

1. Motivation

< insert motivating examples, etc >

2. Background & Notation

A *simplicial complex* $K \subseteq \mathcal{P}(V)$ over a vertex set $V = \{v_1, v_2, \dots, v_n\}$ is a collection of simplices $\{\sigma : \sigma \in \mathcal{P}(V)\}$ such that $\tau \subseteq \sigma \in K \implies \tau \in K$. A *filtration* $K_\bullet = \{K_i\}_{i \in I}$ of a simplicial complex indexed by a totally ordered set I is a family of complexes such that $i < j \in I \implies K_i \subseteq K_j$. K_\bullet is called *simplexwise* if $K_j \setminus K_i = \{\sigma_j\}$, and K_\bullet is *essential* if $i \neq j$ implies $K_i \neq K_j$ whenever j is the immediate successor of i in I :

$$\emptyset = K_0 \subsetneq K_1 \subsetneq \dots \subsetneq K_m = K_\bullet, \quad K_i = K_{i-1} \cup \{\sigma_i\} \quad (1)$$

Filtrations may be equivalently defined as functions $f : K \rightarrow I$ satisfying $f(\tau) \leq f(\sigma)$ whenever $\tau \subseteq \sigma$. Here, we consider two index sets for I : \mathbb{R} and $[n] = \{1, \dots, n\}$. Any finite filtration may be trivially converted into an essential, simplexwise filtration via a set of *condensing*, *refining*, and *reindexing* maps [?]. Thus, without loss of generality, we exclusively consider essential simplexwise filtrations and for brevity refer to them as filtrations.

For K a simplicial complex and \mathbb{F} a field, a p -chain is a formal \mathbb{F} -linear combination of p -simplices of K . The collection of p -chains under addition yields an \mathbb{F} -vector space denoted $C_p(K)$. 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:

$$\partial_p(\sigma) = \partial_p([v_0, v_1, \dots, v_p]) = \sum_{i=0}^p (-1)^i [v_0, \dots, \hat{v}_i, \dots, v_p] \quad (2)$$

where \hat{v}_i indicates the removal of v_i from the i th summand. Similarly, the p -boundary of a p -chain is defined linearly in terms of its constitutive simplices. A p -chain with zero boundary is called a p -cycle, and together they form $Z_p(K) = \text{Ker } \partial_p$. Similarly, the collection of p -boundaries forms $B_p(K) = \text{Im } \partial_{p+1}$. Since $\partial_p \circ \partial_{p+1} = 0$ for all $p \geq 0$, the quotient space $H_p(K) = Z_p(K)/B_p(K)$ is well-defined, and $H_p(K)$ is called the p -th homology of K with coefficients in \mathbb{F} . The dimension of the p -th homology group $\beta_p(K) = \dim(H_p(K))$ of K is called the p -th *Betti number* of K .

Let $K_\bullet = \{K_i\}_{i \in [m]}$ denote a filtration of size $|K_\bullet| = m$. For every pair $i, j \in [m]$ with $i < j$, the inclusions $K_i \subsetneq K_{i+1} \subsetneq \dots \subsetneq K_j$ induce linear transformations $h_p^{i,j}$ at the level of homology:

$$0 = H_p(K_0) \rightarrow \dots \rightarrow H_p(K_i) \xrightarrow[h_p^{i,j}]{} \dots \rightarrow H_p(K_j) \rightarrow \dots \rightarrow H_p(K_m) = H_p(K_\bullet) \quad (3)$$

When \mathbb{F} is a field, this sequence of homology groups admits a unique decomposition of K_\bullet into a pairing of simplices (σ_i, σ_j) demarcating the evolution of homology classes: σ_i marks the creation of a homology class, σ_j marks its destruction, and the difference $|i - j|$ records the lifetime of the class, called its *persistence*. The p -th persistent homology groups are the images of these transformations and the p -th persistent Betti numbers are their dimensions:

$$H_p^{i,j} = \begin{cases} H(K_i) & i = j \\ \text{Im } h_p^{i,j} & i < j \end{cases}, \quad \beta_p^{i,j} = \begin{cases} \beta_p(K_i) & i = j \\ \dim(H_p^{i,j}) & i < j \end{cases} \quad (4)$$

For a fixed $p \geq 0$, the collection of persistent pairs (i, j) together with unpaired simplices (l, ∞) form a summary representation $\text{dgm}_p(K_\bullet)$ called the p -th *persistence diagram* of K_\bullet .

Remark 1. In practice, filtrations often arise from triangulations parameterized by geometric scaling parameters, and references to the “persistence” of a given homology class are with respect to these parameterizations. For example, given a finite metric space $\mathcal{X} = (X, d_X)$, the *Rips complex* at scale $\epsilon \in \mathbb{R}_+$ is the complex given by:

$$\text{Rips}_\epsilon(\mathcal{X}) := \{\sigma \subseteq X : d_X(x, x') \leq \epsilon \text{ for all } x, x' \in \sigma\} \quad (5)$$

Connecting successive complexes via inclusions $\text{Rips}_\epsilon(\mathcal{X}) \hookrightarrow \text{Rips}_{\epsilon'}(\mathcal{X})$ for $\epsilon < \epsilon'$ yields a family of complexes $\text{Rips}_\alpha := \{\text{Rips}_\epsilon(\mathcal{X})\}_{\epsilon \leq \alpha}$ called the *Rips filtration*. As in equation (??), these inclusions induce linear maps at level of homology. Though we consider primarily Rips filtrations in this effort, we will at times keep the notation simple and general by letting K_\bullet denote any simplicial filtration.

3. Methodology

In this section, we derive the relaxed objective function we seek to maximize.

Let $\delta_{\mathcal{X}}$ denote an T -parameterized metric space $\delta_{\mathcal{X}}(\cdot) = (X, d_{\mathcal{X}}(\cdot))$, where $d_{\mathcal{X}} : T \times X \times X \rightarrow \mathbb{R}_+$ is called a *time-varying metric* and X is a finite set with fixed cardinality $|X| = n$. $\delta_{\mathcal{X}}$ as called a *dynamic metric space* (DMS) iff $d_{\mathcal{X}}(\cdot)(x, x')$ is continuous for every pair $x, x' \in X$ and $\delta_{\mathcal{X}}(t) = (X, d_{\mathcal{X}}(t))$ is a pseudo-metric space for every $t \in T$. For a fixed $t \in T$, the Rips complex at scale $\epsilon \in \mathbb{R}$ is the abstract simplicial complex given by

$$\text{Rips}_{\epsilon}(\delta_{\mathcal{X}}(t)) := \{\sigma \subset X : d_{\mathcal{X}}(t)(x, x') \leq \epsilon \text{ for all } x, x' \in \sigma\} \quad (6)$$

As before, the family of Rips complexes for varying $\epsilon > 0$ yields a filtration whose inclusion maps induce linear maps at the level of homology. The time-varying counterpart is analogous. In this context, we write the p -th persistent Betti number with respect to fixed values $i, j \in I$ as a function of $t \in T$:

$$\beta_p^{i,j}(t) = (\dim \circ H_p^{i,j} \circ \text{Rips} \circ \delta_{\mathcal{X}})(t) \quad (7)$$

This quantity can be readily visualized as the number of persistent pairs lying inside the box $[0, i] \times (j, \infty)$, representing the persistent homology groups which were born at or before i and died sometime after j . We consider the problem of maximizing the p -th *persistent* Betti number $\beta_p^{i,j}$ over some set $T \subseteq T$:

$$t_* = \arg \max_{t \in T} \beta_p^{i,j}(t) \quad (8)$$

As an illustrative example, see Figure. < insert SW1Pers vineyards plot >

Persistent Betti Numbers:

As in section ??, let $B_p(K_*) \subseteq Z_p(K_*) \subseteq C_p(K_*)$ denote the p -th boundary, cycle, and chain groups of K_* , respectively. Given a simplicial filtration K_{\bullet} , let $\partial_p : C_p(K_{\bullet}) \rightarrow C_p(K_{\bullet})$ denote the boundary operator sending p -chains to their respective boundaries. With a slight abuse of notation, we also use ∂_p to also denote the filtration boundary matrix with respect to an ordered basis $(\sigma_i)_{1 \leq i \leq m_p}$. The p -th persistent Betti number between scales (i, j) is defined as:

$$\begin{aligned} \beta_p^{i,j} &= \dim(H_p^{i,j}) \\ &= \dim(Z_p(K_i) / (Z_p(K_i) \cap B_p(K_j))) \\ &= \dim(Z_p(K_i)) - \dim(Z_p(K_i) \cap B_p(K_j)) \end{aligned} \quad (9)$$

Note that $\dim(C_p(K_*)) = \dim(B_{p-1}(K_*)) + \dim(Z_p(K_*))$ by the rank-nullity theorem, so we may rewrite (??) as:

$$\beta_p^{i,j} = \dim(C_p(K_i)) - \dim(B_{p-1}(K_i)) - \dim(Z_p(K_i) \cap B_p(K_j)) \quad (10)$$

The dimension of the boundary group $B_{p-1}(K_i)$ may be directly inferred from the rank of ∂_p^i , and the dimension of $C_p(K_i)$ is simply the number of p -simplices with filtration values $f(\sigma) \leq i$. To express the intersection term, we require more notation. If A is a $m \times m$ matrix, then let A_i^j denote the lower-left submatrix of A given by the first j columns and last $m - i + 1$ rows (rows i through m , inclusive). For any $1 \leq i < j \leq m$, define the quantity $r_A(i, j)$:

$$r_A(i, j) = \text{rank}(A_i^j) - \text{rank}(A_{i+1}^j) + \text{rank}(A_{i+1}^{j-1}) - \text{rank}(A_i^{j-1}) \quad (11)$$

One of the seminal results from the Pairing Uniqueness lemma [] 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 (todo: elaborate more):

$$\text{low}_R[j] = i \iff r_{\partial}(i, j) = 1 \iff \text{rank}(R_i^j) = \text{rank}(\partial_i^j)$$

Thus, the lower-left submatrices of the filtered boundary matrix ∂ has the same rank as R . Let $\hat{\partial} = \partial_p$ and $\bar{\partial} = \partial_{p+1}$. Utilizing the result above, observe $\beta_p^{i,j}$ may be written as:

$$\beta_p^{i,j} = |\hat{\partial}_1^i| - \text{rank}(\hat{\partial}_1^i) - \text{rank}(\bar{\partial}_1^j) + \text{rank}(\bar{\partial}_{i+1}^j) \quad (12)$$

where $|A| = \dim(\text{dom}(A))$ denotes the order of the matrix. Thus, we may write the persistent Betti number as a combination of rank computations performed directly on the dimension p and $(p+1)$ boundary matrices.

Boundary Matrix Relaxation

As integer-valued invariants, Betti numbers pose several difficulties to direct optimization. Thus, we require alternative expressions for each of the terms in equation (??) to extend its applicability to the time-varying setting. Towards deriving these expression, we first require a replacement of the standard boundary matrix formulation.

Recall that the boundary operator ∂_p for a finite simplicial filtration K_\bullet with $m = |C_p(K_\bullet)|$ and $n = |C_{p-1}(K_\bullet)|$ can be represented by an $(n \times m)$ boundary matrix ∂_p whose columns and rows correspond to p -simplices and $(p-1)$ -simplices, respectively. The entries of ∂_p depend on the choice of \mathbb{F} ; in general, after orientating the simplices of K arbitrarily, they have the form:

$$\partial_p[i, j] = \begin{cases} c(\sigma) & \text{if } \sigma_i \in \partial_p(\sigma_j) \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

where $c(\sigma) \in \mathbb{F}$ is an arbitrary constant satisfying $c(\sigma) = -c(\sigma')$ if σ and σ' are opposite orientations of the same simplex. In practice, $c(\sigma)$ is typically set to ± 1 . Towards relaxing the persistent Betti computation in dynamic setting, we propose an alternative choice for $c(\sigma)$ which endows continuity in the entries of ∂_p in T .

Definition 1 (Time-varying boundary matrix). *Let $\mathbb{F} = \mathbb{R}$ denote the field, $\delta_X(\cdot) = (X, d_X(\cdot))$ a DMS over a finite set X of fixed size $|X| = n$, and let $(\mathcal{P}(X), \preceq^*)$ be a linear extension of the face poset of the $(n-1)$ -simplex Δ_n . For a fixed $\epsilon > 0$, a time-varying p -th boundary matrix ∂_p^t is an $\binom{n}{p} \times \binom{n}{p+1}$ matrix whose entries $c(\sigma)$ satisfy:*

$$\partial_p^t[i, j] = \begin{cases} |\epsilon - \text{diam}_t(\sigma)|_+ & \text{if } \sigma_i \in \partial_p(\sigma_j) \\ 0 & \text{otherwise} \end{cases}$$

and whose rows and columns are ordered by \preceq^* for all $t \in T$.

We now show a few properties that ∂_p^t exhibits which is advantageous for optimization. Clearly the entries of ∂_p^t must vary continuously in $t \in T$. Moreover, for all $t, t' \in T$, we have:

1. $\text{rank}(\partial_p^t) = \dim(\text{B}_{p-1}(K_t))$ where $K_t = \text{Rips}_\epsilon(\delta_X(t))$
2. $\|\partial_p^t - \partial_p^{t'}\|_F \leq C|t - t'|(p+1)$ when δ_X is C -Lipshitz over T
3. $\|\partial_p^t\|_2 \leq \sqrt{\kappa}(p+1)\epsilon$ where $\kappa = \max_{t \in T} \sum_{\sigma \in K_t} \mathbb{1}(\text{diam}(\sigma) \leq \epsilon)$

Proof. First, consider property (1). For any $t \in T$, applying the boundary operator ∂_p to $K_t = \text{Rips}_\epsilon(\delta_X(t))$ with non-zero entries satisfying (??) by definition yields a matrix ∂_p satisfying $\text{rank}(\partial_p) = \dim(\text{B}_{p-1}(K_t))$. In contrast, definition (??) always produces p -boundary matrices of Δ_n ; however, notice that the only entries which are non-zero are precisely those whose simplices σ that satisfy $\text{diam}(\sigma) < \epsilon$. Thus, $\text{rank}(\partial_p^t) = \dim(\text{B}_{p-1}(K_t))$ for all $t \in T$. < (show proof of (2))> Property (3) follows from the construction of ∂_p and from the inequality $\|A\|_2 \leq \sqrt{m}\|A\|_1$ for an $n \times m$ matrix A , as $\|\partial_p^t\|_1 \leq (p+1)\epsilon$ for all $t \in T$. □

Rank Relaxation (TODO)

In light of the formulation of (??), we may interpret $\beta_p^{i,j}$ from a function composition perspective:

$$t \xrightarrow{f} \partial_*^t \xrightarrow{g} \text{rank}(\partial_*^t)$$

In this sense, by modifying the entries of ∂_p^* via ??, we ensure that f is both continuous and inherits the smoothness of $\delta_X(\cdot)$. We now address g .

Smoothness

Given a function $f : \mathbb{R} \rightarrow (-\infty, \infty]$ and a fixed $\mu > 0$, the *proximal operator* or *prox* of f is given by:

$$\text{prox}_f^\mu(x) = \arg \min_{u \in \mathbb{R}} \left\{ f(u) + \frac{1}{2\mu} \|u - x\|^2 \right\} \quad (14)$$

When f is proper closed convex, $\text{prox}_f^\mu(x)$ is single-valued and yields the solution to the *Moreau envelope* of f :

$$M_f^\mu(x) = \min_{u \in \mathbb{R}} \left\{ f(u) + \frac{1}{2\mu} \|x - u\|^2 \right\} = f(\text{prox}_f^\mu(x)) + \frac{1}{2\mu} \|x - \text{prox}_f^\mu(x)\|^2$$

$M_f^\mu(x)$ exhibits a number of properties suitable for optimization: it is $\frac{1}{\mu}$ -Lipshitz over \mathbb{R} , it retains the same minima as f , and for any $x \in \mathbb{R}$ it admits a gradient $\nabla M_f^\mu(x)$ which again is expressible via the proximal operator:

$$\nabla M_f^\mu(x) = \frac{1}{\mu}(x - \text{prox}_f^\mu(x)) \quad (15)$$

Moreover, if ∇f is L_f -Lipshitz, then $\nabla M_f^\mu(x)$ is L_f/μ -Lipshitz, and if f is L_f -Lipshitz, then $|f(x) - M_f^\mu(x)| \leq L_f^2\mu$. Thus, the Moreau envelope acts as a smooth approximation of f , which makes it an excellent candidate for smoothing (??) if prox_f^μ can be efficiently computed. Fortunately, the prox of $\|\cdot\|_*$ admits a simple characterization:

$$\text{prox}_{\|\cdot\|_*}^\mu(X) = \arg \min_{Y \in \mathbb{R}^{n \times m}} \left\{ \|Y\|_* + \frac{1}{2\mu} \|Y - X\|_F^2 \right\} \quad (16)$$

$$= U\mathcal{D}_\mu(S)V^T \quad (17)$$

where $X = USV^T$ is the SVD of an $(n \times m)$ matrix X and $\mathcal{D}_\mu(S) = \text{diag}(\{|\sigma_i - \mu|_+\}_{1 \leq i \leq r})$ is the application of the *soft-thresholding operator* to the singular values $S = \text{diag}(\{\sigma_i\}_{1 \leq i \leq r})$. Equation (??) yields the prox operator for nuclear norm $\|A\|_*$, which is not necessarily a convex function. However, it is known that $\|\cdot\|_*$ is convex over the set $\{X \in \mathbb{R}^{n \times m} : \|X\|_2 \leq m\}$, thus we may inherit the smoothing properties of M_f^μ by considering the prox operator for the function $f : X \mapsto \frac{1}{m}\|X\|_*$. Letting $\alpha = m^{-1}$, the proximal operator of this function is given by:

$$\text{prox}_{\alpha f}^\mu(X) = \arg \min_{Y \in \mathbb{R}^{n \times m}} \left\{ \alpha \|Y\|_* + \frac{1}{2\mu} \|Y - X\|_F^2 \right\} = \alpha \cdot \text{prox}_f^{\mu\alpha}(X) \quad (18)$$

Remark 2. *Show example of Moreau envelope*

A. Appendix

Dynamic Metric Spaces

Consider an \mathbb{R} -parameterized metric space $\delta_X = (X, d_X(\cdot))$ where X is a finite set and $d_X(\cdot) : \mathbb{R} \times X \times X \rightarrow \mathbb{R}_+$, satisfying:

1. For every $t \in \mathbb{R}$, $\delta_X(t) = (X, d_X(t))$ is a pseudo-metric space¹
2. For fixed $x, x' \in X$, $d_X(\cdot)(x, x') : \mathbb{R} \rightarrow \mathbb{R}_+$ is continuous.

When the parameter $t \in \mathbb{R}$ is interpreted as *time*, the above yields a natural characterization of a “time-varying” metric space. More generally, we refer to an \mathbb{R}^h -parameterized metric space as *dynamic metric space* (DMS). Such space have been studied more in-depth [] and have been shown...

Rank relaxation

A common approach in the literature to optimize quantities involving $\text{rank}(A)$ for some $m \times n$ matrix A is to consider optimizing its *nuclear norm* $\|A\|_* = \text{tr}(\sqrt{A^T A}) = \sum_{i=1}^r |\sigma_i|$, where σ_i denotes the i th singular value of A and $r = \text{rank}(A)$. One of the primary motivations for this substitution is that the nuclear norm is a convex envelope of the rank function over the set:

$$S := \{A \in \mathbb{R}^{n \times m} \mid \|A\|_2 \leq m\}$$

That is, for an appropriate $m > 0$, the function $A \mapsto \frac{1}{m}\|A\|_*$ is a lower convex envelope of the rank function over S . The nuclear norm also admits a subdifferential... thus, we may consider replacing (??) with:

$$\beta_p^{i,j}(t) = |\partial_{p,t}^{1,i}| - m_1^{-1} \|\partial_{p,t}^{1,i}\|_* - m_2^{-1} \|\partial_{p,t}^{1,j}\|_* - m_3^{-1} \|\partial_{\bar{p},t}^{i,j}\|_* \quad (19)$$

where $\bar{c} = c + 1$. Now, if $t \mapsto \partial_p^*(t)$ is a non-decreasing, convex function in t , then the composition ... is convex, as each of the individual terms are convex. Moreover, we have...

< Insert proof about this relaxation always lower-bounding β >

¹This is required so that if one can distinguish the two distinct points $x, x' \in X$ incase $d_X(t)(x, x') = 0$ at some $t \in \mathbb{R}$.

Computation

In this section, we discuss the computation of suitable bases for the subspaces $Z_p(X_*)$, $B_p(K_*)$, and $Z_p(X_*) \cap B_p(X_*)$. In particular, we address two cases: the *dense* case, wherein the aforementioned bases are represented densely in memory, and the *sparse* case, which uses the structure of a particular decomposition of the boundary matrices to derive bases whose size in memory inherits the sparsity pattern of the decomposition.

Sparse case: We require an appropriate choice of bases for the groups $B_{p-1}(K_*)$ and $Z_p(X_*) \cap B_p(X_*)$. For some fixed $t \in T$, let $R_p = \partial_p V_p$ denote the decomposition discussed above, and let $b, d \in \mathbb{R}_+$ be fixed constants satisfying $b \leq d$. Since the boundary group $B_{p-1}(K_b)$ lies in the image of the ∂_p , it can be shown that a basis for the boundary group $B_{p-1}(K_*)$ is given by:

$$M_p^b = \{ \text{col}_{R_{p+1}}(j) \neq 0 \mid j \leq b \} \quad \text{span}() \quad (20)$$

Moreover, since $B_{p-1}(K_b) = \text{Im}(\partial_p^b)$, we have $\text{span}(M_p^b) = B_{p-1}(K_b)$ and thus $\text{rank}(M_p^b) = \text{rank}(\partial_p^b)$. Indeed, it can be shown that every lower-left submatrix of ∂_p^* satisfies $\text{rank}(\partial_p^*) = \text{rank}(R_p^*)$. Thus, although M_p^b does provide a minimal basis for the boundary group $B_{p-1}(K_b)$, it is unneeded here.

A suitable basis for the cycle group can also be read off from the reduced decomposition directly as well. Indeed, let $R_p = \partial_p V_p$ as before. Then the cycle group is spanned by linear combinations of columns of V_p :

$$Z_p^b = \{ \text{col}_{V_p}(j) \mid \text{col}_{R_p}(j) = 0, j \leq b \} \quad (21)$$

The formulation of a basis spanning $Z_p(K_i) \cap B_p(K_j)$ is more subtle, as we can no longer use the fact that every lower-left submatrix of R_p has the same rank as the same lower-left submatrix of ∂_p . Nonetheless, a basis for this group can be obtained by reading off specific columns from R_p :

$$M_p^{b,d} := \{ \text{col}_{R_{p+1}}(k) \neq 0 \mid 1 \leq k \leq d \text{ and } 1 \leq \text{low}_{R_{p+1}}(k) \leq b \} \quad (22)$$

One can show that M_p^d does indeed span $Z_p(X_*) \cap B_p(X_*)$ by using the fact that the non-zero columns of R_p with indices at most d form a basis for $B_p(K_d)$, and that each low-row index for every non-zero is unique.

Dense case: In general, persistent homology groups and its various factor groups are well-defined and computable with the reduction algorithm with coefficients chosen over any ring. By applying operations with respect to a field \mathbb{F} , both the various group structures $Z_p(K_\bullet) \subseteq B_p(K_\bullet) \subseteq C_p(K_\bullet)$ and their induced quotient groups $H_p(K_\bullet)$ are vector spaces; thus, the computation of suitable bases can be approached from a purely linear algebraic perspective. In particular, by fixing $\mathbb{F} = \mathbb{R}$, we inherit not only many useful tools for obtaining suitable bases for these groups, but also access to their corresponding optimized implementations as well.

Consider the p -th boundary operator $\partial_p^* : C_p(K_*) \rightarrow C_{p-1}(K_*)$ whose matrix realization with respect to some choice of simplex ordering $\{\sigma_i\}_{1 \leq i \leq m}$ we also denote with ∂_p . By definition, the boundary group $B_p(K_*)$ is given by the image $\text{Im}(\partial_{p+1}^*) = B_p(K_*)$, thus one may basis for $B_p(K_*)$ by computing the considering the first $r > 0$ columns of the reduced SVD:

$$M_p^* = [u_1 \mid u_2 \mid \cdots \mid u_r] = \{ \} \quad (23)$$