Improved Approximate Rips Filtrations with Shifted Integer Lattices*

Aruni Choudhary[1], Michael Kerber[†2], and Sharath Raghvendra[‡3]

1 **Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany**
aruni.choudhary@mpi-inf.mpg.de
2 **Graz University of Technology, Graz, Austria**
kerber@tugraz.at
3 **Virginia Tech, Blacksburg, VA, USA**
sharathr@vt.edu

Abstract

Rips complexes are important structures for analyzing topological features of metric spaces. Unfortunately, generating these complexes constitutes an expensive task because of a combinatorial explosion in the complex size. For n points in $\mathbb{R}^d$, we present a scheme to construct a $3\sqrt{2}$-approximation of the multi-scale filtration of the L_∞-Rips complex, which extends to a $O(d^{0.25})$-approximation of the Rips filtration for the Euclidean case. The k-skeleton of the resulting approximation has a total size of $n2^{O(d\log k)}$. The scheme is based on the integer lattice and on the barycentric subdivision of the d-cube.

1998 ACM Subject Classification F.2.2 Geometrical problems and computations

Keywords and phrases Persistent homology, Rips filtrations, Approximation algorithms, Topological Data Analysis

Digital Object Identifier 10.4230/LIPIcs.ESA.2017.28

1 Introduction

Persistent homology [4, 10, 11] is a technique to analyze of data sets using topological invariants. The idea is to build a multi-scale representation of the data set and to track its homological changes across the scales.

A standard construction for the important case of point clouds in Euclidean space is the *Vietoris-Rips complex* (or just *Rips complex*): for a scale parameter $\alpha \geq 0$, it is the collection of all subsets of points with diameter at most α. When α increases from 0 to ∞, the Rips complexes form a *filtration*, an increasing sequence of nested simplicial complexes whose homological changes can be computed and represented in terms of a *barcode*.

The computational drawback of Rips complexes is their sheer size: the k-skeleton of a Rips complex (that is, only subsets of size $\leq k+1$ are considered) for n points consists of $\Theta(n^{k+1})$ simplices because every $(k+1)$-subset joins the complex for a sufficiently large scale parameter. This size bound turns barcode computations for large point clouds infeasible even for low-dimensional homological features[1]. This poses the question of what we can say about the barcode of the Rips filtration without explicitly constructing all of its simplices.

* A longer version of the paper is available on arXiv [7].
† MK is supported by the Austrian Science Fund (FWF) grant number P 29984-N35.
‡ SR acknowledges support of NSF CRII grant CCF-1464276.
1 An exception are point clouds in $\mathbb{R}^2$ and $\mathbb{R}^3$, for which *alpha complexes* [10] are an efficient alternative.

25th Annual European Symposium on Algorithms (ESA 2017).
Editors: Kirk Pruhs and Christian Sohler; Article No. 28; pp. 28:1–28:13
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

We address this question using approximation techniques. Barcodes form a metric space: two barcodes are close if the same homological features occur on roughly the same range of scales (see Section 2 for the precise definition). The first approximation scheme by Sheehy [16] constructs a $(1+\varepsilon)$-approximation of the k-skeleton of the Rips filtration using only $n(\frac{1}{\varepsilon})^{O(\lambda k)}$ simplices for arbitrary finite metric spaces, where λ is the doubling dimension of the metric. Further approximation techniques for Rips complexes [9] and the closely related *Čech complexes* [1, 5, 13] have been derived subsequently, all with comparable size bounds. More recently, we constructed an approximation scheme for Rips complexes in Euclidean space that yields a worse approximation factor of $O(d)$, but uses only $n2^{O(d \log k)}$ simplices [8], where d is the ambient dimension of the point set.

Contributions. We present a $3\sqrt{2}$-approximation for the Rips filtration of n points in $\mathbb{R}^d$ in the L_∞-norm , whose k-skeleton has size $n2^{O(d \log k)}$. This translates to a $O(d^{0.25})$-approximation of the Rips filtration in the Euclidean metric and hence improves the asymptotic approximation quality of our previous approach [8] with the same size bound.

On a high level, our approach follows a straightforward approximation scheme: given a scaled and appropriately shifted integer grid on $\mathbb{R}^d$, we identify those grid points that are close to the input points and build an approximation complex using these grid points. The challenge lies in how to connect these grid points to a simplicial complex such that close-by grid points are connected, while avoiding too many connections to keep the size small. Our approach first selects a set of *active faces* in the cubical complex defined over the grid, and defines the approximation complex using the barycentric subdivision of this cubical complex.

We also describe an output-sensitive algorithm to compute our approximation. By randomizing the aforementioned shifts of the grids, we obtain a worst-case running time of $n2^{O(d)} \log \Delta + 2^{O(d)} M$, where Δ is *spread* of the point set (that is, the ratio of the diameter to the closest distance of two points) and M is the size of the approximation.

Additionally, this paper makes the following technical contributions:

- We follow the standard approach of defining a sequence of approximation complexes and establishing an *interleaving* between the Rips filtration and the approximation. We realize our interleaving using *chain maps* connecting a Rips complex at scale α to an approximation complex at scale $c\alpha$, and vice versa, with $c \geq 1$ being the approximation factor. Previous approaches [8, 9, 16] used *simplicial maps* for the interleaving, which induce an elementary form of chain maps and are therefore more restrictive.
  The explicit construction of such maps can be a non-trivial task. The novelty of our approach is that we avoid this construction by the usage of *acyclic carriers* [15]. In short, carriers are maps that assign subcomplexes to subcomplexes under some mild extra conditions. While they are more flexible, they still certify the existence of suitable chain maps, as we exemplify in Section 4. We believe that this technique is of general interest for the construction of approximations of cell complexes.
- We exploit a simple trick that we call *scale balancing* to improve the quality of approximation schemes. In short, if the aforementioned interleaving maps from and to the Rips filtration do not increase the scale parameter by the same amount, one can simply multiply the scale parameter of the approximation by a constant. Concretely, given maps

  $$\phi_\alpha : \mathcal{R}_\alpha \to \mathcal{X}_\alpha \qquad \psi_\alpha : \mathcal{X}_\alpha \to \mathcal{R}_{c\alpha}$$

  interleaving the Rips complex $\mathcal{R}_\alpha$ and the approximation complex $\mathcal{X}_\alpha$, we can define $\mathcal{X}'_\alpha := \mathcal{X}_{\alpha/\sqrt{c}}$ and obtain maps

  $$\phi'_\alpha : \mathcal{R}_\alpha \to \mathcal{X}'_{\sqrt{c}\alpha} \qquad \psi_\alpha : \mathcal{X}'_\alpha \to \mathcal{R}_{\sqrt{c}\alpha}$$

which improves the interleaving from c to $\sqrt{c}$. While it has been observed that the same trick can be used for improving the worst-case distance between Rips and Čech filtrations[2], our work seems to be the first to make use of it in the context of approximations.

Our technique can be combined with dimension reduction techniques in the same way as in [8] (see Theorems 19, 21, and 22 therein), with improved logarithmic factors. We omit the technical details in this paper. Also, we point out that the complexity bounds for size and computation time are for the entire approximation scheme and not for a single scale as in [8]. However, similar techniques as the ones exposed in Section 5 can be used to improve the results of [8] to hold for the entire approximation as well[3].

Outline. We start the presentation by discussing the relevant topological concepts in Section 2. Then, we present few results about grid lattices in Section 3. Building on these ideas, the approximation scheme is presented in Section 4. Computational aspects of the approximation scheme are discussed in Section 5. We conclude in Section 6. Many of the proofs are detailed in the arXiv version of our paper [7].

2 Background

We review the essential topological concepts needed; see [2, 6, 10, 15] for more details.

Simplicial complexes. A *simplicial complex* K on a finite set of elements S is a collection of subsets $\{\sigma \subseteq S\}$ called *simplices* such that each subset $\tau \subset \sigma$ is also in K. The dimension of a simplex $\sigma \in K$ is $k := |\sigma| - 1$, in which case σ is called a *k-simplex*. A simplex τ is a *subsimplex* of σ if $\tau \subseteq \sigma$. We remark that, commonly a subsimplex is called a 'face' of a simplex, but we reserve the word 'face' for a different structure. For the same reason, we do not introduce the common notation of of 'vertices' and 'edges' of simplicial complexes, but rather refer to 0- and 1-simplices throughout. The *k-skeleton* of K consists of all simplices of K whose dimension is at most k. For instance, the 1-skeleton of K is a graph defined by its 0-simplices and 1-simplices.

Given a point set $P \subset \mathbb{R}^d$ and a real number $\alpha \geq 0$, the *(Vietoris-)Rips* complex on P at scale α consists of all simplices $\sigma = (p_0, \ldots, p_k) \subseteq P$ such that $diam(\sigma) \leq \alpha$, where *diam* denotes the diameter. In this work, we write $\mathcal{R}_\alpha$ for the Rips complex at scale α with the Euclidean metric, and $\mathcal{R}_\alpha^\infty$ when using the metric of the L_∞-norm. In either way, a Rips complex is an example of a *flag complex*, which means that whenever a set $\{p_0, \ldots, p_k\} \subseteq P$ has the property that every 1-simplex $\{p_i, p_j\}$ is in the complex, then the k-simplex $\{p_0, \ldots, p_k\}$ is also in the complex.

A simplicial complex K' is a *subcomplex* of K if $K' \subseteq K$. For instance, $\mathcal{R}_\alpha$ is a subcomplex of $\mathcal{R}_{\alpha'}$ for $0 \leq \alpha \leq \alpha'$. Let L be a simplicial complex. Let $\hat{\varphi}$ be a map which assigns to each vertex of K, a vertex of L. A map $\varphi : K \to L$ is called a *simplicial map* induced by $\hat{\varphi}$, if for every simplex $\{p_0, \ldots, p_k\}$ in K, the set $\{\hat{\varphi}(p_0), \ldots, \hat{\varphi}(p_k)\}$ is a simplex of L. For K' a subcomplex of K, the inclusion map $inc : K' \to K$ is an example of a simplicial map. A simplicial map $K \to L$ is completely determined by its action on the 0-simplices of K.

[2] Ulrich Bauer, private communication
[3] An extended version of [8] containing these improvements is currently under submission.

Chain complexes. A *chain complex* $\mathcal{C}_* = (\mathcal{C}_p, \partial_p)$ with $p \in \mathbb{N}$ is a collection of abelian groups $\mathcal{C}_p$ and homomorphisms $\partial_p : \mathcal{C}_p \to \mathcal{C}_{p-1}$ such that $\partial_{p-1} \circ \partial_p = 0$. A simplicial complex K gives rise to a chain complex $\mathcal{C}_*(K)$ by fixing a base field $\mathcal{F}$, defining $\mathcal{C}_p$ as the set of formal linear combinations of p-simplices in K over $\mathcal{F}$, and ∂_p as the linear operator that assigns to each simplex the (oriented) sum of its sub-simplices of codimension one[4].

A *chain map* $\phi : \mathcal{C}_* \to D_*$ between chain complexes $\mathcal{C}_* = (\mathcal{C}_p, \partial_p)$ and $D_* = (D_p, \partial'_p)$ is a collection of group homomorphisms $\phi_p : \mathcal{C}_p \to D_p$ such that $\phi_{p-1} \circ \partial_p = \partial'_p \circ \phi_p$. For example, a simplicial map φ between simplicial complexes induces a chain map $\bar{\varphi}$ between the corresponding chain complexes. This construction is *functorial*, meaning that for φ the identity function on a simplicial complex K, $\bar{\varphi}$ is the identity function on $\mathcal{C}_*(K)$, and for composable simplicial maps φ, φ', we have that $\overline{\varphi \circ \varphi'} = \bar{\varphi} \circ \bar{\varphi}'$.

Homology and carriers. The *p-th homology group* $H_p(\mathcal{C}_*)$ of a chain complex is defined as $\ker \partial_p / \operatorname{im} \partial_{p+1}$. The p-th homology group of a simplicial complex K, $H_p(K)$, is the p-th homology group of its induced chain complex. In either case $H_p(\mathcal{C}_*)$ is a $\mathcal{F}$-vector space because we have chosen our base ring $\mathcal{F}$ as a field. Intuitively, when the chain complex is generated from a simplicial complex, the dimension of the p-th homology group counts the number of p-dimensional holes in the complex (except for $p = 0$, where it counts the number of connected components). We write $H(\mathcal{C}_*)$ for the direct sum of all $H_p(\mathcal{C}_*)$ for $p \geq 0$.

A chain map $\phi : \mathcal{C}_* \to D_*$ induces a linear map $\phi^* : H(\mathcal{C}_*) \to H(D_*)$ between the homology groups. Again, this construction is functorial, meaning that it maps identity maps to identity maps, and it is compatible with compositions.

We call a simplicial complex K *acyclic*, if K is connected and all homology groups $H_p(K)$ with $p \geq 1$ are trivial. For simplicial complexes K and L, an *acyclic carrier* Φ is a map that assigns to each simplex σ in K, a non-empty subcomplex $\Phi(\sigma) \subseteq L$ such that $\Phi(\sigma)$ is acyclic, and whenever τ is a subsimplex of σ, then $\Phi(\tau) \subseteq \Phi(\sigma)$. We say that a chain $c \in \mathcal{C}_p(K)$ is *carried* by a subcomplex K', if c takes value 0 except for p-simplices in K'. A chain map $\phi : \mathcal{C}_*(K) \to \mathcal{C}_*(L)$ is *carried by* Φ, if for each simplex $\sigma \in K$, $\phi(\sigma)$ is carried by $\Phi(\sigma)$. We state the *acyclic carrier theorem* [15]:

▶ **Theorem 1.** *Let $\Phi : K \to L$ be an acyclic carrier.*

- *There exists a chain map $\phi : \mathcal{C}_*(K) \to \mathcal{C}_*(L)$ such that ϕ is carried by Φ.*
- *If two chain maps $\phi_1, \phi_2 : \mathcal{C}_*(K) \to \mathcal{C}_*(L)$ are both carried by Φ, then $\phi_1^* = \phi_2^*$.*

Filtrations and towers. Let $I \subseteq \mathbb{R}$ be a set of real values which we refer to as *scales*. A *filtration* is a collection of simplicial complexes $(K_\alpha)_{\alpha \in I}$ such that $K_\alpha \subseteq K'_\alpha$ for all $\alpha \leq \alpha' \in I$. For instance, $(\mathcal{R}_\alpha)_{\alpha \geq 0}$ is a filtration which we call the *Rips filtration*. A *(simplicial) tower* is a sequence $(K_\alpha)_{\alpha \in J}$ of simplicial complexes with J being a discrete set (for instance $J = \{2^k \mid k \in \mathbb{Z}\}$), together with simplicial maps $\varphi_\alpha : K_\alpha \to K_{\alpha'}$ between complexes at consecutive scales. For instance, the Rips filtration can be turned into a tower by restricting to a discrete range of scales, and using the inclusion maps as φ. The approximation constructed in this paper will be another example of a tower.

We say that a simplex σ is *included* in the tower at scale α', if σ is not the image of $\varphi_\alpha : K_\alpha \to K_{\alpha'}$, where α is the scale preceding α' in the tower. The *size* of a tower is the number of simplices included over all scales. If a tower arises from a filtration, its size is simply the size of the largest complex in the filtration (or infinite, if no such complex exists).

---

[4] To avoid thinking about orientations, it is often assumed that $\mathcal{F} = \mathbb{Z}_2$ is the field with two elements.

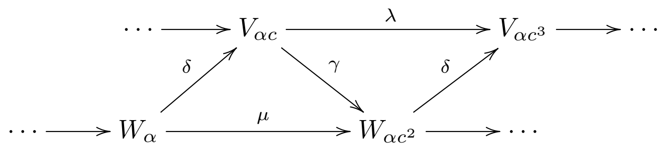
However, this is not true for in general for simplicial towers, since simplices can collapse in the tower and the size of the complex at a given scale may not take into account the collapsed simplices which were included at earlier scales in the tower.

Barcodes and Interleavings. A collection of vector spaces $(V_\alpha)_{\alpha\in I}$ connected with linear maps $\lambda_{\alpha_1,\alpha_2} : V_{\alpha_1} \to V_{\alpha_2}$ is called a *persistence module*, if $\lambda_{\alpha,\alpha}$ is the identity on V_α and $\lambda_{\alpha_2,\alpha_3} \circ \lambda_{\alpha_1,\alpha_2} = \lambda_{\alpha_1,\alpha_3}$ for all $\alpha_1 \leq \alpha_2 \leq \alpha_3 \in I$ for the index set I.

We generate persistence modules using the previous concepts. Given a simplicial tower $(K_\alpha)_{\alpha\in I}$, we generate a sequence of chain complexes $(\mathcal{C}_*(K_\alpha))_{\alpha\in I}$. By functoriality, the simplicial maps φ of the tower give rise to chain maps $\overline{\varphi}$ between these chain complexes. Using functoriality of homology, we obtain a sequence $(H(K_\alpha))_{\alpha\in I}$ of vector spaces with linear maps $\overline{\varphi}^*$, forming a persistence module. The same construction can be applied to filtrations.

Persistence modules admit a decomposition into a collection of intervals of the form $[\alpha, \beta]$ (with $\alpha, \beta \in I$), called the *barcode*, subject to certain tameness conditions. The barcode of a persistence module characterizes the module uniquely up to isomorphism. If the persistence module is generated by a simplicial complex, an interval $[\alpha, \beta]$ in the barcode corresponds to a homological feature (a "hole") that comes into existence at complex K_α and persists until it disappears at K_β.

Two persistence modules $(V_\alpha)_{\alpha\in I}$ and $(W_\alpha)_{\alpha\in I}$ with linear maps $\lambda_{\cdot,\cdot}$ and $\mu_{\cdot,\cdot}$ are said to be *weakly (multiplicatively) c-interleaved* with $c \geq 1$, if there exist linear maps $\gamma_\alpha : V_\alpha \to W_{c\alpha}$ and $\delta_\alpha : W_\alpha \to V_{c\alpha}$, called *interleaving maps*, such that the diagram



(1)

commutes for all $\alpha \in I$, that is, $\mu = \gamma \circ \delta$ and $\lambda = \delta \circ \gamma$ (we have skipped the subscripts of the maps for readability). In such a case, the barcodes of the two modules are $3c$-approximations of each other in the sense of [6]. We say that two towers are *c-approximations* of each other, if their persistence modules that are c-approximations. Under the more stringent conditions of *strong interleaving*, the approximation ratio can be improved. See [7] for more details.

3 Grids and cubes

Let $I := \{\lambda 2^s \mid s \in \mathbb{Z}\}$ with $\lambda > 0$ be a discrete set of scales. For a scale $\alpha_s := \lambda 2^s$, we inductively define a grid G_s on scale α_s which is a scaled and translated (shifted) version of the integer lattice: for $s = 0$, G_s is simply $\lambda\mathbb{Z}^d$, the scaled integer grid. For $s \geq 0$, we choose an arbitrary $O \in G_s$ and define

$$G_{s+1} = 2(G_s - O) + O + \frac{\alpha_s}{2}(\pm 1, \ldots, \pm 1) \tag{2}$$

where the signs of the components of the last vector are chosen uniformly at random (and the choice is independent for each s). For $s \leq 0$, we define

$$G_{s-1} = \frac{1}{2}(G_s - O) + O + \frac{\alpha_{s-1}}{2}(\pm 1, \ldots, \pm 1). \tag{3}$$

It is then easy to check that 2 and 3 are consistent at $s = 0$. A simple instance of the above construction is the sequence of lattices with $G_s := \alpha_s\mathbb{Z}^d$ for even s, and $G_s := \alpha_s\mathbb{Z}^d + \frac{\alpha_{s-1}}{2}(1, \ldots, 1)$ for odd s.

We motivate the shifting next. For a finite point set $Q \subset \mathbb{R}^d$ and $x \in Q$, the *Voronoi region* $Vor_Q(x) \subset \mathbb{R}^d$ is the (closed) set of points in $\mathbb{R}^d$ that have x as one of its closest points in Q. If $Q = G_s$, it is easy to see that the Voronoi region of any grid point x is a cube of side length α_s centered at x. The shifting of the grids ensures that each $x \in G_s$ lies in the Voronoi region of a unique $y \in G_{s+1}$. By an elementary calculation, we show a stronger statement; for shorter notation, we write $Vor_s(x)$ instead of $Vor_{G_s}(x)$.

▶ **Lemma 2.** *Let $x \in G_s, y \in G_{s+1}$ such that $x \in Vor_{s+1}(y)$. Then, $Vor_s(x) \subset Vor_{s+1}(y)$.*

Cubical complexes. The integer grid $\mathbb{Z}^d$ naturally defines a *cubical complex*, where each element is an axis-aligned, k-cube with $0 \leq k \leq d$. Let $\square$ denote the set of all integer translates of faces of the unit cube $[0,1]^d$, considered as a convex polytope in $\mathbb{R}^d$. We call the elements of $\square$ *faces*. Each face has a dimension k; the 0-faces, or *vertices* are exactly the points in $\mathbb{Z}^d$. The *facets* of a k-face f are the $(k-1)$-faces contained in f. We call a pair of facets of f *opposite* if they are disjoint. Obviously, these concepts carry over to scaled and translated versions of $\mathbb{Z}^d$, so we define $\square_s$ as the cubical complex defined by G_s.

We define a map $g_s : \square_s \to \square_{s+1}$ as follows: for vertices, we assign to $x \in G_s$ the (unique) vertex $y \in G_{s+1}$ such that $x \in Vor_{s+1}(y)$ (cf. Lemma 2). For a k-face f of $\square_s$ with vertices $(p_1, \ldots, p_{2^k})$ in G_s, we set $g_s(f)$ to be the convex hull of $\{g_s(p_1), \ldots, g_s(p_{2^k})\}$; the next lemma shows that this is indeed a well-defined map (see [7]).

▶ **Lemma 3.** *$\{g_s(p_1), \ldots, g_s(p_{2^k})\}$ are the vertices of a face e of G_{s+1}. Moreover, if e_1, e_2 are any two opposite facets of e, then there exists a pair of opposite facets f_1, f_2 of f such that $g_s(f_1) = e_1$ and $g_s(f_2) = e_2$.*

Barycentric subdivision. A *flag* in $\square_s$ is a set of faces $\{f_0, \ldots, f_k\}$ of $\square_S$ such that $f_0 \subseteq \ldots \subseteq f_k$. The *barycentric subdivision* sd_s of $\square_s$ is the (infinite) simplicial complex whose simplices are the flags of $\square_s$; in particular, the 0-simplices of sd_s are the faces of $\square_s$. An equivalent geometric description of sd_s can be obtained by defining the 0-simplices as the barycenters of the faces in sd_s, and introducing a k-simplex between $(k+1)$ barycenters if the corresponding faces form a flag. It is easy to see that sd_s is a flag complex. Given a face f in $\square_s$, we write $sd(f)$ for the subcomplex of sd_s consisting of all flags that are formed only by faces contained in f.

4 Approximation scheme

We define our approximation complex at scale α_s as a finite subcomplex of sd_s. To simplify the subsequent analysis, we define the approximation in a slightly generalized form.

Barycentric spans. For a fixed s, let V denote a non-empty subset of G_s. We say that a face $f \in \square_s$ is *spanned* by V if $f \cap V \neq \varnothing$ and is not contained in any facet of f. Trivially, the vertices of $\square_s$ spanned by V are precisely the points in V. We point out that the set of spanned faces is *not* closed under taking sub-faces; for instance, if V consists of two antipodal points of a d-cube, the only faces spanned by V are the d-cube and the two vertices.

The *barycentric span* of V is the subcomplex of sd_s defined by all flags $\{f_0, \ldots, f_k\}$ such that all f_i are spanned by V. This is indeed a subcomplex of sd_s because it is closed under taking subsets. Moreover, for a face $f \in \square_k$, we define the *f-local barycentric span* of V as the set of all flags $\{f_0, \ldots, f_k\}$ in the barycentric span such that $f_i \subseteq f$ for all i. This is a subcomplex both of $sd(f)$ and of the barycentric span of V and is a flag complex.

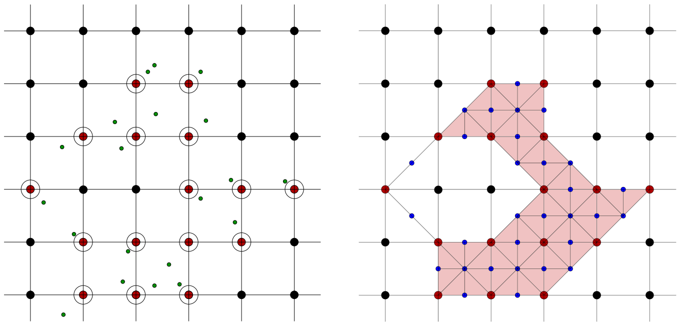


Figure 1 The left figure shows a two-dimensional grid, along with its cubical complex. The green points (small dots) denote the points in P and the red vertices (encircled) are the active vertices. The figure on the right shows the generated simplicial complex. The blue vertices (small dots) are the barycenters of the active faces.

▶ **Lemma 4.** *For each face f, the f-local barycentric span of V is either empty or acyclic.*

Furthermore, if $W \subseteq V$, it is easy to see that faces spanned by W are also spanned by V. Consequently, the barycentric span of W is a subcomplex of the barycentric span of V.

Approximation complex. We denote by $P \subset \mathbb{R}^d$ a finite set of points. For each point $p \in P$, we let $a_s(p)$ denote the grid point in G_s that is closest to p (we assume for simplicity that this closest point is unique). We define the *active vertices of* G_s, V_s, as $a_s(P)$, that is, the set of grid points that are closest to some point in P. The next statement is a direct application of the triangle inequality; let diam_∞ denote the diameter in the L_∞-norm.

▶ **Lemma 5.** *Let $Q \subseteq P$ be such that $\text{diam}_\infty(Q) \leq \alpha_s$. Then, the set $a_s(Q)$ is contained in a face of $\square_s$. Equivalently, for a simplex $\sigma = (p_0, \ldots, p_k) \in \mathcal{R}^\infty_{\alpha_s}$ on P, the set of active vertices $\{a_s(p_0), \ldots, a_s(p_k)\}$ is contained in a face of $\square_s$.*

Vice versa, we define a map $b_s : V_s \to P$ by mapping an active vertex to its closest point in P (again, assuming for simplicity that the assignment is unique). The map b_s is a section of a_s, that is, $a_s \circ b_s$ is the identity on V_s.

Recall that the map $g_s : \square_s \to \square_{s+1}$ from Section 3 maps grid points of G_s to grid points of G_{s+1}. With Lemma 2, it follows at once:

▶ **Lemma 6.** *For all $x \in V_s$, $g_s(x) = (a_{s+1} \circ b_s)(x)$.*

We now define our approximation tower: for scale α_s, we define $\mathcal{X}_{\alpha_s}$ as the barycentric span of the active vertices $V_s \subset G_s$. See Figure 1 for an illustration. To simplify notations, we call the faces of $\square_s$ spanned by V_s *active faces*, and simplices of $\mathcal{X}_{\alpha_s}$ *active flags*.

To complete the construction, we need to define simplicial maps $\mathcal{X}_{\alpha_s} \to \mathcal{X}_{\alpha_{s+1}}$. First, we show:

▶ **Lemma 7.** *Let f be an active face of $\square_s$. Then, $g_s(f)$ is an active face of $\square_{s+1}$.*

Proof. From Lemma 3, $e := g_s(f)$ is a face of G_{s+1}. If e is a vertex, it is active, because f contains at least one active vertex v, and $g_s(v) = e$ in this case. If e is not a vertex, we assume for a contradiction that it is not active. Then, it contains a facet e_1 that contains all active vertices in e. Let e_2 denote the opposite facet. By Lemma 3, f contains opposite facets f_1, f_2 such that $g_s(f_1) = e_1$ and $g_s(f_2) = e_2$. Since f is active, both f_1 and f_2 contain active vertices, in particular, f_2 contains an active vertex v. But then, the active vertex $g_s(v)$ must lie in e_2, contracting the fact that e_1 contains all active vertices of e. ◀

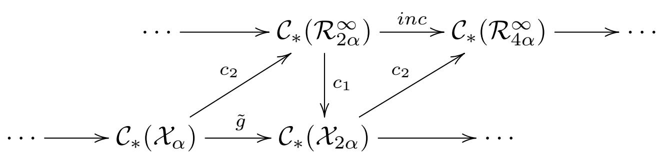
Recall that a simplex $\sigma \in \mathcal{X}_{\alpha_s}$ is a flag $f_0 \subseteq \ldots \subseteq f_k$ of active faces in $\square_s$. We set $\tilde{g}(\sigma)$ as the flag $g(f_0) \subseteq \ldots \subseteq g(f_k)$, which consists of active faces in $\square_{s+1}$ by Lemma 7, and hence is a simplex in $\mathcal{X}_{\alpha_{s+1}}$. It follows that $\tilde{g} : \mathcal{X}_{\alpha_s} \to \mathcal{X}_{\alpha_{s+1}}$ is a simplicial map. This finishes our construction of the simplicial tower $(\mathcal{X}_{\lambda 2^s})_{s\in\mathbb{Z}}$, with simplicial maps $\tilde{g} : \mathcal{X}_{\lambda 2^s} \to \mathcal{X}_{\lambda 2^{s+1}}$.

4.1 Interleaving

To relate our tower with the L_∞-Rips filtration, we start by defining two acyclic carriers. We write $\alpha := \alpha_s = \lambda 2^s$ to simplify notations.

- $C_1 : \mathcal{R}^\infty_\alpha \to \mathcal{X}_\alpha$: let $\sigma = (p_0, \ldots, p_k)$ be any simplex of $\mathcal{R}^\infty_\alpha$. We set $C_1(\sigma)$ as the barycentric span of $U := \{a_s(p_0), \ldots, a_s(p_k)\}$, which is a subcomplex of $\mathcal{X}_\alpha$. U lies in a face f of $\square_s$ by Lemma 5 hence $C_1(\sigma)$ is also the f-local barycentric span of U. Using Lemma 4, $C_1(\sigma)$ is acyclic.
- $C_2 : \mathcal{X}_\alpha \to \mathcal{R}^\infty_{2\alpha}$: let σ be any flag $e_0 \subseteq \ldots \subseteq e_k$ of $\mathcal{X}_\alpha$. Let $\{q_0, \ldots, q_m\}$ be the set of active vertices of e_k. We set $C_2(\sigma) := \{b_s(q_0), \ldots, b_s(q_m)\}$. With a simple triangle inequality, we see that $C_2(\sigma)$ is a simplex in $\mathcal{R}^\infty_{2\alpha}$, hence it is acyclic.

Using the Acyclic Carrier Theorem (Theorem 1), there exist chain maps $c_1 : \mathcal{C}_*(\mathcal{R}^\infty_\alpha) \to \mathcal{C}_*(\mathcal{X}_\alpha)$ and $c_2 : \mathcal{C}_*(\mathcal{X}_\alpha) \to \mathcal{C}_*(\mathcal{R}^\infty_{2\alpha})$, which are carried by C_1 and C_2, respectively. Aggregating the chain maps, we have the following diagram:



(4)

where *inc* corresponds to the inclusion chain map and $\tilde{g}$ denotes the chain map for the corresponding simplicial maps (we removed indices for readability). The chain complexes give rise to a diagram of the corresponding homology groups, connected by the induced linear maps $c_1^*, c_2^*, inc^*, \tilde{g}^*$.

▶ **Lemma 8.** *$inc^* = c_2^* \circ c_1^*$ and $\tilde{g}^* = c_1^* \circ c_2^*$. In particular, the persistence modules $(H(\mathcal{X}_{2^s}))_{s\in\mathbb{Z}}$ and $(H(\mathcal{R}^\infty_\alpha))_{\alpha\geq 0}$ are weakly 2-interleaved.*

Proof. To prove the claim, we consider both triangles separately. We show that the chain maps $\tilde{g}$ and $c_1 \circ c_2$ are carried by a common acyclic carrier. Then we show the same statement for *inc* and $c_2 \circ c_1$. The claim then follows from the Acyclic Carrier Theorem.

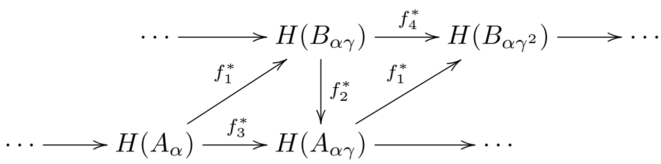
- *Lower triangle*: The map $C_1 \circ C_2 : \mathcal{X}_\alpha \to \mathcal{X}_{2\alpha}$ is an acyclic carrier, because $C_2(\sigma)$ is a simplex for any simplex $\sigma \in \mathcal{X}_\alpha$. Clearly, $C_1 \circ C_2$ carries the map $c_1 \circ c_2$. We show that it also carries $\tilde{g}$.
  Let σ be a flag $f_0 \subseteq \ldots \subseteq f_k$ in $\mathcal{X}_\alpha$ and let $V(f_i)$ denote the active vertices of f_i. Then, $C_1 \circ C_2(\sigma)$ is the barycentric span of $U := \{a_{s+1} \circ b_s(q) \mid q \in V(f_k)\} = \{g_s(q) \mid q \in V(f_k)\}$ (Lemma 6). On the other hand, $V(f_i) \subseteq V(f_k)$ and hence $g(V(f_i)) \subseteq U$. Then, $g(f_i)$ is spanned by U: indeed, since f_i is active, $g(f_i)$ is active and hence spanned by all active vertices, and it remains spanned if we remove all active vertices not in U, since they are not contained in f_i. It follows that the flag $g(f_0) \subseteq \ldots \subseteq g(f_k)$, which is equal to $\tilde{g}(\sigma)$, is in the barycentric span of U.
- *Upper triangle*: We define an acyclic carrier $D : \mathcal{R}^\infty_{2\alpha} \to \mathcal{R}^\infty_{4\alpha}$ which carries both *inc* and $c_2 \circ c_1$. Let $\sigma = (p_0, \ldots, p_k) \in \mathcal{R}^\infty_{2\alpha}$ be a simplex. The active vertices $U := \{a(p_0), \ldots, a(p_k)\} \subset G_{s+1}$ lie in a face f of $G_{2\alpha}$, using Lemma 5. We can assume that f is active, as otherwise, we pass to a facet of f that contains U. We set $D(\sigma)$ as the

simplex on the subset of points in P whose closest grid point in G_{s+1} lies in U. Using a simple application of triangle inequalities, $D(\sigma) \in \mathcal{R}^{\infty}_{4\alpha}$, so D is an acyclic carrier. The 0-simplices of σ are a subset of $D(\sigma)$, so D carries the map *inc*. We next show that D carries $c_2 \circ c_1$.

Let δ be a simplex in $\mathcal{X}_{2\alpha}$ for which the chain $c_1(\sigma)$ takes a non-zero value. Since $c_1(\sigma)$ is carried by $C_1(\sigma)$, $\delta \in C_1(\sigma)$ which is the barycentric span of $V(f)$. Furthermore, for any $\tau \in C_1(\sigma)$, $C_2(\tau)$ is of the form $\{b(q_0), \ldots, b(q_m)\}$ with $\{q_0, \ldots, q_m\} \in V(f)$. It follows that $C_2(\tau) \subseteq D(\sigma)$. In particular, since c_2 is carried by C_2, $c_2(c_1(\sigma)) \subseteq D(\sigma)$ as well. ◀

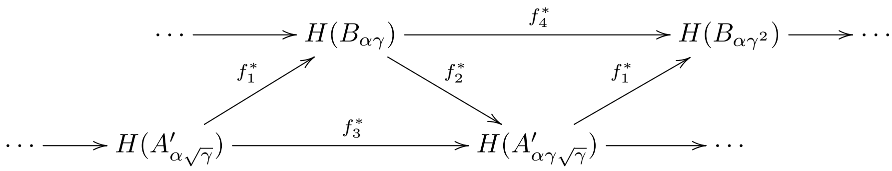
4.2 Scale balancing

We improve the approximation factor with a simple modification. Let $(A_{\lambda\gamma^k})_{k\in\mathbb{Z}}$ and $(B_{\lambda\gamma^k})_{k\in\mathbb{Z}}$ be two simplicial towers with simplicial maps f_3 and f_4 respectively, with $\lambda, \gamma > 0$. Assume that there exist interleaving linear maps f_1^*, f_2^* such that the diagram



(5)

commutes for all scales $\alpha = \lambda\gamma^k$, which implies that the persistence modules are weakly γ-interleaved. Defining another tower $(A'_{\lambda\sqrt{\gamma}\gamma^k})_{k\in\mathbb{Z}}$ with $A'_\alpha := A_{\alpha/\sqrt{\gamma}}$, we obtain a diagram



(6)

which implies that the persistence modules are weakly $\sqrt{\gamma}$-interleaved. Therefore, scale balancing improves the interleaving ratio by only scaling the persistence module.

In our context, we improve the weak 2-interleaving of $(H(\mathcal{X}_{2^k\alpha}))_{k\in\mathbb{Z}}$ and $(H(\mathcal{R}^{\infty}_\alpha))_{\alpha\geq 0}$ to a weak $\sqrt{2}$-interleaving. Using the proximity results for persistence modules [6],

▶ **Theorem 9.** *The persistence module* $\big(H(\mathcal{X}_{2^k/\sqrt{2}})\big)_{k\in\mathbb{Z}}$ *is a* $3\sqrt{2}$*- approximation of the* L_∞*-Rips persistence module* $\big(H(\mathcal{R}^{\infty}_\alpha)\big)_{\alpha\geq 0}$.

For any pair of points $p, p' \in \mathbb{R}^d$, it holds that $\|p - p'\|_2 \leq \|p - p'\|_\infty \leq \sqrt{d}\,\|p - p'\|_2$ which implies that the L_2- and the L_∞-Rips complexes are strongly $\sqrt{d}$-interleaved. The scale balancing technique also works for strongly interleaved persistence modules and yields

▶ **Lemma 10.** $(H(\mathcal{R}_{\alpha/d^{0.25}}))_{\alpha\geq 0}$ *is strongly* $d^{0.25}$*-interleaved with* $(H(\mathcal{R}^{\infty}_\alpha))_{\alpha\geq 0}$.

Using Theorem 9, Lemma 10 and the fact that interleavings satisfy the triangle inequality [3, Theorem 3.3], we see that $(H(\mathcal{X}_{2^k/\sqrt{2}}))_{k\in\mathbb{Z}}$ is weakly $\sqrt{2}d^{0.25}$-interleaved with the scaled Rips module $(H(\mathcal{R}_{\alpha/d^{0.25}}))_{\alpha\geq 0}$. We can remove the scaling in the Rips filtration simply by multiplying both sides with $d^{0.25}$ and obtain our final approximation result.

▶ **Theorem 11.** *The persistence module* $\big(H(\mathcal{X}_{\frac{2^k\sqrt[4]{d}}{\sqrt{2}}})\big)_{k\in\mathbb{Z}}$ *is a* $3\sqrt{2}d^{0.25}$*-approximation of the Euclidean Rips persistence module* $\big(H_*(\mathcal{R}_\alpha)\big)_{\alpha\geq 0}$.

5 Size and computation

Set $n := |P|$ and let $CP(P)$ denote the closest pair distance of P. At scale $\alpha_0 := \frac{CP(P)}{3d}$ and lower, no d-cube of the cubical complex contains more than one active vertex, so the approximation complex consists of n isolated 0-simplices. At scale $\alpha_m := diam(P)$ and higher, points of P map to active vertices of a common face by Lemma 5, so the generated complex is acyclic using Lemma 4. We inspect the range of scales $[\alpha_0, \alpha_m]$ to construct the tower, since the barcode is explicitly known for scales outside this range. The total number of scales is $\lceil \log_2 \alpha_m/\alpha_0 \rceil = \lceil \log_2 \Delta + \log_2 3d \rceil = O(\log \Delta + \log d)$.

5.1 Size of the tower

Recall that the size of a tower is the number of simplices that do not have a preimage. We start by considering the case of 0-simplices.

▶ **Lemma 12.** *The number of* 0*-simplices included in the tower is at most* $n2^{O(d)}$.

The proof can be summarized as follows: 0-simplices in the tower correspond to active faces. Active vertices are only added at the lowest scale, hence they account for n inclusions. Active faces of higher dimensions have at least one active vertex on their boundary. We charge the inclusion of such a face to one point in P that is "close" to the face. In this way, we show that every point in P is charged at most $2^{O(d)}$ times. See [7] for further details.

The next lemma follows from a simple combinatorial counting argument for the number of flags in a d-dimensional cube (see [7]).

▶ **Lemma 13.** *Each* 0*-simplex of* $\mathcal{X}_\alpha$ *has at most* $2^{O(d \log k)}$ *incident* k*-simplices.*

▶ **Theorem 14.** *The* k*-skeleton of the tower has size at most* $n2^{O(d \log k)}$.

Proof. Let $\sigma = f_0 \subseteq \ldots \subseteq f_k$ be a flag included at some scale α. The crucial insight is that this can only happen if at least one face f_i in the flag is included in the tower at the same scale. Indeed, if each f_i has a preimage e_i on the previous scale, then $e_0 \subseteq \ldots \subseteq e_k$ is a flag on the previous scale which maps to σ under $\tilde{g}$.

We charge the inclusion of the flag to the inclusion of f_i. By Lemma 13, the 0-simplex f_i of $\mathcal{X}$ is charged at most $\sum_{i=1}^{k} 2^{O(d \log i)} = 2^{O(d \log k)}$ times in this way, and by Lemma 12, there are at most $n2^{O(d)}$ 0-simplices that can be charged. ◀

5.2 Computing the tower

Recall from the construction of the grids that G_{s+1} is built from G_s using an arbitrary translation vector $(\pm 1, \ldots, \pm 1) \in \mathbb{Z}^d$. In our algorithm, we pick the components of this translation vector uniformly at random, and independently for each scale.

Recall the cubical map $g_s : \square_s \to \square_{s+1}$ from Section 3. For a fixed s, we denote by $g^{(j)} : \square_s \to \square_{s+j}$ the j-fold composition of g, that is $g^{(j)} = g_{s+j-1} \circ g_{s+j-2} \circ \ldots \circ g_s$.

▶ **Lemma 15.** *For a* k*-face* f *of* $\square_s$*, let* Y *be the minimal integer* j *such that* $g^{(j)}(f)$ *is a vertex. Then* $E[Y] \leq 3 \log k$.

The proof idea is as follows. A k-face has a non-zero length in k coordinate direction. In order to map to a point, $g^{(j)}(f)$ has to "collapse" all these dimensions. For a fixed direction x_i, such a collapse happens for $g(f)$ if the random translation moves a grid point in the x_i-range of the face, which happens for exactly half of the translations (depending on the sign at the i-position of the translation vector). The number of steps for which the x_i-direction is not collapsed is thus equivalent to the number of flips of a fair coin until heads shows for the first time, which is 2 in expectation. The entire k-face is collapsed to a point if k coins flipped simultaneously all have shown heads at least once. This takes at most $3 \log k$ steps in expectation. See [7] for details.

As a consequence of the lemma, the expected "lifetime" of k-simplices in our tower with $k > 0$ is rather short: given a flag $e_0 \subseteq \ldots \subseteq e_\ell$, the face e_ℓ will be mapped to a vertex after $O(\log d)$ steps, and so will be all its sub-faces, turning the flag into a vertex. It follows that the total number of k-simplices in the tower is upper bounded by $n2^{O(d \log k)}$ as well.

Algorithm description. We first specify what it means to "compute" the tower. We make use of the fact that a simplicial map between simplicial complexes can be written as a composition of simplex inclusions and contractions of 0-simplices [9, 12]. That is, when passing from a scale α_s to α_{s+1}, it suffices to specify which pairs of 0-simplices in $\mathcal{X}_{\alpha_s}$ are mapped to the same image under $\tilde{g}$ and which simplices in $\mathcal{X}_{\alpha_{s+1}}$ are included.

The input is a set of n points $P \subset \mathbb{R}^d$. The output is a list of *events*, where each event is of one of the three following types: a *scale event* defines a real value α and signals that all upcoming events happen at scale α (until the next scale event). An *inclusion event* introduces a new simplex, specified by the list of 0-simplices on its boundary (we assume that every 0-simplex is identified by an integer). A *contraction event* is a pair of 0-simplices (i, j) and signifies that i and j are identified as the same from that scale.

In a first step, we calculate the range of scales that we are interested in. We compute a 2-approximation of $diam(P)$ by taking any point $p \in P$ and calculating $max_{q \in P} \|p - q\|$. Then we compute $CP(P)$ using a randomized algorithm in $n2^{O(d)}$ expected time [14].

Next, we proceed scale-by-scale and construct the list of events accordingly. On the lowest scale, we simply compute the active vertices by point location for P in a cubical grid, and enlist n inclusion events (this is the only step where the input points are considered in the algorithm). We use an auxiliary container S and maintain the invariant that whenever a new scale is considered, S consists of all simplices of the previous scale, sorted by dimension. In S, for each 0-simplex, we store an id and a coordinate representation of the active face to which it corresponds. Every ℓ-simplex with $\ell > 0$ is stored just as a list of integers, denoting its boundary 0-simplices. We initialize S with the n 0-simplices at the lowest scale.

Let $\alpha < \alpha'$ be any two consecutive scales with $\square, \square'$ the respective cubical complexes and $\mathcal{X}, \mathcal{X}'$ the approximation complexes, with $\tilde{g} : \mathcal{X} \to \mathcal{X}'$ being the simplicial map connecting them. Suppose we have already constructed all events at scale α. We enlist the scale event for α'. Then, we enlist the contraction events. For that, we iterate through the 0-simplices of $\mathcal{X}$ and compute their value under g, using point location in a cubical grid. We store the results in a list S' (which contains the simplices of $\mathcal{X}'$). If for a 0-simplex j, $g(j)$ is found to be equal to $g(i)$ for a previously considered 0-simplex, we choose the minimal such i and enlist a contraction event for i and j.

We turn to the inclusion events and start with the case of 0-simplices. Every 0-simplex is an active face at scale α' and must contain an active vertex, which is also a 0-simplex of $\mathcal{X}'$. We iterate through the elements in S'. For each active vertex v encountered, we go over all faces of the cubical complex $\square'$ that contain v as vertex and check whether they are active. For every active face encountered that is not in S' yet, we add it to S' and enlist an inclusion event of a new 0-simplex. At termination, all 0-simplices of $\mathcal{X}'$ have been detected.

Next, we iterate over the simplices of S of dimension ≥ 1 and compute their image under $\tilde{g}$, and store the result in S'. To find the simplices of dimension ≥ 1 included at $\mathcal{X}'$, we exploit our previous insight that they contain at least one 0-simplex that is included at the same scale (see the proof of Theorem 14). Hence, we iterate over the 0-simplices included in $\mathcal{X}'$ and proceed inductively in dimension. Let v be the current 0-simplex under consideration; assume that we have found all $(p-1)$-simplices in $\mathcal{X}'$ that contain v. Each such $(p-1)$-simplex σ is a flag in $\square'$. We iterate over all faces e that extend σ to a flag of length $p+1$. If e is active, we found a p-simplex in $\mathcal{X}'$. If this simplex is not in S' yet, we add it and enlist an inclusion event for it. We also enqueue the simplex in our inductive procedure, to look for $(p+1)$-simplices in the next iteration. At the end of the procedure, we have detected all simplices in $\mathcal{X}'$ without preimage, and S' contains all simplices of $\mathcal{X}'$. We set $S \leftarrow S'$ and proceed to the next scale. This ends the description of the algorithm.

▶ **Theorem 16.** *To compute the k-skeleton, the algorithm takes time $\left(n2^{O(d)}\log\Delta + 2^{O(d)}M\right)$ time in expectation and M space, where M is the size of the tower. In particular, the expected time is bounded by $\left(n2^{O(d)}\log\Delta + n2^{O(d\log k)}\right)$ and the space is bounded by $n2^{O(d\log k)}$.*

The first summand of the time bound comes from the fact that on each scale, the number of 0-simplices of $\mathcal{X}$ is bounded by $n3^d$, and we employ local searches in the cubical complex to find 0-simplices included in $\mathcal{X}'$. This local search only causes an overhead of $O(2^d)$ per active vertex. The second summand arises because we find the higher-dimensional simplices of $\mathcal{X}'$ inductively and can therefore charge the cost for this search to the number of simplices encountered. Finally, computing the image of $\tilde{g}$ for all simplices in $\mathcal{X}$ can be bounded in expectation by $O(2^{O(d)}M)$, because the total size of all $\mathcal{X}$ in the algorithm is bounded by $O(\log dM)$ (see the remark after Lemma 15). More details are in [7].

6 Conclusion

We gave an approximation scheme for the Rips filtration, with improved approximation ratio, size and computational complexity than previous approaches for the case of high-dimensional point clouds. Moreover, we introduced the technique of using acyclic carriers to prove interleaving results. We point out that, while the proof of the interleaving in Section 4.1 is still technically challenging, it greatly simplifies by the usage of acyclic carriers; defining the interleaving chain maps explicitly significantly blows up the analysis. There is also no benefit in knowing the interleaving maps because they are only required for the analysis, not for the computation.

Our tower is connected by simplicial maps; there are (implemented) algorithms to compute the barcode of such towers [9, 12]. It is also quite easy to adapt our tower construction to a streaming setting [12], where the output list of events is passed to an output stream instead of being stored in memory.

An interesting question is whether persistence can be computed efficiently for more general chain maps, which would allow more freedom in building approximation schemes.

References

1. M. Botnan and G. Spreemann. Approximating Persistent Homology in Euclidean space through collapses. *Applied Algebra in Engineering, Communication and Computing*, 26(1-2):73–101, 2015.
2. P. Bubenik, V. de Silva, and J. Scott. Metrics for Generalized Persistence Modules. *Foundations of Computational Mathematics*, 15(6):1501–1531, 2015.
3. P. Bubenik and J.A. Scott. Categorification of Persistent Homology. *Discrete & Computational Geometry*, 51(3):600–627, 2014.

4 G. Carlsson. Topology and Data. *Bulletin of the American Mathematical Society*, 46:255–308, 2009.

5 N.J. Cavanna, M. Jahanseir, and D. Sheehy. A Geometric Perspective on Sparse Filtrations. In *Proceedings of the 27th Canadian Conference on Computational Geometry (CCCG 2015)*.

6 F. Chazal, D. Cohen-Steiner, M. Glisse, L.J. Guibas, and S.Y. Oudot. Proximity of Persistence Modules and their Diagrams. In *Proceedings of the ACM Symposium on Computational Geometry (SoCG 2009)*, pages 237–246, 2009.

7 A. Choudhary, M. Kerber, and S. Raghvendra. Improved approximate rips filtrations with shifted integer lattices. *CoRR*, abs/1706.07399, 2016. URL: `https://arxiv.org/abs/1706.07399`.

8 A. Choudhary, M. Kerber, and S. Raghvendra. Polynomial-sized topological approximations using the Permutahedron. In *Proceedings of the 32nd International Symposium on Computational Geometry (SoCG 2016)*, pages 31:1–31:16, 2016.

9 T.K. Dey, F. Fan, and Y. Wang. Computing Topological Persistence for Simplicial maps. In *Proceedings of the ACM Symposium on Computational Geometry (SoCG 2014)*, pages 345–354, 2014.

10 H. Edelsbrunner and J. Harer. *Computational Topology - An Introduction*. American Mathematical Society, 2010.

11 H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological Persistence and Simplification. *Discrete & Computational Geometry*, 28(4):511–533, 2002.

12 M. Kerber and H. Schreiber. Barcodes of Towers and a Streaming Algorithm for Persistent Homology. In *Proceedings of the 33rd International Symposium on Computational Geometry (SoCG)*, pages 57:1–57:15, 2017.

13 M. Kerber and R. Sharathkumar. Approximate Čech Complex in Low and High Dimensions. In *International Symposium on Algorithms and Computation (ISAAC)*, pages 666–676, 2013.

14 S. Khuller and Y. Matias. A Simple Randomized Sieve Algorithm for the Closest-Pair Problem. *Information and Computation*, 118(1):34–37, 1995.

15 J.R. Munkres. *Elements of algebraic topology*. Westview Press, 1984.

16 D. Sheehy. Linear-size Approximations to the Vietoris-Rips Filtration. *Discrete & Computational Geometry*, 49(4):778–796, 2013.