,b}(\alpha) \triangleq D_p(\bar{S}_a \circ f_\alpha) \cdot \partial_p(K_{\preceq}) \cdot D_{p+1}(S_b \circ f_\alpha)$$

Though these step functions are discontinuous at their chosen thresholds a and b , we may retain the element-wise continuity of Equation 9 by exchanging them with clamped *smoothstep* functions $\mathcal{S} : \mathbb{R} \rightarrow [0, 1]$ that interpolate the discontinuous step portion of S along a fixed interval $(a, a + \omega)$, for some $\omega > 0$ (see Figure 3).

The observations above collectively motivate our first relaxation. Without loss in generality, assume the orientation of the simplices (K, \preceq) is induced by the order on the vertex set V . To simplify the notation, we write $A^x = A^{*,x}$ to denote the submatrix including all rows of A and all columns of A up to x .

Proposition 1. *Given (K, f_α) , any rectangle $R = ([a, b] \times [c, d]) \subset \Delta_+$, and a constant $\delta > 0$ satisfying $a + \delta < b - \delta$ from Equation 3, the \mathcal{A} -parameterized invariants $\beta_p^{a,b} : \mathcal{A} \times K \rightarrow \mathbb{N}$ and $\mu_p^R : \mathcal{A} \times K \rightarrow \mathbb{N}$ defined by:*

$$\beta_p^{a,b}(\alpha) = \text{rank} \begin{bmatrix} D_p(S_a \circ f_\alpha) & 0 \\ 0 & \hat{\partial}_{p+1}^{a+\delta,b}(\alpha) \end{bmatrix} - \text{rank} \begin{bmatrix} \hat{\partial}_p^a(\alpha) & 0 \\ 0 & \hat{\partial}_{p+1}^b(\alpha) \end{bmatrix} \quad (10)$$

$$\mu_p^R(\alpha) = \text{rank} \begin{bmatrix} \hat{\partial}_{p+1}^{b+\delta,c}(\alpha) & 0 \\ 0 & \hat{\partial}_{p+1}^{a+\delta,d}(\alpha) \end{bmatrix} - \text{rank} \begin{bmatrix} \hat{\partial}_{p+1}^{a+\delta,c}(\alpha) & 0 \\ 0 & \hat{\partial}_{p+1}^{b+\delta,d}(\alpha) \end{bmatrix} \quad (11)$$

yield the correct quantities $\mu_p^R(K, f_\alpha) = \text{card}(\text{dgm}_p(f_\alpha)|_R)$ and $\beta_p^{a,b} = \dim(H_p^{a,b}(K, f_\alpha))$ for all $\alpha \in \mathcal{A}$.

Note that in Equation 2.2.1, we write $\partial_p(K_{\preceq})$ (as opposed to $\partial_p(K, f)$) to emphasize $\partial_p(K_{\preceq})$ is ordered according to a fixed linear ordering (K, \preceq) . The distinction is necessary as evaluating the boundary terms from Corollary 2.1 would require ∂ to be explicitly filtered in the total ordering induced by f_α —which varies in \mathcal{A} —whereas the expressions obtained by replacing the constitutive terms in Equation 7 and Equation 8 with Equation 10 and Equation 11, respectively, requires no such explicit filtering.

Corollary 3. *Given a boundary matrix $\hat{\partial}_p(\alpha) \in \mathbb{R}^{n \times n}$ constructed at time $\alpha \in \mathbb{R}$ from a parameterized family of filtrations (K, f_α) with linear extension \preceq , the time complexity of constructing $\hat{\partial}_p(\alpha')$ for any other $\alpha' \neq \alpha$ is $O(\max(|K^p|, |K^{p+1}|))$, assuming the evaluation of $f_\alpha(\tau)$ is $O(1)$ for every simplex $\tau \in K$.*

2.3 Parameterized Laplacians

For generality's sake, it is important to make the class of expressions for β_p^* and μ_p^* as large as possible. Since we are only concerned with homology over \mathbb{R} , we may exploit another identity of the rank function only applicable to zero characteristic fields:

$$\text{rank}(X) = \text{rank}(XX^T) = \text{rank}(X^TX), \quad \text{for all } X \in \mathbb{F}^{n \times m}$$

In the context of boundary operators, note that $\partial_1 \partial_1^T$ is the well known *graph Laplacian*—more generally, Equation 2.3 suggests we can readily express $\beta_p^*(\alpha)$ and $\mu_p^*(\alpha)$ using the ranks of combinatorial p -Laplacians.²

Following the seminal results from Horak and Jost [25], there are three natural ways to define p -Laplacians over a fixed simplicial complex K : the *up*-Laplacian $L_p^{\text{up}}(K)$, the *down*-Laplacian $L_p^{\text{dn}}(K)$, and their sum, which we refer to as the *combinatorial Laplacian* $\Delta_p(K)$:

$$\Delta_p(K) = \underbrace{\partial_{p+1} \circ \partial_{p+1}^T}_{L_p^{\text{up}}} + \underbrace{\partial_p^T \circ \partial_p}_{L_p^{\text{dn}}}$$

All three operators Δ_p , L_p^{up} , and L_p^{dn} are symmetric, positive semi-definite, and compact [24]—moreover, the *non-zero* multisets $\Lambda(L_p^{\text{up}})$ and $\Lambda(L_{p+1}^{\text{dn}})$ are equivalent, implying they must have identical ranks (see Theorem 2.2 and 3.1 of [25]). Thus, for rank computations, it suffices to consider only one of them.

Let (K, f_α) denote a parameterized family of filtrations of a simplicial complex K equipped with a fixed but arbitrary linear extension \preceq of its face poset and fixed orientations $s(\sigma)$ inherited from the total order on the vertex set (V, \preceq) . Without loss

²By convention, we define $\partial_p = 0$ for all $p \leq 0$.

of generality, we define the weighted p up-Laplacian $\mathcal{L}_p \triangleq L_p^{\text{up}}$ at index (a, b) as follows:

$$\mathcal{L}_{p,\leq}^{a,b}(\alpha) \triangleq D_p(\bar{S}_a \circ f_\alpha) \cdot \partial_{p+1}(K_{\leq}) \cdot D_{p+1}(S_b \circ f_\alpha) \cdot \partial_{p+1}^T(K_{\leq}) \cdot D_p(\bar{S}_a \circ f_\alpha)$$

where $D_p(f)$ denotes a diagonal matrix whose entries represent the application of f to the p -simplices of K . As in Equation 2.2.1, fixing step function S_a and \bar{S}_b at values $a, b \in \mathbb{R}$ yields operators whose ranks correspond to the ranks of certain “lower-left” submatrices of the corresponding full boundary matrix ∂ of (K, f) . In particular, if $R = \partial V$ is the decomposition of (K, f_α) for some fixed choice of $\alpha \in \mathcal{A}$, then for any pair $(a, b) \in \Delta_+$ there exists indices $i = \sum_{\sigma \in K} (\bar{S}_a \circ f_\alpha)(\sigma)$ and $j = \sum_{\sigma \in K} (S_b \circ f_\alpha)(\sigma)$ such that:

$$\text{rank}\left(R_{p+1}^{i,j}\right) = \text{rank}\left(\partial_{p+1}^{i,j}\right) = \text{rank}\left(\hat{\partial}_{p+1}^{a,b}\right) = \text{rank}\left(\left(\hat{\partial}_{p+1}^{a,b}\right)\left(\hat{\partial}_{p+1}^{a,b}\right)^T\right) = \text{rank}\left(\mathcal{L}_{p,\leq}^{a,b}\right)$$

where the second last equality uses the identity $\text{rank}(X) = \text{rank}(X^T X)$, which holds when $X \in \mathbb{F}^{n \times m}$ has coefficients in a zero-characteristic field \mathbb{F} . This confirms that we may substitute any of the parameterized boundary operators used in Proposition 1 with weighted Laplacian operators $\partial_{p+1}^* \mapsto \mathcal{L}_p^*$ equipped with the appropriate down- and up-step functions S_* and \bar{S}_* , respectively.

It is worth noting that composing step functions $\bar{S}_a, S_b : \mathbb{R} \rightarrow [0, 1]$ with f_α is equivalent to endowing a *weight function* $w : K \rightarrow (0, +\infty)$ on a subset $K_{a,b} \subseteq K$, in the sense described by [24]. In particular, if w_p denotes the restriction of w to K^p , then w_p defines an inner product $\langle \cdot, \cdot \rangle_{w_p}$ on space of p -chains $C_p(K, \mathbb{R})$ given by:

$$\langle [\sigma], [\sigma'] \rangle_{w_p} \triangleq \delta_{\sigma\sigma'} \cdot (w_p(\sigma))^{-1}, \quad \forall \sigma, \sigma' \in K^p$$

where $\delta_{\sigma\sigma'} = \mathbf{1}(\sigma = \sigma')$ is the indicator function on σ . Moreover, this inner product $\langle [\sigma], [\sigma'] \rangle_{w_p}$ on $C_p(K, \mathbb{R})$ induces an inner product:

$$\langle\langle f, g \rangle\rangle_w = \sum_{\sigma \in K^p} f([\sigma]) g([\sigma]) w(\sigma), \quad \text{for all } f, g \in C^p(K)$$

In these sense above, Equation 2.3 is simply choosing a particular inner product on the space of p -cochains.

Remark 1. One may interpret the action of sending a subset $S \subseteq K$ of p -simplices to 0 as a restriction of K to a sub-complex $L = K \setminus S$, which suggests e.g. Equation 2.2 and Equation 2.3 could be alternatively defined using the inclusion maps $L \hookrightarrow K$ between simplicial pairs (L, K) , as in [24].

3 Spectral rank relaxation

Under mild assumptions on f_α , the entries of the boundary operators from Equation 2.2.1 are continuous functions of α when S is substituted appropriately with

smooth step functions. In contrast, the quantities from Proposition 1 are by definition discontinuous functions, as they are integer-valued due to the rank function. To circumvent this issue, we consider the spectral characterization of the rank function:

$$\text{rank}(X) = \sum_{i=1}^n \text{sgn}_+(\sigma_i(X)), \quad \text{sgn}_+(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

In the above, $\{\sigma_i\}_{i=1}^n$ are the singular values $\Sigma = \text{diag}(\{\sigma_i\}_{i=1}^n)$ from the singular value decomposition (SVD) $X = U\Sigma V^T$ of $X \in \mathbb{R}^{n \times m}$, and $\text{sgn}_+ : \mathbb{R} \rightarrow \{0, 1\}$ is the one-sided sign function. As the singular values vary continuously under perturbations in X [21], it is clear the discontinuity in Equation 3 manifests from the one-sided sign function—thus, for the purposes of obtaining a continuous substitute for the rank function, a natural approach is to first relax the sgn_+ function.

Our approach follows the seminal work of Mangasarian et al. [26]. Let $p : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ denote a continuous density function and $\nu : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a continuous increasing function satisfying $\nu(0) = 0$. One way to approximate the sgn_+ function is to integrate τ -smoothed variations $\hat{\delta}$ of the Dirac delta measure δ :

$$\phi(x, \tau) = \int_{-\infty}^x \hat{\delta}(z, \tau) dz, \quad \hat{\delta}(z, \tau) = \frac{1}{\nu(\tau)} \cdot p\left(\frac{z}{\nu(\tau)}\right), \quad \forall z \geq 0, \tau > 0$$

In contrast to the sgn_+ function, if p is continuous on \mathbb{R}_+ then $\phi(\cdot, \tau)$ is continuously differentiable on \mathbb{R}_+ , and if p is bounded above on \mathbb{R}_+ , then $\phi(\cdot, \tau)$ is globally Lipschitz continuous on \mathbb{R}_+ . Moreover, varying $\tau \in \mathbb{R}_+$ in Equation 3 yields an τ -parameterized family of continuous sgn_+ relaxations $\phi : \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}_+$, where $\tau > 0$ controls the accuracy of the relaxation.

Many properties of the sign approximation from Equation 3 extend naturally to the rank function when substituted appropriately via Equation 3. In particular, pairing $X = U\Sigma V^T$ with a scalar-valued ϕ that is continuously differentiable at every entry σ of Σ yields a corresponding Löwner operator Φ_τ [27]:

$$\Phi_\tau(X) \triangleq \sum_{i=1}^n \phi(\sigma_i, \tau) u_i v_i^T$$

Note that when X is positive semi-definite, Equation 3 may be interpreted as a particular choice of *matrix function* $f(X) \triangleq Uf(\Lambda)U^T$ from the matrix function calculus perspective [21]. Unlike general matrix functions, however, the restrictions on ϕ Equation 3 grants the operators Φ_τ a variety of attractive properties related to approximation, monotonicity, and differentiability.

Proposition 2 (Bi et al. [27]). *The operator $\Phi_\tau : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{n \times m}$ defined by Equation 3 satisfies:*

1. *For any $\tau \geq 0$, the Schatten-1 norm $\|\Phi_\tau(X)\|_*$ of $\Phi_\tau(X)$ is given by $\sum_{i=1}^n \phi(\sigma_i, \tau)$.*
2. *For any $\tau' \geq \tau$, $\|\Phi_{\tau'}(X)\|_* \leq \|\Phi_\tau(X)\|_*$ for all $X \in \mathbb{R}^{n \times m}$.*

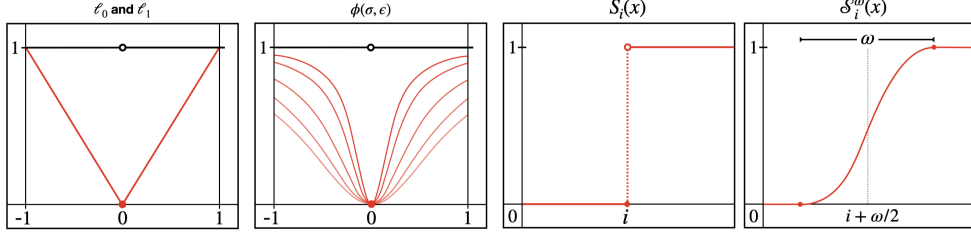


Fig. 2 From left to right—the ℓ_1 (red) norm and ℓ_0 (black) pseudo-norm are shown on the interval $([-1, 1])$; the relaxation $\phi(\cdot, \tau)$ from Equation 3 at various values of $\tau > 0$ (red) and at $\tau = 0$ (black); the step function $S_i(x)$ from Equation 2.2.1; the smoothstep relaxation S_i^ω discussed in Section 2.2.1.

3. For any given $X \in \mathbb{R}^{n \times m}$ with rank $r = \text{rank}(X)$ and positive singular values $\Lambda(X) = \{\sigma_1, \sigma_2, \dots, \sigma_r\}$, we have:

$$0 \leq r - \|\Phi_\tau(X)\|_* \leq r \cdot (1 - \phi(\sigma_r, \tau)) \quad (12)$$

Moreover, if τ satisfies $0 < \tau \leq \sigma_r/r$, then $r - \|\Phi_\tau(X)\|_*$ is bounded above by a constant $c_\phi(r) \geq 0$.

4. $\|\Phi_\tau(X)\|_*$ is globally Lipschitz continuous and semismooth³ on $\mathbb{R}^{n \times m}$.

Noting property (4), since the sum Lipschitz functions is also Lipschitz, it is easy to verify that replacing the rank function in all of the constitutive terms from Proposition 1 yields Lipschitz continuous functions whenever the filter function f_α is itself Lipschitz and the step functions from Equation 2.2.1 are smoothed ($\omega > 0$).

Interpretation #1: In many applications, it is common to regularize an ill-posed objective function to encourage simpler solutions or to prevent overfitting. For example, the classical least-squares approach to solving the linear system $Ax = b$ is often augmented with the *Tikhonov regularization* (TR) for some $\tau > 0$:

$$x_\tau^* = \arg \min_{x \in \mathbb{R}^n} \|Ax - b\|^2 + \tau \|x\|^2 = (A^T A + \tau I)^{-1} A^T b$$

When $\tau = 0$, one recovers the standard ℓ_2 minimization, whereas when $\tau > 0$ solutions x_τ^* with small norm are favored. Similarly, by parameterizing ϕ by $\nu(\tau) = \sqrt{\tau}$ and $p(x) = 2x(x^2 + 1)^{-2}$, one obtains via Equation 3:

$$\phi(x, \tau) = \int_0^z \hat{\delta}(z, \tau) dz = \frac{2}{\tau} \int_0^z z \cdot \left((z/\sqrt{\tau})^2 + 1 \right)^{-2} dz = \frac{x^2}{x^2 + \tau}$$

By substituting $\text{sgn}_+ \mapsto \phi$ and composing with the singular value function Equation 3, the corresponding spectral rank approximation reduces⁴ to the following *trace*

³Here, "semismooth" refers to the existence certain directional derivatives as $\tau \rightarrow 0^+$, see [21, 27].

⁴See Theorem 2 of [28] for a proof of the second equality.

formulation:

$$\|\Phi_\tau(A)\|_* = \sum_{i=1}^n \frac{\sigma_i(A)^2}{\sigma_i(A)^2 + \tau} = \text{tr} \left[(A^T A + \tau I)^{-1} A^T A \right]$$

The relaxation level τ may be thought of as a regularization term that preferences smaller singular values: larger values smooth out $\|\Phi_\tau(\cdot)\|_*$ by making the pseudo-inverse less sensitive to perturbations, whereas smaller values lead to a more faithful⁵ approximations of the rank. In this sense, we interpret the quantities obtained by applying Equation 3 to the terms from Proposition 1 as *regularized RI approximation*.

Interpretation #2: In spectral graph theory, matrix functions are often used to simulate diffusion processes on meshes or graphs embedded in \mathbb{R}^d to obtain information of about their geometry. For example, consider a weighted graph $G = (V, E)$ with $n = |V|$ vertices with graph Laplacian $L_G = \partial_1 \partial_1^T$. The *heat* of every vertex $v(t) \in \mathbb{R}^n$ as a function of time $t \geq 0$ is governed by L_G and the *heat equation*:

$$v'(t) = -L_G v(t) \quad \Leftrightarrow \quad L_G \cdot u(x, t) = -\partial u(x, t) / \partial t$$

To simulate a diffusion process on G from an initial distribution of heat $v(0) \in \mathbb{R}^n$, it suffices to construct the *heat kernel* $H_t \triangleq \exp(-t \cdot L_G)$ via the spectral decomposition $L_G = U \Lambda U^T$ of L_G :

$$v(t) = H_t v(0), \text{ where } H_t = \sum_{i=1}^n e^{-t\lambda_i} u_i u_i^T$$

The heat kernel is invariant under isometric deformations, stable under perturbations, and is known to contain multiscale geometric information due to its close connection to geodesics. As is clear from Equation 3, it is also a matrix function. Now, consider Equation 3 with $\nu(\tau) = \tau$ and $p(\lambda) = \exp(-\lambda_+)$ where $x_+ = \max(x, 0)$:

$$\phi(\lambda, \tau) = \int_0^z \hat{\delta}(z, \tau) dz = \frac{1}{\tau} \int_0^z \exp(-z/\tau) dz = 1 - \exp(-\lambda/\tau), \quad \text{for all } \lambda \geq 0$$

In the context of diffusion, observe the parameter τ is inversely related diffusion time (i.e. $t = 1/\tau$) and that as $t \rightarrow 0$ (or $\tau \rightarrow \infty$) the expression $1 - \exp(-\lambda/\tau)$ approaches the sgn_+ function on the interval $[0, \infty)$. As above, substituting ϕ appropriately into Equation 3 again yields an equivalent trace expression:

$$\|\Phi_\tau(L_G)\|_* = \sum_{i=1}^n 1 - \exp(-\lambda_i/\tau) = n - \text{tr}[H_{1/\tau}]$$

The heat kernel H_t has been shown to fully characterize shapes up to isometry, motivating the creation of various geometric signatures, such as the Heat Kernel Signature

⁵This can be seen directly by Equation 3 as well, wherein increasing τ lowers the condition number of $A^T A + \tau I$ monotonically, signaling a tradeoff in stability at the expense of accuracy.

(HKS) and the Heat Kernel Trace. In this sense, we interpret a spectral rank relaxation using Equation 3 as a *geometrically informative RI approximation*.

4 Computational Implications

4.1 Exact computation

As evidenced by Section 2.2, computing $\hat{\mu}_p^*$ and $\hat{\beta}_p^*$ may be reduced to computing eigenvalues of p -Laplacians. To do this efficiently, we employ the *Lanczos method* [29], which estimates the eigenvalues of any symmetric linear operator A via projection onto successive Krylov subspaces. Given a symmetric $A \in \mathbb{R}^{n \times n}$ with eigenvalues $\lambda_1 \geq \lambda_2 > \dots \geq \lambda_r > 0$ and a vector $v \neq 0$, the Lanczos method generates the triple (K, Q, T) :

$$\begin{aligned} K &= [A^0 v \mid A^1 v \mid A^2 v \mid \dots \mid A^{r-1} v] \\ Q &= [q_1, q_2, \dots, q_r] \leftarrow \text{qr}(K) \\ T &= Q^T A Q \end{aligned}$$

where $K \in \mathbb{R}^{n \times r}$ is the *Krylov matrix* with respect to (A, v) , $Q \in \mathbb{R}^{n \times r}$ is an orthogonal change-of-basis, and $T \in \mathbb{R}^{r \times r}$ is symmetric tridiagonal *Jacobi matrix*. It is well known that Q is a similarity transform, i.e. T preserves the spectrum of A and the problem of finding eigenvalues reduces to diagonalizing T .

The Lanczos method is often called a “matrix free” method due to its only prerequisite for computation being a matrix-vector product operator $v \mapsto Av$ — A need not necessarily be stored in memory explicitly. For this reason, the Lanczos method is often used as a low-memory option for computing eigenvalues. Indeed, due to its *three-term recurrence* [30], the Lanczos method requires just three $O(n)$ -sized vectors and a constant number of $O(n)$ vector operations to obtain T —neither K nor Q need be formed explicitly.

Lemma 2 ([31]). *Given a symmetric rank- r matrix $A \in \mathbb{R}^{n \times n}$ whose matrix-vector operator $x \mapsto Ax$ requires $O(\eta)$ time and $O(\nu)$ space, the Lanczos iteration computes $\Lambda(A) = \{\lambda_1, \lambda_2, \dots, \lambda_r\}$ in $O(\max\{\eta, n\} \cdot r)$ time and $O(\max\{\nu, n\})$ space, when computation is done in exact arithmetic.*

Surprisingly, Lemma 2 suggests the time and space complexities of computing quantities that depend on $\{\lambda_1, \lambda_2, \dots, \lambda_r\}$ are lower for certain classes of matrices, provided ν and η are sufficiently small. In general, the size of these variables depends on the structure and sparsity of the operators⁶ involved. In some cases, both ν and η are $O(n)$; for example, the application $x \mapsto Lx$ for the graph Laplacian $L = \partial_1 \partial_1^T$ is linear in the number of edges $|E|$ due to its graph structure. Though this result is well established for the graph Laplacian, it is not immediately clear whether a similar guarantee generalizes to combinatorial Laplacian operators derived from simplicial complexes—our next result affirms this.

⁶For $n \times n$ sparse matrices with an average of z non-zeros per row and rank r , for example, Lemma 2 implies the Lanczos method has an expected $O(nzr)$ time and $O(nz)$ space complexities [32].

Lemma 3. *For any p -dimensional simplicial complex K with $n = |K^p|$ and $m = |K^{p+1}|$, if there exists a hash function $h : K^p \rightarrow [n]$ constructible in $O(m)$ time supporting $O(1)$ access time and $O(c)$ space, then there exists a two-phase algorithm for computing the product $x \mapsto L_p x$ in $O(m(p+1))$ time and $O(\max(c, m))$ space.*

The algorithm and proof are given in Section A. From a practical perspective, many hash table implementations achieve expected $O(1)$ access time using only a linear amount of storage, and as $p \geq 0$ is typically quite small—the operation $x \mapsto Lx$ in practice exhibits $\approx O(m)$ time and space complexities. We delegate more practical issues regarding the computation to ???. Combining Lemma 2 and Lemma 3 yields our main result for this section.

Proposition 3. *For any constant $p \geq 0$ and box $R = ([a, b] \times [c, d]) \subset \Delta_+$, the persistent multiplicity function $\mu_p^R(K)$ derived from a simplicial complex K with $n_{ad} = |K_d^p| - |K_a^p|$ and $m_{ad} = |K_d^{p+1}| - |K_a^{p+1}|$, can be computed in exact arithmetic with the Lanczos method in the following time and space complexities:*

$$\mu_p^R(K) \stackrel{\text{time}}{=} O(n_{ad} \cdot m_{ad}), \quad \mu_p^R(K) \stackrel{\text{space}}{=} O(\max(n_{ad}, m_{ad}))$$

*In particular, when $R = (-\infty, *] \times [* , +\infty)$, $\mu_p^R(K)$ has time and space complexities of $O(nm)$ and $O(\max(n, m))$, respectively, where $n = |K^p|$ and $m = |K^{p+1}|$.*

It’s worth noting that the standard reduction-family of algorithms computes the p -th persistent homology of a filtration K of dimension $p+1$ and of size $N = |K| \sim O(|K^{p+1}|)$ in $\Theta(N^3)$ time and $\Theta(N^2)$ space. Interestingly, Chen and Kerber [20] have shown that since the persistence diagram contains at most $N/2 = O(N)$ points, it may be constructed using at most $2N-1$ “ μ -queries” (evaluations of μ_p^R) via a divide-and-conquer scheme on the index-persistence plane. Since both $|K^p|$ and $|K^{p+1}|$ are trivially bounded by $O(N)$, by Proposition 3, we may recover the same $O(N^3)$ time complexity of the reduction algorithm using only rank computations, and we improve the space complexity by a factor of N , though at the cost of not having immediate access to cycle representatives.

Remark 2. *A similar result for computing non-persistent Betti numbers of simplicial complexes over finite fields was given by Edelsbrunner and Parsa in [33], wherein the complexity of computing the Betti numbers of a 2-dimensional simplicial complex K with n vertices was shown to be $\Omega(r(n, m))$, where $r(n, m)$ is the complexity of computing the rank of an n -by- n binary matrix with m non-zero entries.*

4.2 Randomized (η, ϵ) -approximation

As in [31], Proposition 3 assumes an exact arithmetic computation model to simplify both the presentation of the theory and the corresponding complexity statements. In practice, finite-precision arithmetic introduces *both* rounding and cancellation errors into the computation affect both the convergence and termination conditions of the

Lanczos method, prohibiting its use practically. Though many improvements have been proposed throughout the decades (e.g. selective re-orthogonalization, implicit restarting, see [34][35] for an overview), it is well-known that the Lanczos method is not efficient at accurately approximating eigenvalues on the interior of the spectrum.

Fortunately, it turns out accurately estimating eigenvalues is not necessary for estimating the rank, largely due to a recent result demonstrating the Lanczos method is stable for matrix function⁷ approximation even in the finite precision arithmetic model. This result was motivated by the fact that Lanczos expansion up to degree k can *exactly* apply any matrix polynomial p with degree $< k$ [36]:

$$p(A)q_1 = Qp(T)e_1 \quad \Leftrightarrow \quad f(A)x = \|x\| \cdot Qf(T)e_1$$

In particular, Paige’s A27 Lanczos variant ([37]), when executed up to degree k , has an error that is bounded by the uniform error the best degree- p polynomial approximation to any bounded function f with degree $p < k$ (see [36] for conditions). For general matrix functions $f(A)$, this implies that finite-precision Lanczos essentially matches the strongest known exact arithmetic bounds. The particular relevance of Equation 4.2 to the computation of our proposed spectral ϕ -approximation Equation 3 is the fact that the quantity of interest $\|\Phi(\mathcal{L}_p)\|_*$ is expressible as a trace-norm, i.e. it may be computed via:

$$\|\Phi_\tau(A)\|_* = \text{tr}(\Phi_\tau(\mathcal{L}_p)) = \sum_{i=1}^n e_i^T \Phi_\tau(\mathcal{L}_p) e_i$$

where $\{e_1, \dots, e_n\}$ is an orthonormal basis. For this reason, we propose the use *stochastic Lanczos quadrature* (SLQ) type methods [23] for estimating quantities of the form $\text{tr}(f(A))$. The simplest such estimator is the Girard-Hutchinson (GH) estimator:

$$\text{tr}(f(A)) \approx \frac{n}{n_v} \sum_{j=1}^{n_v} e_1^T f(T_k^{(j)}) e_1 = \frac{n}{n_v} \sum_{j=1}^{n_v} \left(\sum_{i=1}^k \tau_i^{(j)} f(\theta_i^{(j)}) \right), \quad \tau_i = [e_1^T y_i]^2$$

where $T_k^{(j)}$ represents the Jacobi matrix of the degree- $(k+1)$ Krylov expansion of $(A, v^{(j)})$ and (θ_i, y_i) represent the Rayleigh-Ritz pairs associated with the tridiagonal eigendecomposition $T_k = Y\Theta Y^T$. When $v^{(j)} \sim \mathcal{D}$ derives from a sub-Gaussian distribution satisfying $\mathbb{E}[v^{(j)} \otimes v^{(j)}] = I$, the approximation Equation 4.2 is known to be an unbiased estimator of $\text{tr}(f(A))$. Under mild assumptions (see [23]), if the function of interest $f : [a, b] \rightarrow \mathbb{R}$ is analytic on $[\lambda_{\min}, \lambda_{\max}]$, then for constants $\epsilon, \eta \in (0, 1)$ the GH estimator Γ satisfies:

$$\Pr(|\text{tr}(f(A)) - \Gamma| \leq \epsilon |\text{tr}(f(A))|) \geq 1 - \eta$$

In other words, we can achieve a relative ϵ -approximation of $\text{tr}(f(A))$ with success probability η using on the order of $O(\epsilon^{-2} \log(\eta^{-1}))$ evaluations of $e_1^T f(T_k) e_1$ over random isotropic vectors $v \sim \mathcal{D}$.

⁷Recall that if $A \in \mathbb{R}^{n \times n}$ has eigenvalue decomposition $A = U\Lambda U^T$, the matrix function $f(A) \in \mathbb{R}^{n \times n}$ with respect to a function f is $Uf(\Lambda)U^T$, where $f(\Lambda)$ applies f to each diagonal entry of Λ .

More recent work by [38] has shown that deflation techniques can reduce the number of matrix-vector evaluations needed down to $O(\varepsilon^{-1} \sqrt{\log(\eta^{-1})} + \log(\eta^{-1}))$, though we note in Section 6.1 there are additional limitations to be aware of using these methods (e.g. a $O(mk)$ space complexity).

4.3 Apparent pairs optimization

One of the defining aspects of the above trace-based computation is that, unlike the reduction algorithm, execution does not require matrix decomposition. Unfortunately, one downside to this is that we lose access to certain computational shortcuts developed specifically for the reduction setting; it is not immediately clear whether persistence-specific optimizations like *clearing* and *cohomology* [18] have analogous shortcuts in the matrix-free setting. Nonetheless, one such optimization well known for clique filtrations—the identification of *apparent pairs*—does have a direct translation to the trace estimation setting. On some types of problems, this optimization alone enables us to discard 90-99% of the columns of ∂_p , prior to any rank computations.

Apparent pairs (APs) are a class of persistence pairs which are already reduced in the filtration boundary matrix $\partial(K)$. To motivate their use in our proposed trace-based computation, consider the following lemma:

Lemma 4. *Let $\partial_{p+1}(K)$ denote the dimension $p+1$ filtered boundary matrix obtained from the $R = \partial V$ decomposition of a simplexwise filtration (K, f) . Then, the p -th persistence diagram $\text{dgm}_p(K)$ determines a partitioning of the columns of $\partial_{p+1}(K)$ into submatrices ∂_{p+1}^- and ∂_{p+1}^+ :*

$$\partial_{p+1}^- = \{ \partial[\sigma_j] : \text{col}_R(j) \neq 0, \sigma_j \in K^{p+1} \}, \quad \partial_{p+1}^+ = \{ \partial[\sigma_j] : \text{col}_R(j) = 0, \sigma_j \in K^{p+1} \}$$

and so that:

$$\text{rank}(\partial_{p+1}(K)) = \text{rank}(\partial_{p+1}^-(K))$$

In other words, from a rank-based perspective, we may safely discard $p+1$ simplices whose boundary chains lie in the kernel of R . Of course, if ∂_{p+1}^- is obtained from the full decomposition $R = \partial K$, then the rank of any submatrix of ∂ is fully determined, leaving the utility of the above observation moot.

Fortunately, for simplexwise clique filtrations, we can use APs $(\tau, \sigma) \in \text{dgm}_{p+1}(K)$ to construct an approximation $\hat{\partial}_{p+1}^- \supset \partial_{p+1}^-$ efficiently. The main benefit of using APs here is that they can be readily identified based on a purely local condition, which is evident from their very definition:

Definition 1 (Apparent Pair). *Given a simplexwise filtration K_{\preceq} of a complex K , a pair of simplices (τ, σ) of K is called an apparent pair of K_{\preceq} if (1) τ is youngest facet of σ , and (2) σ is the oldest cofacet of τ .*

Equivalently, a pair (τ, σ) of K is *apparent* if all entries below or to the left of (τ, σ) in the filtered boundary matrix ∂ are zero. Since any persistent pair $(\tau, \sigma) \in \text{dgm}_p(K)$

corresponds to a pair of columns in R_p and R_{p+1} (respectively), the boundary chain $\partial[\sigma]$ must lie in the kernel of $\partial_{p+1}V_{p+1}$, and therefore can be safely discarded. Note that the set of APs does not depend on the coefficient field the homology of the complex is defined over, though they do depend on the simplexwise ordering.

Identifying an AP $(\tau, \sigma) \in \text{dgm}_p(K)$ can be reduced to enumerating cofacets of $\tau \in K^p$. To do this efficiently, we follow the low-memory approach used in the popular software Ripser [39], which restricts its computation to simplexwise filtrations induced by the *reverse colexicographical* vertex order. In this setting, p -simplices $(v_{i_p}, \dots, v_{i_0})$ satisfying $v_{i_p} > \dots > v_{i_0}$ are mapped to integers $r \in [0, C(n, p+1))$ —where $C(n, k)$ is the binomial coefficient—via the *combinatorial number system*; when the colexicographical order is used, the bijection is given by the sum $(v_{i_p}, \dots, v_{i_0}) \mapsto \sum_{j=1}^{p+1} C(i_j, j)$, and cofacets are given by the following relation:

$$(v_{i_p}, \dots, v_{i_{k+1}}, v_j, v_{i_k}, \dots, v_{i_0}) \mapsto \sum_{l=k+1}^p \binom{i_l}{l+2} + \binom{j}{k+1} + \sum_{l=0}^k \binom{i_l}{l+1}$$

By enumerating $j = n-1, \dots, 0$ for $j \notin [n] \setminus \{i_p, \dots, i_0\}$, one recovers all of $n - (p+1)$ cofacets of $(v_{i_p}, \dots, v_{i_0})$ in reverse colexicographic vertex order. Cofacet enumeration in the colexicographical order is particularly efficient due to the fact that the left and right partial sums in Equation 4.3 can be maintained throughout the enumeration: assuming all $O(n \cdot (p+1))$ binomial coefficients are precomputed, finding all the cofacets of any p -simplex τ requires just $O(n-p)$ additions in integer arithmetic.

The most prevalent subset of APs relevant to the clique-filtrations are those with zero persistence. In practice, zero persistence APs may be identified without enumerating all cofacets by using a “shortcut” involving a correspondence between zero-persistence APs and lexicographically minimal (maximal, respectively) facet (cofacet, respectively) pairs (see Proposition 3.12 in [39]). This “shortcut” is particularly useful due to the fact that zero persistence pairs comprise a large proportion of the apparent pairs—indeed, Theorem 3.10 in [39] shows that in dimension 1, the zero persistence pairs of a simplexwise refinement of the Vietoris-Rips filtration are *precisely* the apparent pairs of the same filtration.

5 Applications & Experiments

5.1 Filtration optimization

It is common in TDA for the filter function $f : K \rightarrow \mathbb{R}$ to depend on hyper-parameters. For example, prior to employing persistence, one often removes outliers from point set $X \subset \mathbb{R}^d$ via some density-based pruning heuristic that itself is parameterized. This is typically necessary due to the fact that, though stable under Hausdorff noise [3], diagrams are notably unstable against *strong outliers*—even one point can ruin the summary. As an exemplary use-case of our spectral-based method, we re-cast the problem of identifying strong outliers below as a problem of *filtration optimization*. Consider a Delaunay complex K realized from a point set $X \subset \mathbb{R}^2$ sampled around

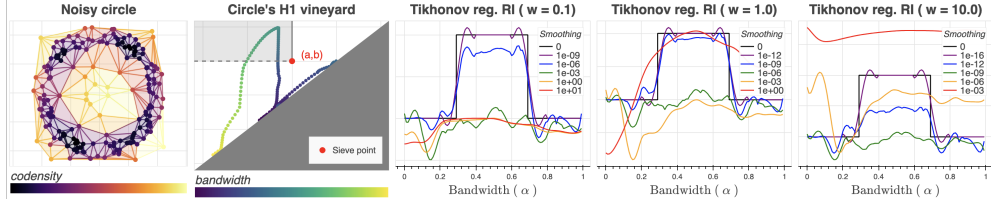


Fig. 3 From left to right: Delaunay complex K realized from point set $X \subset \mathbb{R}^2$ sampled with multiple types of noise around S^1 (colored by codensity at optimal $\alpha^* \approx 1/2$); codensity vineyard of (K, f_α) across varying bandwidths α and a fixed point $(a, b) \in \Delta_+$; Tikhonov relaxations $\hat{\beta}_p^{a,b}(\alpha)$ at varying regularization (τ) and sign width (ω) values.

S^1 affected by both Hausdorff noise and strong outliers, shown in Figure 5.1. One approach to detect the presence of S^1 in the presence of such outliers is maximize $\beta_p^{a,b}(\alpha)$ for some appropriately chosen $(a, b) \in \Delta_+$ over the pair (K, f_α) , where $f_\alpha : X \rightarrow \mathbb{R}_+$ is a kernel (co)-density estimate:

$$\alpha^* = \arg \max_{\alpha \in \mathbb{R}} \beta_p^{a,b}(K, f_\alpha), \quad \text{where } f_\alpha(x) = \frac{1}{n\alpha} \sum_i C(\mathcal{K}) - \mathcal{K}_\alpha(x_i - x)$$

where $C(\mathcal{K}_\alpha)$ is a normalizing constant that depends on the choice of kernel, \mathcal{K}_α . Intuitively, if there exists a choice of bandwidth α^* which distinguishes strong outliers from Hausdorff noise clustered around S^1 , then that choice of bandwidth should exhibit a highly persistent pair $(a^*, b^*) \in \text{dgm}_1(K, f_{\alpha^*})$. Thus, if we place a corner point (a, b) satisfying $a^* \leq a$ and $b < b^*$ for some $\alpha > 0$, we expect $\beta_p^{a,b}(\alpha) = 1$ to near the optimal bandwidth α^* —matching the first Betti number of S^1 —and 0 otherwise.

In Figure 5.1, we depict the dimension-1 vineyard of a simple Delaunay complex and codensity pair (K, f_α) , along with the sieve point (a, b) and the region wherein S^1 is accurately captured by persistence. As $\beta_p^{a,b}$ is an integer-valued invariant, it is discontinuous and difficult to optimize; in contrast, we know from Proposition 2 that we can obtain a continuous and differentiable relaxation of $\beta_p^{a,b}$ by replacing $\beta_p^{a,b} \mapsto \hat{\beta}_p^{a,b}$ in Equation 5.1, enabling the use of first-order optimization techniques. By using the Tikhonov regularization from Equation 3, we obtain continuously varying objective curves from $\hat{\beta}_p^{a,b}(\alpha; \tau)$ which are guaranteed to have the same maxima as $\beta_p^{a,b}(\alpha)$ as $\tau \rightarrow 0$, as shown in Figure 5.1. Observe lower values of τ lead to approximations closer to the rank (black) at the cost of smoothness, while larger values can yield very smooth albeit possibly uninformative relaxations. Practical optimization of these types of objective surfaces can be handled via *iterative thresholding*, a technique which alternates between gradient steps to reduce the objective and thresholding steps to enforce the rank constraints. We leave the tuning of such optimizers to future work.

5.2 Topology-guided simplification

In many 3D computer graphics applications, one would like to simplify a given simplicial or polygonal mesh embedded in \mathbb{R}^3 so as to decrease its level of detail (LOD)

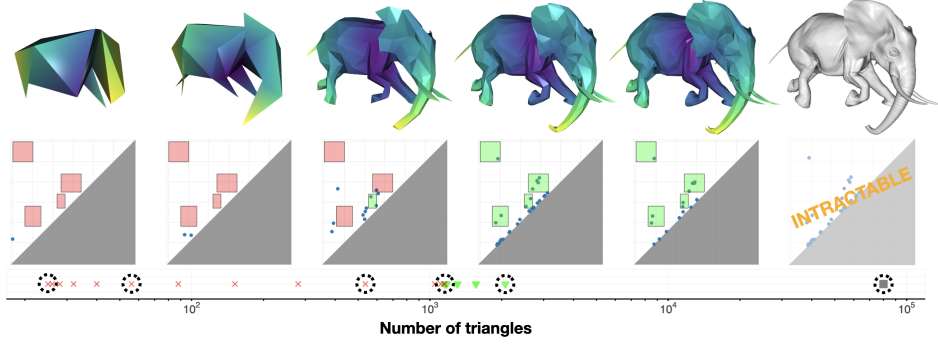


Fig. 4 (Top) Meshes filtered and colored by eccentricity at varying levels of simplification; (middle) their diagrams and topological constraints; (bottom) simplification thresholds tested by an exponential search, on a logarithmic scale. The color/shape of the markers indicate whether the corresponding meshes meet (green triangle) or do not meet (red x) the topological constraints of the sieve—the gray marker identifies the original mesh (not used in the search). Black dashed circles correspond with the meshes in the top row.

while retaining its principal geometric structure(s). Such simplifications are often necessary to improve the efficiency of compute-intensive tasks that depend on the size of the mesh (e.g. rendering). Though many simplification methods developed to preserve geometric criteria (e.g. curvature, co-planarity) are now well known (see [40] for an overview), *topology-preserving* simplification techniques are relatively sparse, especially for higher embedding dimensions. Moreover, such procedures typically restrict to operations that preserve *local* notions of topology, such as the genus of a feature’s immediate neighborhood or the property of being a manifold. These operations are known to greatly limit the amount of detail decimation algorithms can remove.

As a prototypical application of our proposed relaxation, we re-visit the mesh simplification problem under *persistence-based* constraints. In contrast to [41], we forgo the use of persistence-preserving operations and instead opt for a simpler strategy: we perform an exponential search on a given sequence of simplifications, settling on the largest simplification found.

We show an exemplary application of this idea in Figure 4 in sparsifying a mesh of an elephant. To filter the mesh in a geometrically meaningful way, we use the (geodesic) *eccentricity* function, which assigns points $x \in X$ in a metric space (X, d_X) a non-negative value representing the distance that point is from the center of the mesh:

$$E_p(x) = \left(\frac{\sum_{x' \in X} d_X(x, x')^p}{N} \right)^{\frac{1}{p}}$$

To extend $E_p(x)$ to simplices $\sigma \in K$, we take the max $f(\sigma) = \max_{v \in \sigma} d_X(x, v)$. Note this does not require identifying a center point in the mesh. Intuitively, the highly persistent H_1 classes in mesh carrying the most detail corresponds to “tubular” features that persist from the center; the four legs, the trunk, the two ears, and two tusks. In this example, four rectangular-constraints are given which ensure the simplified elephants retain these features with certain minimal persistence $\delta > 0$.

Note that neither persistence diagrams nor deformation-compatible simplicial maps were needed to perform the sparsification, just a priori knowledge of which areas of Δ_+ to check for the existence of persistent topological features. We defer a full comparison of the sparsification application as future work.

5.3 Topological time series analysis

The canonical lens by which time series analysis is performed is through harmonic analysis. Though mature and powerful, it can at times be more illuminating to study time series data in more geometric settings. A classical example of this is Takens delay embedding theorem, which provides sufficiency conditions under which a smooth attractor of d -dimensional manifold can be reconstructed from its dynamics $f : M \rightarrow M$. In this context, the “reconstruction” is a phase space embedding of a given time series topologically equivalent to the original space, as constructed through a *time-delay embedding* of f :

$$\text{SW}_{M,\tau} f(t) = (f(t) \ f(t+\tau) \ \cdots \ f(t+M\tau))^\top$$

where $M+1$ defines the *embedding dimension* into \mathbb{R}^{M+1} and $M\tau$ is the *window size* with respect to some $\tau > 0$. Choosing different values of t yields a collection of points called the sliding window point cloud of f . Under certain conditions, Takens theorem shows this reconstruction indeed has topological equivalence using the Whitney Embedding Theorem, motivating the study of the topology of $\text{SW}_{M,\tau} f$ itself. This methodology has been used to characterize non-linearity or chaotic behavior in ECG-EKG and MEG data.

The time-delay embedding requires fixing two parameters: the dimension $M \in \mathbb{N}$ and delay $\tau \in \mathbb{R}_+$. Takens theorem guarantees topological equivalence if M is more twice as large as d , thus choosing dimensions larger than $2d+1$ poses no problems beyond sampling density and computational concerns. In contrast, once M is fixed, the embedding delay τ strongly affects the shape of the corresponding embedding $\text{SW}_{M,\tau}$. Indeed, for general functions, the geometry of the curve $t \mapsto \text{SW}_{M,\tau} f(t)$ can be quite complicated. However, the analysis by [42] shows that if f is periodic, the roundness⁸ of the sliding window point cloud is maximized as the window-size approaches the length of the period—this periodicity can be recovered by maximizing the maximum 1-dimensional persistence of the sliding window Vietoris-Rips filtration:

$$\tau^* = \arg \max_{\tau \in \mathbb{R}_+} \left(\max_{(a,b) \in \mathcal{D}_{\tau,f}} |b-a| \right), \quad \mathcal{D}_{\tau,f} = \text{dgm}_1(\text{Rips}(\text{SW}_{M,\tau} f))$$

For more details on this connection, see Theorem 4.5 of [42]. On the positive side, the stability of persistence alongside the Whitney embedding theorem suggests that Equation 5.3 is a valid proxy objective to for the purpose of detecting periodic behavior in complex or noisy signals which might otherwise be difficult to detect using e.g.

⁸Here, *roundness* is defined as the largest radius of a ball in \mathbb{R}^{M+1} such that the curve $\text{SW}_{M,\tau}$ is tangent to at least two points from its equator.

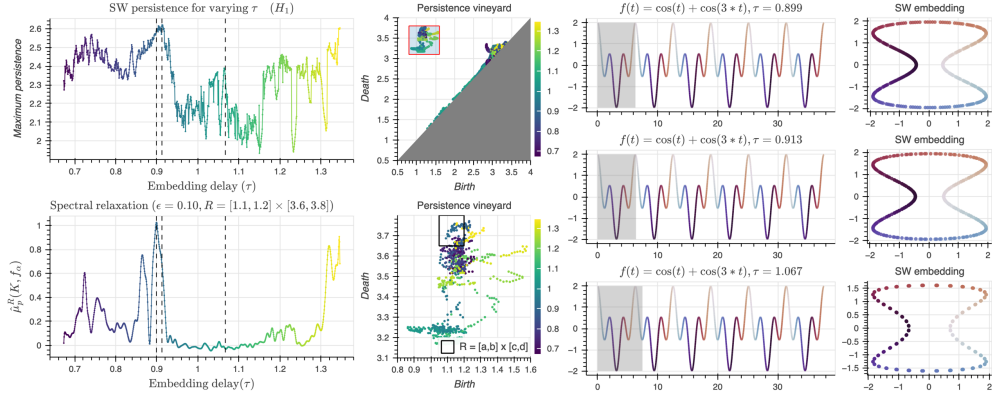


Fig. 5 Maximum SW 1-persistence for $\mathcal{D}_{\tau, f}$ at varying $\tau \in [2\pi c/8, 2\pi c/4]$ with $c = 7/8$ alongside the spectral multiplicity $\hat{\mu}_p^R(K, f_\alpha)$ for regularization $\varepsilon = 0.10$ and box $R = [1.1, 1.2] \times [3.6, 3.8]$ (left); two vineyard plots, one showing the region that determines maximum persistence (middle-top), the bottom showing the chosen box R (middle-bottom); time series $f(t)$ with window-lengths $M\tau$ highlighted for three values of τ and their embedding plotted with PCA (right).

autocorrelation. On the negative side, solving Equation 5.3 directly presents a computational complexity issue, due to the large size of Rips and the high asymptotic complexity of persistence.

To illustrate another application of our spectral relaxation, we consider an proxy optimization problem that is both easier and more efficient to optimize than Equation 5.3. We begin with a simple periodic function f given by:

$$f(t) = \cos(t) + \cos(3t), \quad t \in [0, 12\pi]$$

Regarding the parameterization, we consider optimizing τ over the interval $[2\pi c/8, 2\pi c/4]$ using the Tikhonov regularized spectral multiplicity $\hat{\mu}_p^R$ with a (large) regularization from Equation 3 with $\varepsilon = 0.10$. To determine the proxy objective function $\hat{\mu}_1^R$, we fix a rectangle $R = [1.1, 1.2] \times [3.6, 3.8] \in \Delta_+$, which we determined contains the correct maximizer by inspection of the persistence vineyard⁹. All other parameters determined, we plot both the (true) maximum persistence and the optimization surface of $\hat{\mu}_1^R$ on the top and bottom left subplot of Figure 5, respectively.

Analogous to the Takens and Nyquist-Shannon theorems, [42] show that for trigonometric polynomials, one loses no information if the embedding dimension is greater than twice the maximum frequency, thus we set $M = 7$ and sample $n = 300 \gg 2(3 \cdot 6)$ points to ensure no issues related dimension or sampling density occur. Then, we measure the maximum dimension-1 persistence of $\mathcal{D}_{\tau, f}$ across 900 equispaced delay values $\tau \in [0, 12\pi]$, which we plot on the top left of Figure 5. We observed the maximum persistence was attained at around $\tau \approx 0.913$; the theory

⁹In practice, the full vineyard wouldn't be available, and the box R would need to be determined using apriori knowledge or by probing random subsets of $R \subset \Delta_+$ for non-trivial persistence.

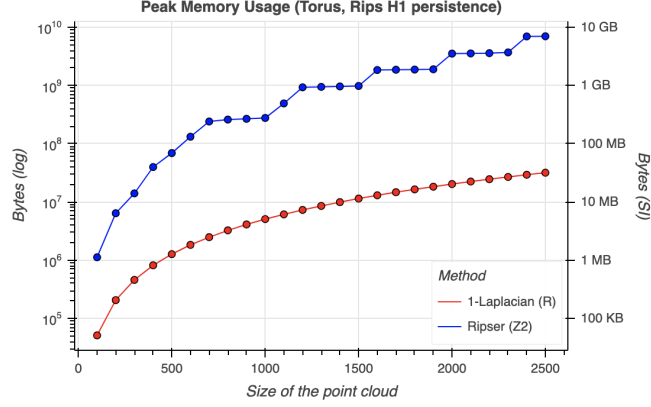


Fig. 6 (Blue) The “high watermark” maximum (heap) memory allocated by Ripser in computing persistence of a Rips filtration K up to the enclosing radius. (Red) the maximum memory allocated by the Jacobi matrix constructed from a degree- $|K|$ Lanczos expansion of $\mathcal{L}_p(K)$. Note that the Jacobi matrix of $\mathcal{L}_p(K)$ (or of some subset thereof) contains all the information needed to reconstruct the persistence invariants.

from [42] suggests that if f is L -periodic, the delay value τ which achieves the maximum persistence should be around $\tau = 2\pi c/L \approx 0.916$ where $c = M/(M+1)$ is the proportionality constant.

5.4 Low memory persistence computations

One particular limitation of the persistence computation is its space usage. Though the initial boundary matrices (when explicitly constructed) are known to be sparse, both the constitutive matrices storing the reduced boundary chains (R) and the cycle representatives (V) are known to lose their sparsity throughout the duration of the reduction. As the number of non-zeros in R affects the performance of reduction, analyzing the worst and average-case amount of “fill-in” is a subject of recent research [19]. Though the average case complexity is far better than the worst case, space usage is a well known to be one of the barriers antagonizing the scalability of the persistence computation. As the persistence diagram may be entirely determined by rank computations, the time and space complexity results from Proposition 3 extend directly to the computation of persistence diagrams via Chen and Kerbers divide-and-conquer approach [43].

To demonstrate the space efficiency of the matrix-free approach, we revisit the persistence computation in the context of the homology inference problem, one of the original use-cases of persistence [44]. To measure the scalability of the Lanczos method, we measure its space usage via the persistence computation on an increasingly dense point cloud sampled from the uniform measure on the 2-torus \mathbb{T} (via [45]):

$$f(\theta, \varphi) = ((R + r \cos(\theta)) \cos(\varphi), (R + r \cos(\theta)) \sin(\varphi), r \sin(\theta))$$

To ensure uniform coverage of \mathbb{T} , we build the point clouds using k -prefixes of the *greedy permutation* of a dense point sample, for k varying along evenly spaced points

from 50 to 2500 (yielding complexes with sizes ranging from $2.0 \cdot 10^4$ to $2.6 \cdot 10^9$ simplices, respectively). For each k , we compute both the persistence of the Rips filtration K up to the enclosing radius using the popular software *Ripser* [39] and the Jacobi matrix of the K ’s corresponding combinatorial Laplacian using the Lanczos method, up to degree $|K|$. To maximize space efficiency, like *Ripser*, we use the *apparent pairs* optimization to discard zero-persistence pairs (see Section 4.3). For a fair comparison in terms of space usage, we record the maximum (heap) memory allocated by tracing all system allocations using the software *Memray*¹⁰.

It’s worth noting the results in Figure 5.4 carry certain limitations. In particular, we only report the memory usage of both methods, rather than the time usage, as the method by Chen and Kerber has simply too large of a compute overhead to be practical. Additionally, *Ripser* is computing persistence here with respect to $\mathbb{Z}/2$ coefficients, whereas we rely on IEEE 64-bit floating precision arithmetic as a proxy for \mathbb{R} (thus the persistence diagrams are not identical). For more limitations, see Section 6.

5.5 Shape comparison via featurization

Analogous to its historical use in comparing shapes (i.e. manifolds, curves, etc.) through size functions, persistence has increasingly found use in machine learning pipelines through the use of mappings from diagrams to function spaces (e.g. Hilbert spaces). One notable such featurization is the persistent homology transform (PHT) [10], an injective transform which filters a triangulated space \mathcal{X} embedded in \mathbb{R}^d via the sublevel set parameterization $\mathcal{X}(v)_r = \{x \in \mathcal{X} : x \cdot v \leq r\}$, where $v \in S^{d-1}$. The intuition here is that by “looking” at the sublevel-set persistence of the shape \mathcal{X} from all possible directions $v \in S^1$, one can recover enough information to fully reconstruct \mathcal{X} from the collection of diagrams alone.

Aside from sparking an inverse theory for persistence, the injectivity of the PHT also provides the foundation necessary for equipping shape spaces with bona fide distance metrics. For example, one such metric is given by the integrated Wasserstein distance $\text{dist}_{\mathcal{X}}$ between the sublevel-set persistence diagrams $\text{dgm}_p(X, v)$:

$$\text{dist}_{\mathcal{X}}(X, X') = \sum_{p=0}^d \int_{S^{d-1}} d_W(\text{dgm}_p(X, v), \text{dgm}_p(X', v)) dv$$

where $X, X' \in \mathcal{X}$. In most machine learning applications, distance metrics are often difficult—if not intractable—to analytically determine; instead, most practitioners either craft or learn task-specific pseudometrics that are easy to compute and discriminative enough for the given application.

We seek to learn shape-sensitive, discriminative pseudometrics from weaker persistence invariants. Towards this end, consider the following parameterization, for some

¹⁰See <https://bloomberg.github.io/memray/memory.html> for an overview of how memory is measured.

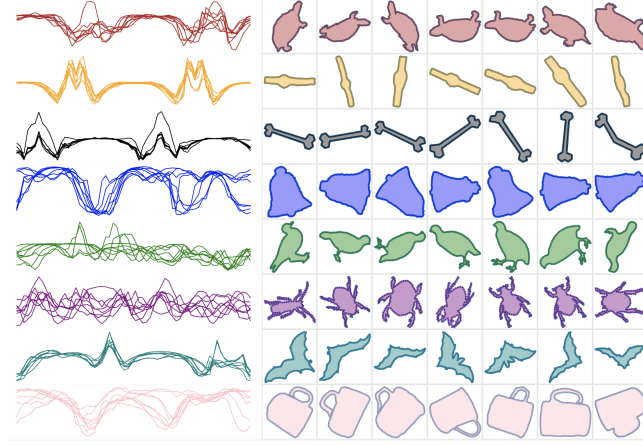


Fig. 7 Spectral signatures generated from filtering combinatorial Laplacians of MPEG7 2D shape datasets using the directional transform. By using the relaxation from Equation 3, these signatures can be interpreted as combinations of heat traces over the interval $[0, 2\pi)$. Observe that highly similar shapes generate curves that maintain a large degree of intra-class similarity under the ℓ_2 norm whereas outliers are detected and retain dissimilarity between classes (e.g. the bent bone in the black class exhibits higher peaks in its signature compared to its class, but is highly class-similar.)

fixed simplicial complex K and pair $(a, b) \in \Delta_+$:

$$\begin{aligned} \mathcal{B}_{a,b} : S^{d-1} &\rightarrow \mathbb{R} \\ v &\mapsto \beta_p^{a,b}(X, v) \end{aligned}$$

where $\beta_p^{a,b}(X, v) = |\{(a', b') \in \text{dgm}_p(K, v) : a' \leq a, b < b'\}|$. When $d = 2$, this is a piecewise-constant function periodic over the interval $[0, 2\pi)$. The collection of all such curves over Δ_+ characterizes all the information in PHT, thus it is a natural candidate hypothesis space for learning algorithms. However, some choices of pairs $(a, b) \in \Delta_+$ may yield functions which have little-or-no information (e.g. identically zero or constant functions), suggesting that some choices of $(a, b) \in \Delta_+$ may be more useful than others for the purpose of classification.

To demonstrate another application of our methodology for shape classification, we apply our spectral relaxation to the MPEG-7 shape matching data set from [46]. To build a classifier, we generate curves $\mathcal{B}_{a,b} = \{\hat{\beta}_p^{a,b}(K, v) : v \in S^1\}$ for random $(a, b) \in \Delta_+$, which act as hypotheses functions $h_*(x)$ in an ensemble model H :

$$H(x) = \text{sign}(\alpha_1 h_1(x) + \alpha_2 h_2(x) + \cdots + \alpha_t h_t(x))$$

To determine the coefficients α_* , we use the classical AdaBoost method [47], sampling new hypothesis functions h_* uniformly at random from a bounded subset of Δ_+ . We choose the spectral relaxation $\phi(\lambda, \tau) = 1 - \exp(-\lambda/\tau)$ with a smoothing parameter $\tau = 1 \cdot 10^{-1}$. Following Equation 3, the corresponding curves $\mathcal{B}_{a,b}$ may be interpreted as combinations of *heat kernel traces*, which are known to be useful for characterizing graphs [48].

In Figure 7, we plot the curves associated to one such hypothesis function generated for the ensemble. Observe that highly similar shapes generate curves that maintain a large degree of intra-class similarity under the ℓ_2 norm whereas outliers are detected and retain dissimilarity between classes (e.g. the bent bone in the black class exhibits higher peaks in its signature compared to its class, but is highly class-similar). This suggests that, for some choice $(a, b) \in \Delta_+$ and certain choices of ϕ , composing the spectrally-relaxed persistent rank function with the direction transform is viable strategy for generating *weak learners* for machine learning purposes.

6 Concluding Remarks and Limitations

In summary, we have introduced various spectral relaxations of the persistent rank function, which has a wealth of useful properties, including differentiability on the positive semi-definite cone, a matrix-free representation, and natural connections to variational constructions common in machine learning applications, such as the Tikhonov regularization and the Heat kernel. By focusing on coefficients in \mathbb{R} , we were able to exploit the connection between the inner product spaces on cochain spaces and the theory of persistence measures, providing an avenue to introduce differentiability to an otherwise discontinuous function. Moreover, by focusing on the spectral characterization of the rank function, we were able to study persistence in the setting of *iterative methods*, a pursuit which lead us to the Lanczos method of Krylov expansion. Surprisingly, these computational techniques turned out to pair well with the output-sensitive algorithm from [20], which we demonstrated may be used to compute Γ -persistence pairs in an output sensitive fashion. As cycle representatives can be obtained once a pairing has been constructed, the iterative approach may be used to compute essentially any persistence invariant.

6.1 Limitations

There are few limitations to the proposed work that may prevent its practical use, depending on the problem and invariant of interest. We mention the limitations to the approaches proposed that we are aware of below.

Function parameterization: Choosing the matrix function Φ from Equation 3 and tuning any of its associated hyper-parameters (i.e. τ and ω) can be very application-dependent. We demonstrate this in Figure 5.1—for some values of $\tau > 0$, the relaxation can be either have high variance due to instability or high bias due over smoothing. In the optimization setting, we addressed these issues via an iterative shrinkage approach; in the more general setting, these choices depend on a multitude of factors, such as the relative precision the target invariant needs to achieve, the amount of compute resources available, etc.

Rank instability: Though the persistence rank invariant is considered a “stable function” in the multi-parameter setting [2], this stability is generally proven by reduction to a matching-type distance between the pairings themselves. In particular, the Betti number function β_p^* and its various related invariants—such as the

sequence containing the Betti numbers of the homology groups at all scales (the “*Betti sequence*”)—is in fact *unstable* with respect to the 1-Wasserstein distance (see Theorem 2.5 in [49]). We sought to address this instability by a spectral approach, but other approaches that could be used; for example, both [49] and [7] provide a Gaussian-smoothing type stabilization procedure, though they also both require access to the pairings. Another approach—which depends only on the rank—is the *hierarchical stabilization* method from [50], though this remains unexplored.

Randomized estimation: Computing some of the persistence invariants exactly—such as the pairing itself—poses some difficulty in practice, especially when the exact rank is required. For example, in the divide-and-conquer algorithm from [20], a single wrong rank value can lead to unpredictable output. If a randomized estimator with fixed success probability $p > 0$ is used, then one may need to be re-run k times to achieve a higher success probability $(1 - (1 - p)^k)$. Fortunately, for any constant $\delta \in (0, 1)$, Lemma 10 of [20] shows the number of such rank computations needed to compute Γ -persistent pairs is not larger than:

$$\rho = \rho(n, \left\lceil \frac{1}{\delta} \right\rceil) = 4n \left(2 \left\lceil \frac{1}{\delta} \right\rceil + \log n + 1 \right)$$

Indeed, the analysis from section 6 of [20] shows that if the rank estimator has success probability p and $\hat{p} = 1 - p$, then for the entire algorithm to achieve an arbitrary success probability $q > 0$, every rank computation needs to be repeated at most k times, where:

$$\log(1/\hat{p})^{-1} \cdot \log(\omega) \leq k, \quad \omega = \left(1 - q^{\frac{1}{\hat{p}}}\right)^{-1}$$

Since $\log(\omega) = \Theta(\rho)$, every rank computation requires $O(\log \rho) = O(\log n + \log \frac{1}{\delta})$ repeated trials. In other words, the randomized aspect of the algorithm does not significantly affect its computational complexity, though it does increase the complexity of a practical implementation.

Estimator efficiency: In theory, perhaps the largest limitation of approximating the rank function via the GH Monte-carlo estimator via Equation 4.2 is its convergence rate, which is dominated by $O(\epsilon^{-2})$. If ϵ is small, the number of iterations the estimator needs to *provably* obtain a $(1 \pm \epsilon)$ -approximation can be astronomical. To make matters worse, estimating the numerical rank may require shrinking ϵ to the order of $O(1/n)$. Thus its practical utility lies in applications where ϵ need not be too small, i.e. only a relatively coarse approximation of $\text{tr}(f(A))$ is needed. From our view, this slow convergence and high precision combination of the GH estimator Equation 4.2 is the only major barrier preventing the use of matrix-free techniques in more practical application settings.

In practice, the efficiency of the iterative implicit trace estimator from Section 4.2 depends on a variety of nuanced factors, such as the condition number of the operator, the spectral gap, etc. Fortunately, implicit trace estimation is an active area of research; recent results by Meyers et al. [38], for example, show how the estimator Equation 4.2 can be modified to obtain a $(1 \pm \epsilon)$ approximation using only

$O\left(\varepsilon^{-1} \cdot \sqrt{\log(\eta^{-1})} + \log(\eta^{-1})\right)$ samples using deflation techniques, though this does require additional re-orthogonalization and storage costs. Even more recently, optimal trace estimators based on the *exchangeability principle* and the *Nyström approximation* have been shown to not only match the $O(1/n^2)$ variance reduction rate of Hutch++, but to also empirically achieve much faster convergence [51].

Interestingly, using a reduction to the Gap-Hamming problem, it was shown that achieving a $(1 \pm \epsilon)$ trace approximation with probability greater than $3/4$ requires at least $\Omega(\epsilon^{-1})$ matrix-vector queries $v \mapsto Av$ if the input vectors $v \in \mathbb{R}^n$ are chosen *non-adaptively* [38] (i.e. independent of the structure of A). This lower bound suggests it may be possible to achieve a $O(1 \pm \epsilon)$ approximation of the rank-based invariants from Proposition 3 with the same space complexity and without exact arithmetic, but no algorithm is known to the author that achieves this non-adaptive bound.

Supplementary information.

Acknowledgements.

Declarations

Appendix A

A.1 Proofs

Proof of Lemma 1 The Pairing Uniqueness Lemma[1] asserts that if $R = \partial V$ is a decomposition of the total $m \times m$ boundary matrix ∂ , then for any $1 \leq i < j \leq m$ we have $\text{low}_R[j] = i$ if and only if $r_\partial(i, j) = 1$. As a result, for $1 \leq i < j \leq m$, we have:

$$\text{low}_R[j] = i \Leftrightarrow r_R(i, j) \neq 0 \Leftrightarrow r_\partial(i, j) \neq 0$$

Extending this result to Equation 6 can be seen by observing that in the decomposition, $R = \partial V$, the matrix V is full-rank and obtained from the identity matrix I via a sequence of rank-preserving (elementary) left-to-right column additions. \square

Proof of Proposition 1 We first need to show that $\beta_p^{i,j}$ can be expressed as a sum of rank functions. Note that by the rank-nullity theorem, so we may rewrite Equation 2.2 as:

$$\beta_p^{i,j} = \dim(C_p(K_i)) - \dim(B_{p-1}(K_i)) - \dim(Z_p(K_i) \cap B_p(K_j))$$

The dimensions of groups $C_p(K_i)$ and $B_p(K_i)$ are given directly by the ranks of diagonal and boundary matrices, yielding:

$$\beta_p^{i,j} = \text{rank}(I_p^{1,i}) - \text{rank}(\partial_p^{1,i}) - \dim(Z_p(K_i) \cap B_p(K_j))$$

To express the intersection term, note that we need to find a way to express the number of p -cycles born at or before index i that became boundaries before index j . Observe that the non-zero columns of R_{p+1} with index at most j span $B_p(K_j)$, i.e. $\{\text{col}_{R_{p+1}}[k] \neq 0 \mid k \in [j]\} \in \text{Im}(\partial_{p+1}^j)$. Now, since the low entries of the non-zero columns of R_{p+1} are unique, we have:

$$\dim(Z_p(K_i) \cap B_p(K_j)) = |\Gamma_p^{i,j}|$$

where $\Gamma_p^{i,j} = \{ \text{col}_{R_{p+1}}[k] \neq 0 \mid 1 \leq \text{low}_{R_{p+1}}[k] \leq i \}$. Consider the complementary matrix $\bar{\Gamma}_p^{i,j}$ given by the non-zero columns of R_{p+1} with index at most j that are not in $\Gamma_p^{i,j}$, i.e. the columns satisfying $\text{low}_{R_{p+1}}[k] > i$. Combining rank-nullity with the observation above, we have:

$$\bar{\Gamma}_p^{i,j} = \dim(B_p(K_j)) - |\Gamma_p^{i,j}| = \text{rank}(R_{p+1}^{i+1,j})$$

Combining equations Equation A.1 with Equation A.1 yields:

$$\dim(Z_p(K_i) \cap B_p(K_j)) = |\Gamma_p(i,j)| = \dim(B_p(K_j)) - |\bar{\Gamma}_p^{i,j}| = \text{rank}(R_{p+1}^{1,j}) - \text{rank}(R_{p+1}^{i+1,j})$$

Observing the final matrices in Equation A.1 are *lower-left* submatrices of R_{p+1} , the final expression Equation 7 follows by applying Lemma 1 repeatedly. \square

A.2 Combinatorial Laplacians

The natural extension of the graph Laplacian L to simplicial complexes is the p -th *combinatorial Laplacian* Δ_p , whose explicit matrix representation is given by Equation 2.3. Indeed, when $p = 0$, $\Delta_0(K) = \partial_1 \partial_1^T = L$ recovers the graph Laplacian. As with boundary operators, $\Delta_p(K)$ encodes simplicial homology groups in its nullspace, a result known as the discrete Hodge Theorem [52]:

$$\tilde{H}_p(K; \mathbb{R}) \simeq \ker(\Delta_p(K)), \quad \beta_p = \text{nullity}(\Delta_p(K))$$

The fact that the Betti numbers of K may be recovered via the nullity of $\Delta_p(K)$ has been well studied (see e.g. Proposition 2.2 of [25]). In fact, one need not only consider Δ_p as the spectra of Δ_p , L_p^{up} , and L_p^{dn} are intrinsically related by the identities:

$$\Lambda(\Delta_p(K)) \doteq \Lambda(L_p^{\text{up}}) \cup \Lambda(L_p^{\text{dn}}), \quad \Lambda(L_p^{\text{up}}) \doteq \Lambda(L_{p+1}^{\text{dn}})$$

where $A \doteq B$ and $A \cup B$ denotes equivalence and union between the *non-zero* elements of the multisets A and B , respectively, counted with multiplicity. Moreover, all three operators Δ_p , L_p^{up} , and L_p^{dn} are symmetric, positive semidefinite, and compact—thus, for the purpose of estimating β_p , it suffices to consider only one family of operators.

A.3 Laplacian matvec

Given a simple undirected graph $G = (V, E)$, let $A \in \{0, 1\}^{n \times n}$ denote its binary adjacency matrix satisfying $A[i, j] = 1 \Leftrightarrow i \sim j$ if the vertices $v_i, v_j \in V$ are adjacent in G , and let $D = \text{diag}(\{\deg(v_i)\})$ denote the diagonal *degree* matrix, where $\deg(v_i) = \sum_{j \neq i} A[i, j]$. The *graph Laplacian's* adjacency, incidence, and element-wise definitions are:

$$L = D - A = \partial_1 \circ \partial_1^T, \quad L[i, j] = \begin{cases} \deg(v_i) & \text{if } i = j \\ -1 & \text{if } i \sim j \\ 0 & \text{if } i \not\sim j \end{cases}$$

By using the adjacency relation $i \sim j$ as in [53], the linear and quadratic forms of L may be succinctly expressed as:

$$L(x)_i = \deg(v_i) \cdot x_i - \sum_{i \sim j} x_j, \quad x^T Lx = \sum_{i \sim j} (x_i - x_j)^2$$

If G has m edges and n vertices taking labels in the set $[n]$, observe computing the product from Equation A.3 requires just $O(m)$ time and $O(n)$ storage via two edge traversals: one to accumulate vertex degrees and one to remove components from incident edges. By precomputing the degrees, the operation reduces further to a single $O(n)$ product and $O(m)$ edge pass, which is useful when repeated evaluations for varying values of x are necessary.

To extend the two-pass algorithm outlined above for $p > 0$, we first require a generalization of the connected relation $i \sim j$ from Equation A.3. Denote with $\text{co}(\tau) = \{\sigma \in K^{p+1} \mid \tau \subset \sigma\}$ the set of proper cofaces of $\tau \in K^p$, or *cofacets*, and the (weighted) *degree* of $\tau \in K^p$ with:

$$\deg_w(\tau) = \sum_{\sigma \in \text{co}(\tau)} w(\sigma)$$

Note setting $w(\sigma) = 1$ for all $\sigma \in K$ recovers the integral notion of degree representing the number of cofacets a given p -simplex has. Now, since K is a simplicial complex, if the faces τ, τ' share a common cofacet $\sigma \in K^{p+1}$, this cofacet σ is in fact *unique* in the sense that $\{\sigma\} = \text{co}(\tau) \cap \text{co}(\tau')$ (see [54] for more details). Thus, we may use a relation $\tau \stackrel{\sigma}{\sim} \tau'$ to rewrite L_p^{up} element-wise:

$$L_p^{\text{up}}(\tau, \tau') = \begin{cases} \deg_w(\tau) \cdot w^+(\tau) & \text{if } \tau = \tau' \\ s_{\tau, \tau'} \cdot w^{+/2}(\tau) \cdot w(\sigma) \cdot w^{+/2}(\tau') & \text{if } \tau \stackrel{\sigma}{\sim} \tau' \\ 0 & \text{otherwise} \end{cases}$$

where $s_{\tau, \tau'} = \text{sgn}([\tau], \partial[\sigma]) \cdot \text{sgn}([\tau'], \partial[\sigma])$. Ordering the p -faces $\tau \in K^p$ along a total order and choosing an indexing function $h : K^p \rightarrow [n]$ enables explicit computation of the corresponding matrix-vector product:

$$(L_p^{\text{up}} x)_i = \deg_w(\tau_i) \cdot w^+(\tau_i) \cdot x_i + w^{+/2}(\tau_i) \sum_{\tau_j \stackrel{\sigma}{\sim} \tau_i} s_{\tau_i, \tau_j} \cdot x_j \cdot w(\sigma) \cdot w^{+/2}(\tau_j)$$

Observe $v \rightarrow L_p^{\text{up}} v$ can be evaluated via a similar two-pass algorithm as described for the graph Laplacian by simply enumerating the boundary chains of the simplices of K^{p+1} .

A.4 Choosing a weight function

Given Equation 2.3, a natural question to ask whether there exists a weight function w that is more computationally amenable for rank computations. For the purpose of improving the condition number of the underlying Laplacian, consider the following optimization problem:

$$\begin{aligned} & \max_{w \in \mathbb{R}^n} \quad \lambda_{\min}(\mathcal{L}_p(w)) \\ & \text{subject to} \quad w > 0, \mathbf{1}^T w = 1 \end{aligned}$$

In spectral graph theory, a specialization of this problem (for $p = 1$) arises under the guise of related problems, such as maximizing *algebraic connectivity* under a fixed total edge weight or finding the *fastest mixing time* of a Markov process. Fortunately, Equation A.4 is a convex optimization problem, which can be formulated as a semi-definite program (SDP) with variables $\gamma, \beta \in \mathbb{R}, w \in \mathbb{R}^n$ under positivity and sum-to-one constraints:

$$\begin{aligned} & \max_{\gamma \in \mathbb{R}} \quad \gamma \\ & \text{subject to} \quad \gamma I \preceq \partial_1 D_1(w) \partial_1^T + \beta \mathbf{1} \mathbf{1}^T \\ & \quad \mathbf{1}^T w = 1, \quad w > 0 \end{aligned}$$

In practice, we found the weight function w^* that optimizes Equation A.4 indeed tends to be lead to less iterations than e.g. random weights, though we leave further investigation of this open ended.

References

- [1] Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. In: Proceedings 41st Annual Symposium on Foundations of Computer Science, pp. 454–463 (2000). IEEE
- [2] Cerri, A., Fabio, B.D., Ferri, M., Frosini, P., Landi, C.: Betti numbers in multi-dimensional persistent homology are stable functions. *Mathematical Methods in the Applied Sciences* **36**(12), 1543–1557 (2013)
- [3] Cohen-Steiner, D., Edelsbrunner, H., Harer, J.: Stability of persistence diagrams. In: Proceedings of the Twenty-first Annual Symposium on Computational Geometry, pp. 263–271 (2005)
- [4] Chazal, F., Cohen-Steiner, D., Guibas, L.J., Mémoli, F., Oudot, S.Y.: Gromov-hausdorff stable signatures for shapes using persistence. In: Computer Graphics Forum, vol. 28, pp. 1393–1403 (2009). Wiley Online Library
- [5] Scoccola, L., Perea, J.A.: Fibered: Fiberwise dimensionality reduction of topologically complex data with vector bundles. In: 39th International Symposium on Computational Geometry (SoCG 2023) (2023). Schloss Dagstuhl-Leibniz-Zentrum für Informatik
- [6] Perea, J.A.: Persistent homology of toroidal sliding window embeddings. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6435–6439 (2016). IEEE

- [7] Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushtanova, S., Hanson, E., Motta, F., Ziegelmeier, L.: Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research* **18** (2017)
- [8] Bubenik, P., *et al.*: Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.* **16**(1), 77–102 (2015)
- [9] Perea, J.A., Munch, E., Khasawneh, F.A.: Approximating continuous functions on persistence diagrams using template functions. *Foundations of Computational Mathematics*, 1–58 (2022)
- [10] Turner, K., Mukherjee, S., Boyer, D.M.: Persistent homology transform for modeling shapes and surfaces. *Information and Inference: A Journal of the IMA* **3**(4), 310–344 (2014)
- [11] Piekenbrock, M., Perea, J.A.: Move schedules: Fast persistence computations in coarse dynamic settings. *arXiv preprint arXiv:2104.12285* (2021)
- [12] Chazal, F., De Silva, V., Glisse, M., Oudot, S.: *The Structure and Stability of Persistence Modules* vol. 10. Springer, ??? (2016)
- [13] Edelsbrunner, H., Harer, J.L.: *Computational Topology: an Introduction*. American Mathematical Society, ??? (2022)
- [14] Zomorodian, A., Carlsson, G.: Computing persistent homology. In: *Proceedings of the Twentieth Annual Symposium on Computational Geometry*, pp. 347–356 (2004)
- [15] Kim, W., Mémoli, F.: Spatiotemporal persistent homology for dynamic metric spaces. *Discrete & Computational Geometry* **66**, 831–875 (2021)
- [16] Lesnick, M., Wright, M.: Interactive visualization of 2-d persistence modules. *arXiv preprint arXiv:1512.00180* (2015)
- [17] Cohen-Steiner, D., Edelsbrunner, H., Morozov, D.: Vines and vineyards by updating persistence in linear time. In: *Proceedings of the Twenty-second Annual Symposium on Computational Geometry*, pp. 119–126 (2006)
- [18] Dey, T.K., Wang, Y.: *Computational Topology for Data Analysis*. Cambridge University Press, ??? (2022)
- [19] Bauer, U., Masood, T.B., Giunti, B., Houry, G., Kerber, M., Rathod, A.: Keeping it sparse: Computing persistent homology revised. *arXiv preprint arXiv:2211.09075* (2022)
- [20] Chen, C., Kerber, M.: An output-sensitive algorithm for persistent homology. In: *Proceedings of the Twenty-seventh Annual Symposium on Computational*

Geometry, pp. 207–216 (2011)

- [21] Bhatia, R.: Matrix Analysis vol. 169. Springer, ??? (2013)
- [22] Teng, S.-H.: The laplacian paradigm: Emerging algorithms for massive graphs. In: Theory and Applications of Models of Computation: 7th Annual Conference, TAMC 2010, Prague, Czech Republic, June 7–11, 2010. Proceedings 7, pp. 2–14 (2010). Springer
- [23] Ubaru, S., Saad, Y.: Fast methods for estimating the numerical rank of large matrices. In: International Conference on Machine Learning, pp. 468–477 (2016). PMLR
- [24] Mémoli, F., Wan, Z., Wang, Y.: Persistent laplacians: Properties, algorithms and implications. SIAM Journal on Mathematics of Data Science **4**(2), 858–884 (2022)
- [25] Horak, D., Jost, J.: Spectra of combinatorial laplace operators on simplicial complexes. Advances in Mathematics **244**, 303–336 (2013)
- [26] Mangasarian, O., Chen, C.: A class of smoothing functions for nonlinear and mixed complementarity problems. Technical report (1994)
- [27] Bi, S., Han, L., Pan, S.: Approximation of rank function and its application to the nearest low-rank correlation matrix. Journal of Global Optimization **57**(4), 1113–1137 (2013)
- [28] Zhao, Y.-B.: An approximation theory of matrix rank minimization and its application to quadratic equations. Linear Algebra and its Applications **437**(1), 77–93 (2012)
- [29] Lanczos, C.: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators (1950)
- [30] Simon, H.D.: Analysis of the symmetric lanczos algorithm with reorthogonalization methods. Linear algebra and its applications **61**, 101–131 (1984)
- [31] Parlett, B.N.: Do we fully understand the symmetric lanczos algorithm yet. Brown et al **3**, 93–107 (1994)
- [32] Golub, G.H., Van Loan, C.F.: Matrix Computations. JHU press, ??? (2013)
- [33] Edelsbrunner, H., Parsa, S.: On the computational complexity of betti numbers: reductions from matrix rank. In: Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 152–160 (2014). SIAM
- [34] Sorensen, D.: Implicitly restarted arnoldi/lanczos methods and large scale svd applications. SVD and Signal Processing III, 21–31 (1995)

- [35] Lehoucq, R.B., Sorensen, D.C., Yang, C.: ARPACK Users' Guide: Solution of Large-scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. SIAM, ??? (1998)
- [36] Musco, C., Musco, C., Sidford, A.: Stability of the lanczos method for matrix function approximation. In: Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1605–1624 (2018). SIAM
- [37] Paige, C.C.: Computational variants of the lanczos method for the eigenproblem. IMA Journal of Applied Mathematics **10**(3), 373–381 (1972)
- [38] Meyer, R.A., Musco, C., Musco, C., Woodruff, D.P.: Hutch++: Optimal stochastic trace estimation. In: Symposium on Simplicity in Algorithms (SOSA), pp. 142–155 (2021). SIAM
- [39] Bauer, U.: Ripser: efficient computation of vietoris–rips persistence barcodes. Journal of Applied and Computational Topology **5**(3), 1–33 (2021)
- [40] Heckbert, P.S., Garland, M.: Survey of polygonal surface simplification algorithms. (1997). Siggraph
- [41] Fugacci, U., Kerber, M., Manet, H.: Topology-preserving terrain simplification. In: Proceedings of the 28th International Conference on Advances in Geographic Information Systems, pp. 36–47 (2020)
- [42] Perea, J.A., Harer, J.: Sliding windows and persistence: An application of topological methods to signal analysis. Foundations of computational mathematics **15**, 799–838 (2015)
- [43] Chen, C., Kerber, M.: An output-sensitive algorithm for persistent homology. Computational Geometry **46**(4), 435–447 (2013)
- [44] Perea, J.A.: A brief history of persistence. arXiv preprint arXiv:1809.03624 (2018)
- [45] Diaconis, P., Holmes, S., Shahshahani, M.: Sampling from a manifold. In: Advances in Modern Statistical Theory and Applications: a Festschrift in Honor of Morris L. Eaton vol. 10, pp. 102–126. Institute of Mathematical Statistics, ??? (2013)
- [46] Bai, X., Yang, X., Latecki, L.J., Liu, W., Tu, Z.: Learning context-sensitive shape similarity by graph transduction. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(5), 861–874 (2009)
- [47] Mukherjee, I., Rudin, C., Schapire, R.E.: The rate of convergence of adaboost. The Journal of Machine Learning Research **14**(1), 2315–2347 (2013)
- [48] Xiao, B., Hancock, E.R., Wilson, R.C.: Graph characteristics from the heat kernel trace. Pattern Recognition **42**(11), 2589–2606 (2009)

- [49] Johnson, M., Jung, J.-H.: Instability of the betti sequence for persistent homology and a stabilized version of the betti sequence. arXiv preprint arXiv:2109.09218 (2021)
- [50] Gäfvert, O., Chachólski, W.: Stable invariants for multiparameter persistence. arXiv preprint arXiv:1703.03632 (2017)
- [51] Epperly, E.N., Tropp, J.A., Webber, R.J.: Xtrace: Making the most of every sample in stochastic trace estimation. SIAM Journal on Matrix Analysis and Applications **45**(1), 1–23 (2024)
- [52] Lim, L.-H.: Hodge laplacians on graphs. Siam Review **62**(3), 685–715 (2020)
- [53] Chung, F.R.: Spectral Graph Theory vol. 92. American Mathematical Soc., ??? (1997)
- [54] Goldberg, T.E.: Combinatorial laplacians of simplicial complexes. Senior Thesis, Bard College **6** (2002)