CrossMark

Polynomial-Sized Topological Approximations Using the Permutahedron

Aruni Choudhary[1] ⬤ · **Michael Kerber[2]** ·
Sharath Raghvendra[3]

Abstract Classical methods to model topological properties of point clouds, such as the Vietoris–Rips complex, suffer from the combinatorial explosion of complex sizes. We propose a novel technique to approximate a multi-scale filtration of the Rips complex with improved bounds for size: precisely, for n points in $\mathbb{R}^d$, we obtain a $O(d)$-approximation whose k-skeleton has size $n2^{O(d \log k)}$ per scale and $n2^{O(d \log d)}$ in total over all scales. In conjunction with dimension reduction techniques, our approach yields a $O(\text{polylog}(n))$-approximation of size $n^{O(1)}$ for Rips filtrations on arbitrary metric spaces. This result stems from high-dimensional lattice geometry and exploits properties of the permutahedral lattice, a well-studied structure in discrete geometry. Building on the same geometric concept, we also present a lower bound result on the size of an approximation: we construct a point set for which every $(1 + \varepsilon)$-approximation of the Čech filtration has to contain $n^{\Omega(\log \log n)}$ features, provided that $\varepsilon < \frac{1}{\log^{1+c} n}$ for $c \in (0, 1)$.

Aruni Choudhary
aruni.choudhary@mpi-inf.mpg.de

Michael Kerber
kerber@tugraz.at

Sharath Raghvendra
sharathr@vt.edu

[1] Max Planck Institute for Informatics, Saarland Informatics Campus, 66123 Saarbrücken, Germany

[2] Institut für Geometrie, Graz University of Technology, 8010 Graz, Austria

[3] Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA

🖄 Springer

1 Introduction

1.1 Motivation and Previous Work

Topological data analysis aims at finding and reasoning about the underlying topological features of metric spaces. The idea is to represent a data set by a set of discrete structures on a range of scales and to track the evolution of homological features as the scale varies. The theory of *persistent* homology allows for a topological summary, called the *persistence diagram* which summarizes the lifetimes of topological features in the data as the scale under consideration varies monotonously. A major step in the computation of this topological signature is the question of how to compute a multi-scale representation of a given data set.

For data in the form of finite point clouds, two frequently used constructions are the (*Vietoris–)Rips* complex $\mathcal{R}_\alpha$ and the *Čech* complex $\mathcal{C}_\alpha$ which are defined with respect to a scale parameter $\alpha \geq 0$. Both are simplicial complexes capturing the proximity of points at scale α, with different levels of accuracy. Increasing α from 0 to ∞ yields a nested sequence of simplicial complexes called a *filtration*.

Unfortunately, Rips and Čech filtrations can be uncomfortably large to handle. For homological features in low dimensions, it suffices to consider the k-skeleton of the complex, that is, all simplices of dimension at most k. Still, the k-skeleton of Rips and Čech complexes can be as large as n^{k+1} for n points, which is already impractical for small k when n is large.

One remedy is to approximate the topological signature of the filtration. A common way to do this is to construct a *tower*, which is a sequence of simplicial complexes connected by simplicial maps. The size of a tower is the number of simplices which are included in the sequence of simplicial complexes. A tower is said to approximate a filtration, if it yields a topological signature similar to that of the original filtration, while having a significantly smaller size. The notion of "similarity" in this context can be made formal through a distance measure on persistence diagrams. The most frequently used similarity measure is the *bottleneck distance*, which finds correspondences between topological features of two towers, such that the lifetimes of each pair of matched features are as close as possible. A related notion is the *log-scale bottleneck distance* which allows a larger discrepancy for larger scales and thus can be seen as a relative approximation, with usual bottleneck distance as its absolute counterpart. We call an approximate tower a *c-approximation* of the original, if their persistence diagrams have log-scale bottleneck distance at most c.

Sheehy [28] gave the first such approximate tower for Rips complexes with a formal guarantee. For $0 < \varepsilon \leq 1/3$, he constructs a $(1 + \varepsilon)$-approximate tower of the Rips filtration. The approximation tower of [28] consists of a filtration of simplicial complexes, whose k-skeleton has size only $n(1/\varepsilon)^{O(\lambda k)}$, where λ is the doubling

dimension of the metric. Since then, several alternative techniques have been explored for Rips [12] and Čech complexes [5,22], all arriving at the same complexity bound.

While the above approaches work well for instances where λ and k are small, we focus on high-dimensional point sets. This has two reasons: first, one might simply want to analyze data sets for which the intrinsic dimension is high, but the existing methods do not succeed in reducing the complex size sufficiently. Second, even for medium-size dimensions, one might not want to restrict its scope to the low-homology features, so that $k = \lambda$ is not an unreasonable parameter choice. To adapt the afore-mentioned schemes to play nice with high dimensional point clouds, it makes sense to use dimension reduction results to eliminate the dependence on λ. Indeed, it has been shown, in analogy to the famous Johnson–Lindenstrauss Lemma [18], that an orthogonal projection of a point set of $\mathbb{R}^d$ to a $O(\log n/\varepsilon^2)$-dimensional subspace yields a $(1 + \varepsilon)$ approximation of the Čech filtration [20,29]. Combining these two approximation schemes, however, yields an approximation of size $O(n^{k+1})$ (ignoring ε-factors) and does not improve upon the exact case.

1.2 Our Contributions

We present two results about the approximation of Rips and Čech filtrations: we give a scheme for approximating the Rips filtration with smaller complex size than existing approaches, at the price of guaranteeing only an approximation quality of $\mathrm{polylog}(n)$. Since Rips and Čech filtrations approximate each other by a constant factor, our result also extends to the Čech filtration, with an extra constant factor in the approximation quality. Second, we prove that any approximation scheme for the Čech filtration has superpolynomial size in n if high accuracy is required. For this result, our proof technique does not extend to Rips complexes. In more detail, our results are as follows:

Upper bound. We present a $6(d+1)$-approximation of the Rips filtration for n points in $\mathbb{R}^d$. The construction scheme is output-sensitive in the size of the approximate tower. We show that the k-skeleton of the approximate tower has size $n2^{O(d \log k)}$ per scale. Through a randomized selection of the origin at each scale, we show that the tower has size $n2^{O(d \log d)}$ over all scales, in expectation. Here, the expectation is over the random choice of the origin, and not on the input point set. This shows that by using a more rough approximation, we can achieve asymptotic improvements on the complex size. We present two algorithms for computing the approximation, whose expected runtimes are upper bounded by $n2^{O(d)} \log \Delta + n2^{O(d \log d)}$ and $(O(nd^2) + \mathrm{poly}(d)) \log \Delta + n \log n2^{O(d)} + n2^{O(d \log d)}$, where Δ is the spread of the point set. The expected space required by the algorithms is upper bounded by $n2^{O(d \log d)}$. The real power of our approach reveals itself in high dimensions, in combination with dimension reduction techniques. In conjunction with the lemma of Johnson and Lindenstrauss [18], we obtain a $O(\log n)$-approximation of expected size $n^{O(\log \log n)}$, which is much smaller than the original tower; however, the bound is still super-polynomial in n. Combined with a different dimension reduction result of Matoušek [25], we obtain a $O(\log^{3/2} n)$-approximation of expected size $n^{O(1)}$. This is the first polynomial bound in n of an approximate tower, independent of the dimensionality of the point set. For inputs

from arbitrary metric spaces (instead of points in $\mathbb{R}^d$), the same results hold with an additional $O(\log n)$ factor in the approximation quality.

Lower bound. We construct a point set of n points in $d = \Theta(\log n)$ dimensions whose Čech filtration has $n^{\Omega(\log \log n)}$ persistent features with "relatively long" lifetime. Precisely, that means that any $(1 + \delta)$-approximation has to contain a bar of non-zero length for each of those features if $\delta < O(\frac{1}{\log^{1+c} n})$ with $c \in (0, 1)$. This shows that it is impossible to define an approximation scheme that yields an accurate approximation of the Čech complexes as well as polynomial size in n.

Methods. Our results follow from a link to lattice geometry: the A^*-lattice is a configuration of points in $\mathbb{R}^d$ which realizes the thinnest known coverings for low dimensions [11]. The dual Voronoi polytope of a lattice point is the *permutahedron*, whose vertices are obtained by all coordinate permutations of a fixed point in $\mathbb{R}^d$.

Our technique resembles the perhaps simplest approximation scheme for point sets: if we digitize $\mathbb{R}^d$ with d-dimensional pixels, we can take the union of pixels that contain input points as our approximation. Our approach does the same, except that we use a tessellation of permutahedra for digitization. In $\mathbb{R}^2$, our approach corresponds to the common approach of replacing the square tiling by a hexagonal tiling. We utilize the fact that the permutahedral tessellation is in generic position, that is, no more than $d + 1$ polytopes have a common intersection. At the same time, permutahedra are still relatively round, that is, they have small diameter and non-adjacent polytopes are well-separated. These properties ensure good approximation quality and a small complex. In comparison, a cubical tessellation yields a $O(\sqrt{d})$-approximation of the Rips filtration which looks like an improvement over our $O(d)$-approximation, but the highly degenerate configuration of the cubes yields a complex size of $n2^{O(dk)}$, and therefore does not constitute an improvement over Sheehy's approach [28].

For the lower bound, we arrange n points in a way that one center point has the permutahedron as Voronoi polytope, and we consider simplices incident to that center point in a fixed dimension. We show a superpolynomial number of these simplices create or destroy topological features of non-negligible persistence.

1.3 Updates from the Conference Version

An earlier version of this work appeared at the 32nd International Symposium on Computational Geometry [10]. There, we gave a naive upper bound on the size of the tower: if Δ denotes the spread of the point set, then a simple upper bound on the size of the tower is $n2^{O(d \log k)} \log \Delta$. In the current version, we show that the dependence on spread can be removed, by a randomized translation of the A_d^* lattice at each scale of the tower. With this technique, we show that there are $n2^{O(d \log d)}$ simplices in the tower in expectation, which is independent of the spread Δ. In the conference version, an upper bound for the time to compute the approximate tower was shown to be $n2^{O(d \log k)} \log \Delta$. In this version, we present two algorithms to compute the approximation tower, with better upper bounds for the runtime. Also, the current

version of the paper contains the proofs which were missing from the conference version.

1.4 Outline of the Paper

We begin by reviewing basics of persistent homology in Sect. 2. Next, we study several relevant properties of the A^* lattice in Sect. 3. An approximation scheme based on concepts from Sect. 3 is presented in Sect. 4. We present the computational aspects of the scheme from Sect. 4 in Sect. 5. In Sect. 6, we present the effects of dimension reduction techniques on our approximation scheme. In Sect. 7, we present the lower bound result on the size of approximations of Čech filtration. We conclude in Sect. 8.

2 Topological Background

We review some topological concepts needed in our argument. More extensive treatments covering most of the material can be found in the textbooks [13,17,26].

2.1 Simplicial Complexes

For an arbitrary set V, called *vertices*, a *simplicial complex* over V is a collection of non-empty subsets which is closed under taking non-empty subsets. The elements of a simplicial complex K are called *simplices* of K. A simplex σ is a *face* of τ if $\sigma \subseteq \tau$. A *facet* is a face of co-dimension 1. The dimension of σ is $k := |\sigma| - 1$; we also call σ a k-simplex in this case. The *k-skeleton* of K is the collection of all simplices of dimension at most k. For instance, the 1-skeleton of K is a graph defined by its 0- and 1-simplices.

We discuss two ways of generating simplicial complexes. In the first one, take a collection $\mathcal{S}$ of sets over a common universe (for instance, polytopes in $\mathbb{R}^d$), and define the *nerve* of $\mathcal{S}$ as the simplicial complex whose vertex set is $\mathcal{S}$, and a k-simplex σ is in the nerve if the corresponding $k + 1$ sets have a non-empty common intersection. The *nerve theorem* [4] states that if all sets in $\mathcal{S}$ are convex subsets of $\mathbb{R}^d$, their nerve is homotopically equivalent to the union of the sets (the statement can be generalized significantly; see [17, Sect. 4.G]). The second construction that we consider are *flag complexes*: Given a graph $G = (V, E)$, we define a simplicial complex K_G over the vertex set V such that a k-simplex σ is in K if for every distinct pair of vertices $v_1, v_2 \in \sigma$, the edge (v_1, v_2) is in E. In other words, K_G is the maximal simplicial complex with G as its 1-skeleton. In general, a complex K is called a flag complex, if $K = K_G$ with G being the 1-skeleton of K.

Given a set of points P in $\mathbb{R}^d$ and a parameter r, the *Čech complex at scale r*, $\mathcal{C}_r$ is defined as the nerve of the balls centered at the elements of P, each of radius r. This is a collection of convex sets. Therefore, the nerve theorem is applicable and it asserts that the nerve agrees homotopically with the union of balls. In the same setup, we can as well consider the intersection graph G of the balls (that is, we have an edge between two points if their distance is at most $2r$). The flag complex of G is called

the (*Vietoris–*)*Rips complex at scale r*, denoted by $\mathcal{R}_r$. The relation $\mathcal{C}_r \subseteq \mathcal{R}_r \subseteq \mathcal{C}_{\sqrt{2}r}$ follows from Jung's Theorem [19].

2.2 Persistence Modules and Simplicial Towers

A *persistence module* $(V_\alpha)_{\alpha \in G}$ for a totally ordered index set $G \subseteq \mathbb{R}$ is a sequence of vector spaces with linear maps $F_{\alpha, \alpha'} \colon V_\alpha \to V_{\alpha'}$ for any $\alpha \leq \alpha'$, satisfying $F_{\alpha, \alpha} = \mathrm{id}$ and $F_{\alpha', \alpha''} \circ F_{\alpha, \alpha'} = F_{\alpha, \alpha''}$. Persistence modules can be decomposed into *indecomposable intervals* giving rise to a *persistent barcode* which is a complete discrete invariant of the corresponding module.

A distance measure between persistence modules is defined through interleavings: we call two modules (V_α) and (W_α) with linear maps $F_{\cdot, \cdot}$ and $G_{\cdot, \cdot}$ *additively ε-interleaved*, if there exist linear maps $\phi \colon V_\alpha \to W_{\alpha + \varepsilon}$ and $\psi \colon W_\alpha \to V_{\alpha + \varepsilon}$ such that the maps ϕ and ψ commute with F and G (see [9]). We call the modules *multiplicatively c-interleaved* with $c \geq 1$, if there exist linear maps $\phi \colon V_\alpha \to W_{c\alpha}$ and $\psi \colon W_\alpha \to V_{c\alpha}$ with the same commuting properties. Equivalently, this means that the modules are additively $(\log c)$-interleaved when switching to a logarithmic scale. In this case, we also call the module (G_α) a *c-approximation* of (F_α) (and vice versa). Note that the case $c = 1$ implies that the two modules give rise to the same persistent barcode, which is usually referred to as the *persistence equivalence theorem* [13].

The most common way to generate persistence modules is through the homology of sequences of simplicial complexes. Let K, L be two simplicial complexes with a vertex map $f \colon K \to L$, such that for each simplex $\sigma = (v_0, \ldots, v_k) \in K$, there is a simplex $(f(v_0), \ldots, f(v_k)) \in L$. The linear extension of f to simplices of K is called a *simplicial map* induced by f. A (*simplicial*) *tower* $(K_\alpha)_{\alpha \in G}$ over a totally ordered index set $G \subseteq \mathbb{R}$ is a sequence of simplicial complexes connected by simplicial maps $f_{\alpha, \alpha'} \colon K_\alpha \to K_{\alpha'}$ for any $\alpha \leq \alpha'$, such that $f_{\alpha, \alpha} = \mathrm{id}$ and $f_{\alpha', \alpha''} \circ f_{\alpha, \alpha'} = f_{\alpha, \alpha''}$. By the functorial properties of homology (using some fixed field $\mathbb{F}$ and some fixed dimension $p \geq 0$), such a tower gives rise to a persistence module $(H_p(K_\alpha, \mathbb{F}))_{\alpha \in G}$ [26]. We call a tower a *c-approximation* of another tower if the corresponding persistence modules induced by homology are *c-approximations* of each other.

The standard way of obtaining a tower is through a nested sequence of simplicial complexes, where the simplicial maps are induced by inclusion. Such towers are called *filtrations*. Examples are the *Čech filtration* $(\mathcal{C}_\alpha)_{\alpha \in \mathbb{R}}$ and the *Rips filtration* $(\mathcal{R}_\alpha)_{\alpha \in \mathbb{R}}$. By the relation of Rips and Čech complexes from above, the Rips filtration is a $\sqrt{2}$-approximation of the Čech filtration.

Any simplicial tower can be written in the form $K_1 \to \cdots \to K_M$ where each K_i differs from K_{i-1} in an elementary fashion [12,21]. Either each K_i contains one more simplex than K_{i-1} (which is called an *elementary inclusion*), or a pair of vertices of K_{i-1} collapse into one in K_i (*elementary contraction*). In particular, if $K_i \setminus K_{i-1}$ is a vertex, then we call the elementary inclusion as a *vertex inclusion*. Each contraction collapses simplices which were introduced at an earlier scale. A simplex which contracts does not re-appear in the tower. As a result, the number of inclusions is at least the number of contractions in the tower. The size of a tower is then defined as the total number of elementary inclusions in this sequence of simplicial complexes.

2.3 Simplex-Wise Čech Filtration and (Co-)Face Distances

In the Čech filtration $(\mathcal{C}_\alpha)$, every simplex has an *alpha value* $\alpha_\sigma := \min\{\alpha \geq 0 \mid \sigma \in \mathcal{C}_\alpha\}$, which equals the radius of the minimal enclosing ball of its boundary vertices. If the point set P is finite, the Čech filtration consists of a finite number of simplices, and we can define a *simplex-wise tower*

$$\emptyset = \mathcal{C}^0 \subsetneq \mathcal{C}^1 \subsetneq \cdots \subsetneq \mathcal{C}^m,$$

where exactly one simplex is added from $\mathcal{C}^i$ to $\mathcal{C}^{i+1}$, and where σ is added before τ whenever $\alpha_\sigma < \alpha_\tau$. The tower is not unique and ties can be broken arbitrarily.

In a simplex-wise tower, passing from $\mathcal{C}^i$ to $\mathcal{C}^{i+1}$ means adding the k-simplex $\sigma := \sigma_{i+1}$. The effect of this addition is that either a k-homology class comes into existence, or a $(k-1)$-homology class is destroyed. Depending on the case, we call σ *positive* or *negative*, accordingly. In terms of the corresponding persistent barcode, there is exactly one interval *associated to* σ either starting at i (if σ is positive) or ending at i (if σ is negative). We define the *(co-)face distance* L_σ (L_σ^*) of σ as the minimal distance between α_σ and its (co-)facets,

$$L_\sigma := \min_{\tau \text{ facet of } \sigma} \alpha_\sigma - \alpha_\tau, \qquad L_\sigma^* := \min_{\tau \text{ co-facet of } \sigma} \alpha_\tau - \alpha_\sigma.$$

Note that L_σ and L_σ^* can be zero. Nevertheless, they constitute lower bounds for the persistence of the associated barcode intervals. An alternative to our proof is to argue using structural properties of the matrix reduction algorithm for persistent homology [13].

Lemma 2.1 *If σ is negative, the barcode interval associated to σ has persistence at least L_σ.*

Proof σ kills a $(k-1)$-homology class by assumption, and this class is represented by the cycle $\partial\sigma$. However, this cycle came into existence when the last facet τ of σ was added. Therefore, the lifetime of the cycle destroyed by σ is at least $\alpha_\sigma - \alpha_\tau$. $\square$

Lemma 2.2 *If σ is positive, the homology class created by σ has persistence at least L_τ^*.*

Proof σ creates a k-homology class; every representative cycle of this class is non-zero for σ. To turn such a cycle into a boundary, we have to add a $(k+1)$-simplex τ with σ in its boundary (otherwise, any $(k+1)$-chain formed will be zero for σ). Therefore, the cycle created at σ persists for at least $\alpha_\tau - \alpha_\sigma$. $\square$

3 The A^*-Lattice and the Permutahedron

A *lattice L* in $\mathbb{R}^d$ is the set of all integer-valued linear combinations of d independent vectors, called the *basis* of the lattice. Note that the origin belongs to every lattice. The *Voronoi polytope* of a lattice L is the closed set of all points in $\mathbb{R}^d$ for which the origin

is among the closest lattice points. Since lattices are invariant under translations, the Voronoi polytopes for other lattice points are just translations of the one at the origin, and these polytopes tile $\mathbb{R}^d$. An elementary example is the integer lattice, spanned by the unit vectors $(e_1, \ldots, e_d)$, whose Voronoi polytope is the unit d-cube, shifted by $-1/2$ in each coordinate direction.

We are interested in a different lattice, called the A_d^*-lattice, whose properties are also well-studied [11]. First, we define the A_d lattice as the set of points $(x_1, \ldots, x_{d+1}) \in \mathbb{Z}^{d+1}$ satisfying $\sum_{i=1}^{d+1} x_i = 0$. A_d is spanned by vectors of the form $(e_i, -1)$, $i = 1, \ldots, d$. While it is defined in $\mathbb{R}^{d+1}$, all points lie on the hyperplane H defined by $\sum_{i=1}^{d+1} y_i = 0$. After a suitable change of basis, we can express A_d by d vectors in $\mathbb{R}^d$; thus, it is indeed a lattice. In low dimensions, A_2 is the hexagonal lattice, and A_3 is the FCC lattice that realizes the best sphere packing configuration in $\mathbb{R}^3$ [15].

The *dual lattice* L^* of a lattice L is defined as the set of points $(y_1, \ldots, y_d)$ in $\mathbb{R}^d$ such that $y \cdot x \in \mathbb{Z}$ for all $x \in L$ [11]. Both the integer lattice and the hexagonal lattice are self-dual, while the dual of A_3 is the BCC lattice that realizes the thinnest sphere covering configuration among lattices in $\mathbb{R}^3$ [2].

We are mostly interested in the Voronoi polytope Π_d generated by A_d^*. Again, the definition becomes easier when embedding $\mathbb{R}^d$ one dimension higher as the hyperplane H. In that representation, it is known [11] that Π_d has $(d + 1)!$ vertices obtained by all permutations of the coordinates of

$$\frac{1}{2(d + 1)} (d, d - 2, d - 4, \ldots, -d + 2, -d).$$

Π_d is known as the *permutahedron* [32, Lect. 0].[1] Our approximation results in Sects. 4 and 7 are based on various combinatorial and geometric properties of Π_d, which we describe next. We will fix d and write $A^* := A_d^*$ and $\Pi := \Pi_d$ for brevity.

3.1 Combinatorics

The k-faces of Π correspond to ordered partitions of the coordinate indices $[d+1] := \{1, \ldots, d + 1\}$ into $(d + 1 - k)$ non-empty ordered subsets $\{S_1, \ldots, S_{d+1-k}\}$ such that all coordinates in S_i are smaller than all coordinates in S_j for $i < j$ [32]. For example, with $d = 3$, the partition $(\{1, 3\}, \{2, 4\})$ is the 2-face spanned by all points for which the two smallest coordinates appear at the first and the third position. This is an example of a facet of Π, for which we need to partition the indices in exactly two subsets; equivalently, the facets of Π are in one-to-one correspondence to non-empty proper subsets of $[d + 1]$ so Π has $2^{d+1} - 2$ facets. The vertices of Π are the $(d + 1)$-fold ordered partitions which correspond to permutations of $[d + 1]$, reassuring the fact that Π has $(d + 1)!$ vertices. Moreover, two faces σ, τ of Π with $\dim \sigma < \dim \tau$ are incident if the partition of σ is a refinement of the partition of τ. Continuing

---

[1] Often, a scaled, translated and rotated version is considered, in which all permutations of the point $(1, \ldots, d + 1)$ are taken.

our example from before, the four 1-faces bounding the 2-face $(\{1, 3\}, \{2, 4\})$ are $(\{1\}, \{3\}, \{2, 4\}), (\{3\}, \{1\}, \{2, 4\})$, $(\{1, 3\}, \{2\}, \{4\})$, and $(\{1, 3\}, \{4\}, \{2\})$. Vice versa, we obtain co-faces of a face by combining consecutive partitions into one larger partition. For instance, the two co-facets of $(\{1, 3\}, \{4\}, \{2\})$ are $(\{1, 3\}, \{2, 4\})$ and $(\{1, 3, 4\}, \{2\})$.

Lemma 3.1 *Let σ, τ be two facets of Π, defined using the two partitions $(S_\sigma, [d+1]\backslash S_\sigma)$ and $(S_\tau, [d+1]\backslash S_\tau)$, respectively. Then σ and τ are adjacent in Π iff $S_\sigma \subseteq S_\tau$ or $S_\tau \subseteq S_\sigma$.*

Proof Two facets are adjacent if they share a common face. By the properties of the permutahedron, this means that the two facets are adjacent if and only if their partitions permit a common refinement, which is only possible if one set is contained in the other. □

We have already established that Π has "few" $(2^{d+1} - 2 = O(2^d))$ $(d-1)$-faces and "many" $((d+1)! = O(2^{d\log d}))$ 0-faces. We give an interpolating bound for all intermediate dimensions.

Lemma 3.2 *The number of $(d-k)$-faces of Π is bounded by $2^{3(d+1)\log_2(k+1)}$.*

Proof By our characterization of faces of Π, it suffices to count the number of ordered partitions of $[d+1]$ into $k+1$ subsets. That number equals $(k+1)!$ times the number of unordered partitions. The number of unordered partitions, in turn, is known as *Stirling number of the second kind* [27] and is bounded by $\frac{1}{2}\binom{d+1}{k+1}(k+1)^{d-k}$. To get an upper bound for the number of $(d-k)$-faces, we multiply the Stirling number with $(k+1)!$ and get

$$\frac{1}{2}\binom{d+1}{k+1}(k+1)^{d-k}(k+1)! \leq (d+1)^{k+1}(k+1)^{d-k}(k+1)!$$
$$\leq (d+1)^{k+1}(k+1)^{d-k}(k+1)^{k+1} \leq (d+1)^{k+1}(k+1)^{d+1}$$
$$\leq (k+1)^{3(d+1)} = 2^{3(d+1)\log_2(k+1)},$$

where we have used the fact that $(d+1)^{k+1} \leq (k+1)^{2(d+1)}$ for $k \leq d$. □

3.2 Geometry

All vertices of Π are equidistant from the origin, and it can be checked with a simple calculation that this distance is $\sqrt{\frac{d(d+2)}{12(d+1)}}$. Using the triangle inequality, we obtain:

Lemma 3.3 *The diameter of Π is at most $\sqrt{d}$.*

The permutahedra centered at all lattice points of A^* define the Voronoi tessellation of A^*. Its nerve is the Delaunay triangulation $\mathcal{D}$ of A^*. An important property of A^* is that, unlike for the integer lattice, $\mathcal{D}$ is non-degenerate – this will ultimately ensure small upper bounds for the size of our approximation scheme.

Lemma 3.4 *Each vertex of a permutahedral cell has precisely $d + 1$ cells adjacent to it. In other words, the A_d^* lattice points are in general position. As a consequence, we can identify Delaunay simplices incident to the origin with faces of Π.*

Proof To prove the claim, the idea is to look at any vertex of the Voronoi cell and argue that it has precisely $d + 1$ equidistant lattice points. See [1, Thm. 2.5] for a concise argument. Here, we rephrase the proof idea of [1, Thm. 2.5] in slightly simplified terms.

The basis vectors of A_d^* are of the form

$$g_t = \frac{1}{(d+1)} \underbrace{(t, \ldots, t}_{d+1-t}, \underbrace{t - (d+1), \ldots, t - (d+1))}_{t}$$

for $1 \le t \le d$ [11]. It can be seen that each component of the numerator of g_t is congruent to t modulo $(d + 1)$. Hence, we call the numerator of g_t a *remainder-t point*. Since any lattice point x in A_d^* can be written as $x = \sum m_t \cdot g_t$, it follows that the numerator of x is a remainder-$\{(\sum m_t \cdot t) \text{ modulo } (d + 1)\}$ point.

Now, we show that the Delaunay cells of the A_d^* lattice are all d-simplices, which will prove our claim. Let $\mathbf{v}$ be a vertex of the permutahedron which is the Voronoi cell of the origin. Without loss of generality, we can assume that $\mathbf{v} = \frac{1}{2(d+1)}(d, d-2, \ldots, -d)$. The A_d^* lattice points closest to $\mathbf{v}$ define the Delaunay cell of $\mathbf{v}$. We have seen that the lattice points have the form $\mathbf{y} = \frac{1}{d+1}(\mathbf{m}(d+1) + k\mathbf{1})$, where $\mathbf{m} \in \mathbb{Z}^{d+1}$. Also, $\mathbf{m} \cdot \mathbf{1} = -k$ because $\mathbf{y} \cdot \mathbf{1} = 0$.

We wish to minimize the distance between v and y by choosing a suitable value for m. In other words, we wish to find $\text{argmin}_{\mathbf{m}} \|\mathbf{y} - \mathbf{v}\|^2$ for a fixed k:

$$\begin{aligned}
\text{argmin}_{\mathbf{m}} \|\mathbf{y} - \mathbf{v}\|^2 &= \text{argmin}_{\mathbf{m}} \sum \left(m_i + \frac{k}{d+1} - v_i \right)^2 \\
&= \text{argmin}_{\mathbf{m}} \sum (m_i - v_i)^2 + 2(m_i - v_i) \frac{k}{d+1} \\
&= \text{argmin}_{\mathbf{m}} \sum (m_i - v_i)^2 + \frac{2k}{d+1} \sum m_i \\
&= \text{argmin}_{\mathbf{m}} \sum (m_i - v_i)^2 + \frac{2k}{d+1} \cdot (-k) \\
&= \text{argmin}_{\mathbf{m}} \sum (m_i - v_i)^2 \\
&= \text{argmin}_{\mathbf{m}} \|\mathbf{m} - \mathbf{v}\|^2 \\
&= \text{argmin}_{\mathbf{m}} \left\| \mathbf{m} - \frac{1}{2(d+1)}(d, \ldots, -d) \right\|^2.
\end{aligned}$$

Using $\mathbf{m} \cdot \mathbf{1} = -k$ and an elementary calculation, we see that $\|\mathbf{y} - \mathbf{v}\|^2$ is minimized for

$$\mathbf{m} = (\underbrace{0, \ldots, 0}_{d+1-k}, \underbrace{-1, \ldots, -1}_{k}),$$

given any fixed k. This means that there is a unique remainder-k nearest lattice point to $\mathbf{v}$, for $k \in \{0, \ldots, d\}$. Moreover, the corresponding lattice points $\mathbf{y}$ are Delaunay neighbors of the origin, and are equidistant from $\mathbf{v}$. The Delaunay cell corresponding to $\mathbf{v}$ contains precisely $(d + 1)$ points, one for each value of k, which proves the claim for the vertex $\mathbf{v}$.

Recall that any other vertex $\mathbf{u}$ of Π can be written as some permutation π of $\mathbf{v}$, that is, $\mathbf{u} = \pi(\mathbf{v})$. Following the above derivation, the nearest lattice points for $\mathbf{u}$ can be found by simply applying the permutation π on the nearest lattice points for $\mathbf{v}$. As a result, the vertex $\mathbf{u}$ also has $d + 1$ nearest lattice points, and the corresponding d-simplices are congruent for all $\mathbf{u}$. This proves the claim. □

Proposition 3.5 *The $(k - 1)$-simplices in $\mathcal{D}$ that are incident to the origin are in one-to-one correspondence to the $(d - k + 1)$-faces of Π and, hence, in one-to-one correspondence to the ordered k-partitions of $[d + 1]$.*

Let V denote the set of lattice points that share a Delaunay edge with the origin. The following statement shows that the point set V is in convex position, and the convex hull encloses Π with some "safety margin". The proof is a mere calculation, deriving an explicit equation for each hyperplane supporting the convex hull and applying it to all vertices of V and of Π.

Lemma 3.6 *For each d-simplex attached to the origin, the facet τ opposite to the origin lies on a hyperplane which is at least a distance $\frac{1}{\sqrt{2}(d+1)}$ to Π and all points of V are either on the hyperplane or on the same side as the origin.*

Proof Consider the d-simplex σ incident to the origin that is dual Voronoi vertex of Π with coordinates

$$v = \frac{1}{d+1}\left(\frac{d}{2}, \frac{d}{2} - 1, \ldots, \frac{d}{2} - (d-1), \frac{d}{2} - d\right).$$

The $(d-1)$-facet τ of σ opposite to the origin is spanned by lattice points of the form

$$\ell_k = \frac{1}{(d+1)} \underbrace{(k, \ldots, k,}_{d+1-k} \underbrace{k - (d+1), \ldots, k - (d+1))}_{k}, \quad 1 \le k \le d$$

(see the proof of Lemma 3.4 above). All points in V can be obtained by permuting the coordinates of ℓ_k.

We can verify at once that all these points lie on the hyperplane $-x_1 + x_{d+1} + 1 = 0$, so this plane supports τ. The origin lies on the positive side of the plane. All points in V either lie on the plane or are on the positive side as well, as one can easily check. For the vertices of Π, observe that the value $x_1 - x_{d+1}$ is minimized for the point v above, for which $x_1 - x_{d+1} + 1 = 1/(d + 1)$ is obtained. It follows that v as well as any vertex of V is at least in distance $\frac{1}{\sqrt{2}(d+1)}$ from H (the $\sqrt{2}$ comes from the length of the normal vector). This proves the claim for the simplex dual to v.

Any other choice of σ is dual to a permuted version of v. Let π denote the permutation on v that yields the dual vertex. The vertices of τ are obtained by applying

the same permutation on the points ℓ_k from above. Consequently, the plane equation changes to $-x_{\pi(1)} + x_{\pi(d+1)} + 1 = 0$. The same reasoning as above applies, proving the statement in general. □

Lemma 3.7 *If two lattice points are not adjacent in $\mathcal{D}$, then the corresponding Voronoi polytopes have a distance of at least $\frac{\sqrt{2}}{d+1}$.*

Proof Lemma 3.6 shows that Π is contained in a convex polytope C and the distance of Π to the boundary of C is at least $\frac{1}{\sqrt{2}(d+1)}$. Moreover, if Π' is the Voronoi polytope of a non-adjacent lattice point o', the corresponding polytope C' is interior-disjoint from C. To see that, note that the simplices in $\mathcal{D}$ incident to the origin triangulate the interior of C, and likewise for o' any interior intersection would be covered by a simplex incident to o and one incident to o', and since they are not connected, the simplices are distinct, contradicting the fact that $\mathcal{D}$ is a triangulation. Having established that C and C' are interior-disjoint, the distance between Π and Π' is at least $\frac{2}{\sqrt{2}(d+1)}$, as required. □

Recall the definition of a flag complex as the maximal simplicial complex one can form from a given graph. We next show that $\mathcal{D}$ is of this form. While our proof exploits certain properties of A^*, we could not exclude the possibility that the Delaunay triangulation of any lattice is a flag complex.

Lemma 3.8 *$\mathcal{D}$ is a flag complex.*

Proof The proof is based on two claims: consider two facets f_1 and f_2 of Π that are disjoint, that is, do not share a vertex. In the tessellation, there are permutahedra Π_1 and Π_2 that are adjacent to Π, such that $\Pi \cap \Pi_1 = f_1$ and $\Pi \cap \Pi_2 = f_2$. The first claim is that Π_1 and Π_2 are disjoint. We prove this explicitly by constructing a hyperplane separating Π_1 and Π_2. See the Appendix for further details.

The second claim is that if k facets of Π are pairwise intersecting, they also have a common intersection. Another way to phrase this statement is that the link of any vertex in $\mathcal{D}$ is a flag complex. This is a direct consequence of Lemma 3.1. See the Appendix for more details.

The lemma follows directly with these two claims: consider $k + 1$ vertices of $\mathcal{D}$ which pairwise intersect. We can assume that one point is the origin, and the other k points are the centers of permutahedra that intersect Π in a facet. By the contrapositive of the first claim, all these facets have to intersect pairwisely, because all vertices have pairwise Delaunay edges. By the second claim, there is some common vertex of Π to all these facets, and the dual Delaunay simplex contains the k-simplex spanned by the vertices. □

Lemma 3.9 *The shortest lattice vector of the A_d^* lattice has length $\sqrt{\frac{d}{d+1}}$.*

Proof The lattice vectors of the A_d^* lattice are permutations of the vectors

$$v_t = \frac{1}{d+1} \underbrace{(t, \ldots, t}_{d+1-t}, \underbrace{t - (d+1), \ldots, t - (d+1))}_{t}$$

for $1 \leq t \leq d$ [11]. The lengths of the vectors are of the form

$$|v_t| = \frac{1}{d+1}\sqrt{t^2(d+t-1) + (d+1-t)^2 t} = \sqrt{\frac{t(d+1-t)}{d+1}}.$$

This length is minimum for $t = 1$ and $t = d$, so $|v_1| = |v_d| = \sqrt{\frac{d}{d+1}}$ is the shortest length of any lattice vector. □

For any $\beta > 0$, by scaling the lattice vectors of the A_d^* lattice by β, we get a scaled A_d^* lattice. The Voronoi cells of this scaled lattice are scaled permutahedra. For scaled permutahedra we show an additional property:

Lemma 3.10 *Let π and π' denote the permutahedral cells at the origin at scales β and β', respectively where $0 < \beta' < \beta$. Then,*

- *$\pi' \subset \pi$, and*
- *the minimum distance between any facet of π' and any facet of π is at least $\frac{(\beta-\beta')}{2}\sqrt{\frac{d}{d+1}}$.*

In particular, this implies that the Minkowski sum of π' with a ball of radius $\frac{(\beta-\beta')}{2}\sqrt{\frac{d}{d+1}}$ (with the ball's center being the reference point) lies within π.

Proof The first claim, $\pi' \subset \pi$, follows since both permutahedra are scalings of a convex object centered at the origin.

For the second claim, consider any lattice vector v of the standard A_d^* lattice. The corresponding vectors at scales β and β' are $v\beta$ and $v\beta'$, respectively. Let f and f' be facets of π and π', corresponding to $v\beta$ and $v\beta'$, respectively. Then f and f' lie in parallel hyperplanes, which are separated by distance $|(v\beta - v\beta')/2| = |v|(\beta - \beta')/2$. From Lemma 3.9, we know that the shortest lattice vector has length $\sqrt{\frac{d}{d+1}}$ for the standard A_d^* lattice. This quantity scales linearly for any scaling of the lattice. This means that the minimal distance between facets of the form f, f' is $\delta := \frac{(\beta-\beta')}{2}\sqrt{\frac{d}{d+1}}$. Let f' be a facet of π' and g be a facet of π. Then there is a facet g' of π' which is a scaled version of g. Let H be the supporting hyperplane of g'. Since π' is convex, f' lies in the half-space of H (on H if $f' = g'$) containing the origin. On the other hand, g lies in other half-space. Moreover, g is at a distance at least δ from g'. Therefore, f' is separated from g by distance at least δ. This is true for any choice of f' or g, so the second claim follows. □

4 Approximation Scheme

Given a point set P of n points in $\mathbb{R}^d$, we describe our approximation complex X_β for a fixed scale $\beta > 0$. For that, let L_β denote the A_d^* lattice in $\mathbb{R}^d$, with each lattice vector scaled by β. Recall that the Voronoi cells of the lattice points are scaled permutahedra which tile $\mathbb{R}^d$. The bounds for the diameter (Lemma 3.3) as well as for the distance between non-intersecting Voronoi polytopes (Lemma 3.7) remain valid
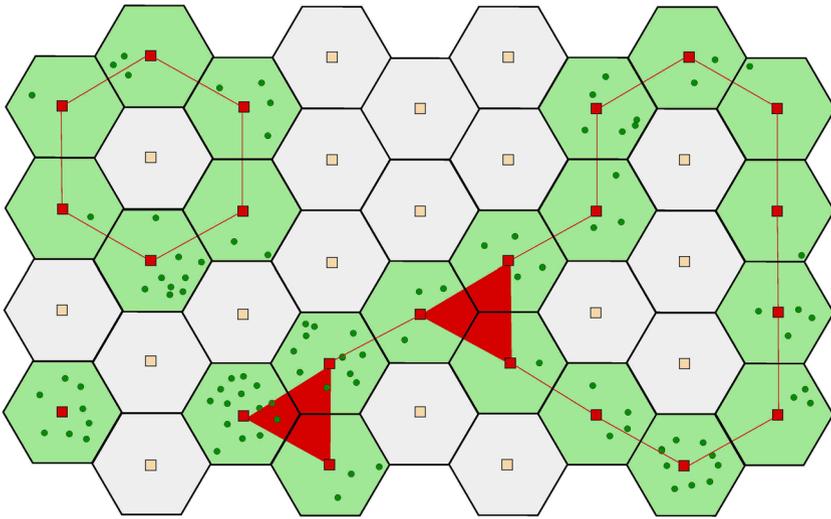
Fig. 1 An example of X_β: the darkly shaded hexagons are the *full* permutahedra, which contain input points marked as dark disks. Each dark square corresponds to a full permutohedron and represents a vertex of X_β. If two full permutahedra are adjacent, there is an edge between the corresponding vertices. The clique completion on the edge graph constitutes the complex X_β

when multiplying them with the scale factor. Hence, any cell of L_β has diameter at most $\beta\sqrt{d}$. Moreover any two non-adjacent cells have a distance at least $\beta\frac{\sqrt{2}}{d+1}$.

We call a permutahedron *full*, if it contains a point of P, and *empty* otherwise (we assume for simplicity that each point in P lies in the interior of some permutahedron; this can be ensured with well-known methods [14]). Clearly, there are at most n full permutahedra for a given P. We define X_β as the nerve of the full permutahedra defined by L_β. An equivalent formulation is that X_β is the subcomplex of $\mathcal{D}$ defined in Sect. 3 induced by the lattice points of full permutahedra. This implies that X_β is also a flag complex. We usually identify the permutahedron and its center in L_β and interpret the vertices of X_β as a subset of L_β. See Fig. 1 for an example in 2D.

4.1 Interleaving

To prove that X_β approximates the Rips filtration, we define simplicial maps connecting the complexes on related scales.

Let V_β denote the subset of L_β corresponding to full permutohedra. To construct X_β, we use a map $v_\beta \colon P \to V_\beta$, which maps each point $p \in P$ to its closest lattice point. Vice versa, we define $w_\beta \colon V_\beta \to P$ to map a vertex in V_β to the closest point of P. Note that $v_\beta \circ w_\beta$ is the identity map, while $w_\beta \circ v_\beta$ is not.

Lemma 4.1 *The map v_β induces a simplicial map $\phi_\beta \colon \mathcal{R}_{\frac{\beta}{\sqrt{2}(d+1)}} \to X_\beta$.*

Proof Because X_β is a flag complex, it is enough to show that for any edge (p, q) in $\mathcal{R}_{\frac{\beta}{\sqrt{2(d+1)}}}$, $(v_\beta(p), v_\beta(q))$ is an edge of X_β. This follows at once from the contrapositive of Lemma 3.7. □

Lemma 4.2 *The map w_β induces a simplicial map $\psi_\beta \colon X_\beta \to \mathcal{R}_{\beta 2\sqrt{d}}$.*

Proof It is enough to show that for any edge (p, q) in X_β, $(w_\beta(p), w_\beta(q))$ is an edge of $\mathcal{R}_{\beta 2\sqrt{d}}$. Note that $w_\beta(p)$ lies in the permutahedron of p and similarly, $w_\beta(q)$ lies in the permutahedron of q, so their distance is bounded by twice the diameter of the permutahedron. The statement follows from Lemma 3.3. □

Since $\beta 2\sqrt{d} < \beta 2(d + 1)$, we can compose the map ψ_β from the previous lemma with an inclusion map to a simplicial map $X_\beta \to \mathcal{R}_{\beta 2(d+1)}$ which we denote by ψ_β as well. Composing the simplicial maps ψ and ϕ, we obtain simplicial maps

$$\theta_\beta \colon X_\beta \to X_{\beta(2(d+1))^2}$$

for any β, giving rise to a discrete tower

$$\left(X_{\beta(2(d+1))^{2k}}\right)_{k\in\mathbb{Z}}.$$

The maps define the following diagram of complexes and simplicial maps between them (we omit the indices in the maps for readability):

$$
\begin{array}{ccccccc}
\cdots & \longrightarrow & \mathcal{R}_{\beta 2(d+1)} & \xrightarrow{\quad g \quad} & \mathcal{R}_{\beta 8(d+1)^3} & \longrightarrow & \cdots \\
 & {\scriptstyle \psi}\nearrow & {\scriptstyle \phi}\searrow & & {\scriptstyle \psi}\nearrow & & \\
\cdots & \longrightarrow & X_\beta & \xrightarrow{\quad \theta \quad} & X_{\beta 4(d+1)^2} & \longrightarrow & \cdots
\end{array}
\tag{1}
$$

Here, g is the inclusion map of the corresponding Rips complexes. Applying the homology functor yields a sequence of vector spaces and linear maps between them.

Lemma 4.3 *Diagram (1) commutes on the homology level, that is, $\theta_* = \phi_* \circ \psi_*$ and $g_* = \psi_* \circ \phi_*$, where the asterisk denotes the homology map induced by the simplicial map.*

Proof For the first statement, since θ is defined as $\phi \circ \psi$, so the maps commute at the simplicial level. The second identity is not true on a simplicial level; we show that the maps g and $h := \psi \circ \phi$ are *contiguous*, that means, for every simplex $(x_0, \ldots, x_k) \in \mathcal{R}_{\beta 2(d+1)}$, the simplex $(g(x_0), \ldots, g(x_k), h(x_0), \ldots, h(x_k))$ forms a simplex in $\mathcal{R}_{\beta 8(d+1)^3}$. Contiguity implies that the induced homology maps g_* and $h_* = \psi_* \circ \phi_*$ are equal [26, Sect. 12].

For the second statement, it suffices to prove that any pair of vertices among $\{g(x_0), \ldots, g(x_k), h(x_0), \ldots, h(x_k)\}$ is at most $\beta 16(d + 1)^3$ apart. This is immediately clear for any pair $(g(x_i), g(x_j))$ and $(h(x_i), h(x_j))$, so we can restrict to pairs

of the form $(g(x_i), h(x_j))$. Note that $g(x_i) = x_i$ since g is the inclusion map. Moreover, $h(x_j) = \psi(\phi(x_j))$, and $\ell := \phi(x_j)$ is the closest lattice point to x_j in $X_{\beta 4(d+1)^2}$. Since $\psi(\ell)$ is the closest point in P to ℓ, it follows that $\|x_j - h(x_j)\| \leq 2\|x_j - \ell\|$. With Lemma 3.3, we know that $\|x_j - \ell\| \leq \beta 4(d+1)^2\sqrt{d}$, which is the diameter of the permutahedron cell. Using triangle inequality, we obtain

$$\|g(x_i) - h(x_j)\| \leq \|x_i - x_j\| + \|x_j - h(x_j)\|$$
$$\leq \beta 4(d+1) + \beta 8(d+1)^2\sqrt{d} < \beta 16(d+1)^3. \qquad \square$$

Theorem 4.4 *The persistence module* $\left(H_*(X_{\beta(2(d+1))^{2k}})\right)_{k\in\mathbb{Z}}$ *approximates the persistence module* $(H_*(\mathcal{R}_\beta))_{\beta\geq 0}$ *by a factor of* $6(d+1)$.

Proof Lemma 4.3 proves that on the logarithmic scale, the two towers are *weakly* ε-*interleaved* with $\varepsilon = 2(d+1)$, in the sense of [9]. Theorem 4.3 of [9] asserts that the bottleneck distance of the towers is at most 3ε. $\qquad \square$

Remark 4.5 The simplicial maps ϕ and ψ are still valid when the lattices L_β are subject to arbitrary unitary transformations independently at each scale. Consequently, Lemma 4.3 and Theorem 4.4 remain valid in such settings. For instance, a random translation can be applied to L_β without altering its approximation qualities.

With a minor modification in the construction, we can improve the approximation factor by $O(d^{1/4})$. The main observation is that the simplicial maps $\phi: \mathcal{R}_{\frac{\beta}{\sqrt{2}(d+1)}} \to X_\beta$ and $\psi: X_\beta \to \mathcal{R}_{\beta 2\sqrt{d}}$ do not increase the scale parameters of the Rips and the approximate complexes by the same amount: ϕ increases the scale by a factor of $\sqrt{2}(d+1)$ while ψ increases it by $2\sqrt{d}$. We balance this jump in scales by redefining the approximation complex.

Let c be the constant

$$c := \left(\sqrt{2}(d+1) \cdot 2\sqrt{d}\right)^{1/2} = 2^{3/4} d^{1/4} (d+1)^{1/2}.$$

We define a new approximation complex,

$$X'_\beta := X_{\beta \frac{(d+1)^{1/2}}{2^{1/4} d^{1/4}}}, \quad \forall \beta > 0,$$

that is a scaled version of X_β. It is straightforward to verify that the simplicial maps get the form

$$\phi: \mathcal{R}_\beta \to X'_{\beta c} \quad \text{and} \quad \psi: X'_\beta \to \mathcal{R}_{\beta c},$$

and their composition gives rise to a tower $(X_{\beta c^{2k}})_{k\in\mathbb{Z}}$. Aggregating the maps, we get a diagram

$$\cdots \longrightarrow \mathcal{R}_{\beta c} \xrightarrow{\ \ g\ \ } \mathcal{R}_{\beta c^3} \longrightarrow \cdots$$

$$\cdots \longrightarrow X'_{\beta} \xrightarrow{\ \ \theta\ \ } X'_{\beta c^2} \longrightarrow \cdots$$
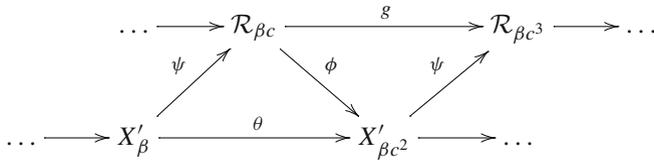
which commutes on the homology level. The proof is very similar to the proof of Lemma 4.3 and is hence omitted. As a result, $(X_{\beta c^{2k}})_{k \in \mathbb{Z}}$ is weakly c-interleaved with the Rips filtration, and both are $3c$-approximations of each other. Since $3c = O(d^{3/4})$, this improves upon the approximation factor of Theorem 4.4 by a factor of $O(d^{1/4})$.

For simplicity, we consider the original definition of the approximation complexes in the rest of the paper.

5 Computational Aspects

We utilize the non-degenerate configuration of the permutahedral tessellation to prove that X_{β} is not too large. We let $X_{\beta}^{(k)}$ denote the k-skeleton of X_{β}. In the rest of the section, we make no distinction between a vertex of X_{β} and the corresponding permutahedron, when it is clear from the context.

Theorem 5.1 *For any scale β, each vertex of X_{β} has at most $2^{O(d \log k)}$ incident k-simplices. This means that $X_{\beta}^{(k)}$ has at most $n2^{O(d \log k)}$ simplices.*

Proof We fix k and a vertex v of V_{β}. Recall that v represents a permutahedron, which we also denote by $\Pi(v)$. By definition, any k-simplex containing v corresponds to an intersection of $k + 1$ permutahedra, involving $\Pi(v)$. By Proposition 3.5, such an intersection corresponds to a $(d - k)$-face of $\Pi(v)$. Therefore, the number of k-simplices involving v is bounded by the number of $(d-k)$-faces of the permutahedron, which is $2^{O(d \log k)}$ using Lemma 3.2. The bound follows because X_{β} has at most n vertices. □

5.1 Range of Scales

Let $\mathrm{CP}(P)$ denote the closest-pair distance of P and $\mathrm{diam}(P)$ the diameter of P. The *spread* of the point set P is defined as $\Delta = \frac{\mathrm{diam}(P)}{\mathrm{CP}(P)}$. At the scale $\beta_0 = \frac{\mathrm{CP}(P)}{3d}$ and lower, no two points of P lie in adjacent cells. Therefore, the complex at such scales consists of n isolated vertices. At scale $\beta_m := \mathrm{diam}(P)(d + 1)$ and higher, all points of P lie in a collection of adjacent cells, and the nerve of these cells is a contractible simplicial complex. As a result, the persistence barcodes for scales lower than β_0 and greater than β_m are known explicitly. We restrict our attention only to the range of scales $[\beta_0, \beta_m]$ to construct the tower. In our discrete tower, the scales jump by a factor of $c := (2(d + 1))^2$ from one scale to the next. The total number of scales to

be inspected is at most

$$\lceil \log_c \beta_m / \beta_0 \rceil = \left\lceil \log_c \frac{\operatorname{diam}(P)}{\operatorname{CP}(P)} 3d(d+1) \right\rceil = \lceil \log_c \Delta + \log_c 3d(d+1) \rceil$$
$$\leq \lceil \log_c \Delta + 1 \rceil = O(\log \Delta).$$

In Theorem 5.1, we showed that the size of the k-skeleton at each scale is upper bounded by $n2^{O(d \log k)}$. A simple upper bound on the size of the k-skeleton of the tower is $n2^{O(d \log k)} \log \Delta$. The spread of a point set can be arbitrarily large, independent of the number of points or the ambient dimension.

To mitigate this undesirable dependence, we introduce a slight modification in the construction: at each scale β of the tower, we apply a random translation to the A_d^* lattice. More specifically, let π be the permutahedron at the origin at scale β. We translate the origin uniformly at random inside π, so that the lattice and the cells translate by the same amount. With this randomization, we show that the expected size of the tower is independent of the spread. Specifically, we use random translations to bounding the expected number of vertex inclusions in the tower, which then leads to the main result. The expectation is taken over the random translation of the origin, and does not depend on the choice of the input. We emphasize that the selection of the origin is the only randomized part of our construction.

5.2 Well-Separated Pair Decomposition

Given a set of n points P in $\mathbb{R}^d$, a well-separated pair decomposition (WSPD) [8] of P is a collection of pairs of subsets of P, such that for each pair, the diameter (denoted as $\operatorname{diam}()$) of the subsets is much smaller than the distance between the subsets. More formally, given a parameter $\varepsilon > 0$, an ε-WSPD consists of pairs of the form $(A_i, B_i) \subset P$ such that $\max(\operatorname{diam}(A_i), \operatorname{diam}(B_i)) \leq \varepsilon d(A_i, B_i)$ where $d(A, B)$ is the minimum separation between points of A_i and points of B_i. Additionally, for each pair of points $p, q \in P$, there exists a pair (A_j, B_j) such that either $(p \in A_j, q \in B_j)$ or $(p \in B_j, q \in A_j)$. In other words, a WSPD covers each pair of points of P. An ε-WSPD of size at most $n(1/\varepsilon)^{O(d)}$ can be computed in time $n \log n 2^{O(d)} + n(1/\varepsilon)^{O(d)}$ (see, for instance [8,16,30]).

Let W be an ε-WSPD on P with $\varepsilon = \frac{1}{6d^2}$. For each pair $(A, B) \in W$, let $P_A \subset P$ denote the set of points of A and $P_B \subset P$ denote the set of points of B. We select a representative point for A, which we call $\operatorname{rep}(A)$, by taking an arbitrary point $\operatorname{rep}(A) \in P_A$. Similarly, we select a representative $\operatorname{rep}(B) \in P_B$ for B. For the pair (A, B), we denote the distance between the representatives by $\hat{d}(A, B) := \|\operatorname{rep}(A) - \operatorname{rep}(B)\|$. We have that $d(A, B) \leq \hat{d}(A, B) \leq d(A, B) + \operatorname{diam}(A) + \operatorname{diam}(B)$, which can be simplified to $d(A, B) \leq \hat{d}(A, B) \leq d(A, B)(1 + 2\varepsilon)$ or $\frac{\hat{d}(A,B)}{1+2\varepsilon} \leq d(A, B) \leq \hat{d}(A, B)$.

5.3 Critical Scales

For any pair $(A, B) \in W$, let i be the largest integer such that $\hat{d}(A, B) > (1 + 2\varepsilon)\beta_i 2\sqrt{d}$. We say that the scales $\{\beta_{i+1}, \beta_{i+2}\}$ are *critical* for (A, B). All higher scales are *non-critical* for (A, B).

For any permutahedron π, let $\mathcal{NBR}(\pi)$ denote the union of π and its neighboring cells.

Lemma 5.2 *Let $(A, B) \in W$ be any WSPD pair. Let $(\beta < \delta)$ denote the critical scales for (A, B).*

- *At scale β, let π, π' denote the permutahedra containing $\mathrm{rep}(A), \mathrm{rep}(B)$, respectively. Then P_A lies in $\mathcal{NBR}(\pi)$. Similarly, P_B lies in $\mathcal{NBR}(\pi')$.*
- *At scale δ, let Π denote the permutahedron that contains $\mathrm{rep}(A)$. Then, $P_A \cup P_B$ lies in $\mathcal{NBR}(\Pi)$.*

Proof For the first claim, we have $\frac{\hat{d}(A,B)}{1+2\varepsilon} \leq \beta 2\sqrt{d}$ by definition, which implies

$$\mathrm{diam}(A) \leq \frac{d(A, B)}{6d^2} \leq \frac{\hat{d}(A, B)}{6d^2} \leq \frac{(1+2\varepsilon)\beta 2\sqrt{d}}{6d^2} < \frac{\beta\sqrt{2}}{d+1}.$$

Using Lemma 3.7, we get that P_A lies in $\mathcal{NBR}(\pi)$. The argument for P_B follows similarly.

For the second claim, we have

$$\begin{aligned} \mathrm{diam}(P_A \cup P_B) &\leq \mathrm{diam}(A) + d(A, B) + \mathrm{diam}(B) \leq (1 + 2\varepsilon)d(A, B) \\ &\leq (1 + 2\varepsilon)\hat{d}(A, B) \leq (1 + 2\varepsilon)^2 \beta 2\sqrt{d} \\ &\leq \frac{(1+2\varepsilon)^2 \delta 2\sqrt{d}}{c}. \end{aligned}$$

Since $\varepsilon = \frac{1}{6d^2}$ and $c = (2(d+1))^2$, we get $\mathrm{diam}(P_A \cup P_B) < \frac{\delta\sqrt{2}}{d+1}$. Using Lemma 3.7, we get that $P_A \cup P_B$ lies in $\mathcal{NBR}(\Pi)$. $\qquad\square$

Lemma 5.3 *Let $(A, B) \in W$ be any WSPD pair. Let $(\beta < \delta)$ denote the critical scales for (A, B). Consider any arbitrary pair of points $(a \in P_A, b \in P_B)$. Let $\alpha' < \alpha$ be a pair of consecutive scales such that at α', a and b lie in distinct non-adjacent permutahedra but at α, they lie in adjacent (or the same) permutahedra. Then α is a critical scale for (A, B), that is, $\alpha = \beta$ or $\alpha = \delta$.*

Proof We prove the claim by contradiction. There are two cases:

- $\alpha < \beta$: From the definition of critical scales, we have that $d(A, B) \geq \frac{\hat{d}(A,B)}{1+2\varepsilon} > \alpha(2\sqrt{d})$, that is, the minimum distance between points of P_A and P_B is more than twice the diameter of the cells at scale α. This means that for all $(a \in P_A, b \in P_B)$, the cells containing a and b are not adjacent. This contradicts our assumption that at α, there exists a pair of points $(a \in P_A, b \in P_B)$, such that they lie in adjacent (or the same) cells.

– $\alpha > \delta$: In such a case, we have $\alpha' \geq \delta$. From Lemma 5.2, we know that if rep(A) lies in cell π at scale δ or higher, then $P_A \cup P_B$ lies in $\mathcal{NBR}(\pi)$. This contradicts our assumption that at α', there exists a pair of points ($a \in P_A, b \in P_B$) which lies in distinct non-adjacent cells.

The claim follows. □

5.4 Size of the Tower

5.4.1 Splits

In the permutahedral tessellation, a cell at a given scale may not be entirely contained within a single cell at larger scales. This can lead to cases where the input points contained in a single cell map to several distinct cells at a higher scale. Formally, at a given scale β, let π be a non-empty cell and denote by $P_\pi \subset P$ the set of input points contained in π. At the next scale β', let $\{\pi_0, \pi_1, \ldots, \pi_m\}$ be the collection of cells to which P_π maps, with π mapping to π_0. We call each pair (π_0, π_i) for $1 \leq i \leq m$ a *split* at scale β'. For each split (π_0, π_i), there exists at least one pair of points $(a, b) \subset P$ such that $a, b \in \pi$, $a \in \pi_0$, $b \in \pi_i$. We call such a pair a *split inducing pair (SIP)*. Each split is induced by some SIP. Also, several SIPs may induce the same split.

Let $(A, B) \in W$ be a WSPD pair. We upper bound the number of splits induced by SIPs of the form (a, b) where $(a \in P_A, b \in P_B)$ over all scales. Counting this for each pair of W gives an upper bound on the number of splits for all SIPs, since each pair of points is covered by some pair of W.

Lemma 5.4 *Let $(A, B) \in W$ be a pair of the WSPD. The expected number of splits for SIPs of the form $(a \in P_A, b \in P_B)$ is upper bounded by $2^{O(d)}$.*

Proof First, we count the number of scales at which splits can be induced by pairs of points of the form $(a \in P_A, b \in P_B)$. From Lemma 5.3, at scales below the critical scales for (A, B), there are no splits induced by such SIPs, so we ignore those scales. There are two relevant cases:

1. *Critical scales.* Let $\beta < \delta$ be the two critical scales for (A, B). Suppose there is a split at scale β. Then there exists a SIP $(a \in P_A, b \in P_B)$ which was in a single cell at the scale immediately lower than β, but is in different cells at scale β. By Lemma 5.3, this is not possible. Therefore, there are no splits at β.

At the next critical scale δ, if rep(A) lies in cell π, then the points of $P_A \cup P_B$ lie in $\mathcal{NBR}(\pi)$, using Lemma 5.2. An upper bound on the number of full cells occupied by points $P_A \cup P_B$ at scale δ is therefore the number of cells in $\mathcal{NBR}(\pi)$, which is $2^{O(d)}$.

2. *Non-critical scales.* We denote these scales by $\mu_i = c^i \delta$, $i \geq 1$. Let π denote the permutahedron at scale μ_i that contains rep(A). Using the proof of Lemma 5.2, it holds that

$$\text{diam}(P_A \cup P_B) \leq \frac{(1+2\varepsilon)^2 \delta 2\sqrt{d}}{c} \leq \frac{(1+2\varepsilon)^2 \mu_i 2\sqrt{d}}{c^{i+1}} < \frac{\mu_i}{d^i}.$$

Points of $P_A \cup P_B$ lie in $\mathcal{NBR}(\pi)$. An upper bound on the number of cells occupied by $P_A \cup P_B$ is $2^{O(d)}$. We show that with a high probability, points of $P_A \cup P_B$ lie in π, so that it is unlikely that they occupy many cells. We give an upper bound for the expected number of scales where $P_A \cup P_B$ does not lie in π.

If rep(A) has distance greater than diam$(P_A \cup P_B)$ from all facets of π, then all points of $P_A \cup P_B$ lie in π. Without loss of generality, assume that π is centered at the origin. Set $x := \mu_i - 3\text{diam}(P_A \cup P_B)$ and let π' denote the permutahedron at the origin at scale x. From Lemma 3.10, the Minkowski sum of π' with a ball of radius $\frac{\mu_i - x}{2}\sqrt{\frac{d}{d+1}}$ lies inside π. We see that

$$\frac{\mu_i - x}{2}\sqrt{\frac{d}{d+1}} \geq \frac{3\,\text{diam}(P_A \cup P_B)}{2}\sqrt{\frac{d}{d+1}} > \text{diam}(P_A \cup P_B).$$

Because of the random translation at each scale, the location of rep(A) inside π is uniformly distributed. Let Q_i be the probability that rep(A) lies in π', and $Q_i' = 1 - Q_i$ its complement. As before, diam$(P_A \cup P_B) < \frac{\mu_i}{d^i}$, so $\frac{\text{diam}(P_A \cup P_B)}{\mu_i} < \frac{1}{d^i}$. Using this fact,

$$Q_i = \frac{\text{Vol}(\pi')}{\text{Vol}(\pi)} = \left(\frac{x}{\mu_i}\right)^d = \left(1 - \frac{3\text{diam}(P_A \cup P_B)}{\mu_i}\right)^d \implies Q_i > \left(1 - \frac{3}{d^i}\right)^d.$$

Using Bernoulli's inequality [3], $Q_i > 1 - \frac{3d}{d^i}$, so $Q_i' < \frac{3}{d^{i-1}}$. Let T_i denote the probability that at scale μ_i, $P_A \cup P_B$ lies in π, with $T_i' = 1 - T_i$ denoting the complement. Since $T_i \geq Q_i$, we have that $T_i' \leq Q_i' < \frac{3}{d^{i-1}}$. The expected number of scales where $P_A \cup P_B$ does not lie in π, implying that splits can occur, is

$$\sum_{i=1}^{\infty} T_i' < \sum_{i=1}^{\infty} \frac{3}{d^{i-1}} < 6.$$

The total number of scales where splits can occur for (A, B) is seven in expectation, one being a critical scale and six being non-critical scales. At each such scale, a simple upper bound for the number of splits is the number of permutations of two cells from the full cells, which is $2\binom{2^{O(d)}}{2}$. This is again $2^{O(d)}$, so the claim follows. $\qquad\square$

Lemma 5.5 *The expected number of vertex inclusions in the tower is upper bounded by $n2^{O(d \log d)}$.*

Proof At scale β_0, there are n vertex inclusions in the tower due to n full permutahedra. First, we show that each vertex inclusion at higher scales is caused by a split.

Let $\alpha' < \alpha$ be any two consecutive scales in the tower, with the set of full vertices being V' and V, respectively and let θ be the simplicial map from the complex at α' to the complex at α. Let $v \in V \backslash \theta(V')$ denote a full cell. There is an input point $p \in v$, since v is full. Let u denote the full cell at scale α', which contains p. Since $\theta(u) \neq v$, there exists another input point $p' \in u$ at scale α' such that p' is the closest input point

to u's center. Then (u, v) is a split induced by the SIP (p, p'), implying that v was created from a split.

There are at most $n(6d^2)^{O(d)} = n2^{O(d \log d)}$ pairs in the WSPD, so the total number of expected splits is upper bounded by $n2^{O(d \log d)} \cdot 2^{O(d)}$, using Lemma 5.4. The claim follows. □

Theorem 5.6 *The expected size of the tower is upper bounded by $n2^{O(d \log d)}$.*

Proof Recall that the size of a tower is the number of simplex inclusions involved. From Lemma 5.5, we know that the expected number of vertex inclusions in the tower is upper bounded by $n2^{O(d \log d)}$. Each simplex included in the filtration is attached to one of these vertices. From Theorem 5.1 we know that each vertex has at most $2^{O(d \log k)}$ k-simplices attached to it. Therefore, the expected number of simplex inclusions in the tower is upper bounded by $n2^{O(d \log d)}2^{O(d \log k)} = n2^{O(d \log d)}$. □

Note that we do not explicitly construct the WSPD W to argue about the size of the tower. Existence of W suffices for our claims. We next show that any simplex that is included in the tower, collapses to a vertex very soon, within the next few scales:

Lemma 5.7 *Let σ be any k-simplex ($k \geq 1$), and δ_1 denote the scale at which it is included in the tower. Let $\delta_{i+1} = c^i\delta_1$, $i \geq 1$, denote the next scales. Let the induced simplicial complexes and simplicial maps be*

$$X_{\delta_1} \xrightarrow{\theta_1} X_{\delta_2} \xrightarrow{\theta_2} \ldots$$

Let $\theta^i = \theta_i \circ \theta_{i-1} \circ \ldots \circ \theta_1$ denote the composition of simplicial maps over i consecutive scales. Then,

- *$\theta^1(\sigma)$ is a vertex with probability greater than $1/2$.*
- *Let j denote the smallest integer such that $\theta^j(\sigma)$ is a vertex. We say that σ survives for j scales. Then, the expected value of j is at most four.*

Proof Let σ be a simplex with vertices being full cells $(\pi_0, \ldots, \pi_k)$ at scale δ_1. The diameter of this collection of cells is no more than $2\delta_1\sqrt{d}$. Let s be any point of $\mathbb{R}^d$ in π_0 and denote by Π the permutahedron at scale δ_2 that contains s. Let Pr denote the probability that $\theta^1(\sigma) = \Pi$. If s lies at distance at least $2\delta_1\sqrt{d}$ from the facets of Π, then $\pi_0 \cup \cdots \cup \pi_k$ lies inside Π, guaranteeing that $\theta^1(\sigma) = \Pi$. Set $x := \delta_2 - 6\delta_1\sqrt{d}$ and denote by Π' the permutahedron centered at the origin at scale x. From Lemma 3.10, the Minkowski sum of Π' with a ball of radius $\frac{\delta_2 - x}{2}\sqrt{\frac{d}{d+1}}$ lies inside Π. We see that $\frac{\delta_2 - x}{2}\sqrt{\frac{d}{d+1}} > 2\delta_1\sqrt{d}$, so if s lies in Π', then it is further than $2\delta_1\sqrt{d}$ from the facets of Π. Since the origin is randomly translated at each scale, the position of s is uniformly distributed in Π. Let Qr denote the probability that s lies in Π'. Then,

$$Qr = \frac{\text{Vol}(\Pi')}{\text{Vol}(\Pi)} = \left(\frac{x}{\delta_2}\right)^d = \left(1 - \frac{6\delta_1\sqrt{d}}{\delta_2}\right)^d$$
$$= \left(1 - \frac{6\sqrt{d}}{c}\right)^d = \left(1 - \frac{3\sqrt{d}}{2(d+1)^2}\right)^d.$$

Using Bernoulli's inequality [3], $Qr > 1 - \frac{3d\sqrt{d}}{2(d+1)^2} > 1/2$. Since $Pr \geq Qr > 1/2$, the first claim follows.

If $\theta^i(\sigma)$ is a vertex, it remains so for all higher scales. Because the origin is chosen uniformly at random at each scale, and the ratio of any two consecutive scales is a constant, the probability that $\theta^i(\sigma)$ is a vertex given that $\theta^{i-1}(\sigma)$ was not, also has the value Pr. Let Pr' denote the complement of Pr. Then, the probability that σ survives for j scales is $(Pr')^{j-1}Pr$. Since $Pr > 1/2$, $(Pr')^j < 1/2^j$. The expected number of scales for which σ survives is

$$\sum_{j=1}^{\infty} j(Pr')^{j-1}Pr < \sum_{j=1}^{\infty} j(Pr')^{j-1} < \sum_{j=1}^{\infty} j/2^{j-1} = 4.$$

The claim follows. □

5.5 Computing the Tower

5.5.1 Determining the Range of Scales

If the range of scales $[\beta_0, \beta_m]$ is provided as an input, we build the tower at each of the relevant scales. If the range is not provided, we calculate the spread of the point set to determine the relevant scales. For our purpose, it suffices to calculate constant-factor approximations of $\text{diam}(P)$ and $\text{CP}(P)$. Taking an arbitrary point $p \in P$ and calculating $\max_{q \in P} \|p - q\|$ gives a $1/2$-approximation of $\text{diam}(P)$. $\text{CP}(P)$ can be computed exactly using a randomized algorithm in $n2^{O(d)}$ expected time [23]. Using this information, we calculate the range of scales and call them $[\beta_0, \beta_m]$. The scales of the tower can then be written in the form $\beta_i = \beta_0 c^i$, $i \geq 0$, where $c = (2(d+1))^2$.

Algorithm 5.8 We construct the tower scale-by-scale. Let $\alpha' < \alpha$ be any two consecutive scales and X', X the respective complexes, with $\theta : X' \to X$ being the induced simplicial map. Suppose we have already constructed X'. There are two steps in constructing the complex X:

Adding vertices and edges. We translate the lattice by picking a point uniformly at random from the cell at the origin, which can be done using random walks in polytopes [24]. We compute the set of full permutahedra by finding the closest lattice point for each point in P [11, Alg. 4, Chap. 20]. Then, for each full cell π, we go over $\mathcal{NBR}(\pi)\backslash\pi$ to find neighboring full cells. If a full neighbor is found, we add an edge between π and its neighbor. This completes the 1-skeleton X.

Adding simplices. Each simplex in X is one of two kinds:

- Those which are images of simplices of X' under θ: To construct these, we first construct θ for vertices of X. Then we go over each simplex $\sigma = (v_0, \ldots, v_k) \in X'$ and add the simplex $\theta(\sigma)$ on the vertices $(\theta(v_0), \ldots, \theta(v_k))$ of X.
- Those which are not in the image of θ, that is, those simplices which are included in the tower at scale α: Each such simplex σ must contain at least one edge which is not in the image of θ, since otherwise all edges of σ and hence σ itself would be

under the image of θ. We first enumerate all the edges which are not in the image of θ. To do this, for each edge $(u, v) \in X'$, we calculate $(\theta(u), \theta(v))$ and exclude them from the list of edges of X, to get the list of new edges.

To complete the k-skeleton, we go over the new edges of the complex in an arbitrary order, and at each step we add the new simplices induced by the current edge. Let (u, v) be the edge under consideration. We construct the simplices incident to (u, v) inductively by dimension. The base case is the 1-skeleton, with simplex (u, v). Assume that we have completed the $(j - 1)$-simplices incident to (u, v). Let σ be a j-simplex incident to (u, v). Then, σ is of the form $\sigma = w * \gamma$, where γ is a $(j - 1)$-simplex incident to (u, v) and w is a full cell which is a common neighbor of u and v. To find σ, we go over each $2^{O(d)}$ common neighbors of u and v and each $(j - 1)$-simplex γ containing (u, v), and test whether $w * \gamma$ is a j-simplex in the complex. The test works by checking whether each $w * \gamma_i$ is a $(j - 1)$-simplex in the complex, where γ_i is a facet of γ. Since we enumerate all the simplices attached to each new edge, this step generates all simplices included in the tower at scale α.

Theorem 5.9 *Algorithm* 5.8 *takes* $n2^{O(d)} \log \Delta + M2^{O(d)}$ *time in expectation and* M *space to compute the k-skeleton, where M is the size of the tower. Additionally, the expected runtime is upper bounded by* $n2^{O(d)} \log \Delta + n2^{O(d \log d)}$ *and the expected space is upper bounded by* $n2^{O(d \log d)}$.

Proof At each scale, picking the origin takes $\text{poly}(d)$ time [24]. Finding the closest lattice vertex for any given input point takes $O(d^2)$ time [11, Alg. 4, Chap. 20]. Therefore, finding the full vertices at each scale takes $O(nd^2)$ time per scale, and in total $O(nd^2 \log \Delta)$ time. Each cell has $2^{O(d)}$ neighbors, so finding the full neighbors and adding the edges takes $n2^{O(d)}$ per scale. Computing the map θ for the vertices of X_1 takes time $O(nd^2)$ per scale. In total, these steps take $n2^{O(d)} \log \Delta$ time.

For each simplex of X', we compute the image under θ. This takes time $O(d)$ per simplex of X', since the vertex map has already been established. From Lemma 5.7, each simplex in the tower survives at most four scales (in expectation) under θ, until it collapses to a vertex. Therefore, for each simplex in the tower, we compute its related images four times in expectation. This step takes $4M O(d)$ time over the tower, in expectation.

Computing θ for the edges of X' takes time $O(1)$ time per edge, since we already computed the vertex map. Finding new edges takes $n2^{O(d)}$ time, since that is the maximum number of edges at any scale. In total, finding new edges takes $n2^{O(d)} \log \Delta$ time. To complete the k-skeleton, the testing technique requires an overhead of $k^2 2^{O(d)} = 2^{O(d)}$ for each simplex in the tower. Since we do the k-completion only for newly added edges, the test is not repeated for any simplex. The time bound follows.

The space complexity follows by storing the tower. The expected size of the tower is upper bounded by $n2^{O(d \log d)}$, from Theorem 5.6. The claims follow. $\qquad\square$

It is possible to compute the persistence barcode of towers in a streaming setting [21], where instead of storing the entire tower in memory, the complex is constructed at each scale and fed to the output stream. In this setting, Algorithm 5.8 only needs to store the k-skeleton of the current scale in memory, to complete the

k-skeleton of the next scale. So, the maximum memory consumption of Algorithm 5.8 comes down to $n2^{O(d \log k)}$, which is the maximum size of the k-skeleton per scale, using Theorem 5.1.

In Algorithm 5.8, to construct the edges of the complex at each scale, we scan the neighborhood of each full cell. By adding the edges in more careful method, we reduce the complexity of this step. Let W denote a $\frac{1}{6d^2}$-WSPD on P. Let $\alpha' < \alpha$ be any two consecutive scales of the tower, with X', X being the complexes at the respective scales. Let $\theta: X' \to X$ be the induced simplicial map. For any permutahedron π, let $\mathcal{NBR}(\mathcal{NBR}(\pi))$ denote the union of the collections of cells $\mathcal{NBR}(\pi_i)$ for each cell $\pi_i \in \mathcal{NBR}(\pi)$.

Lemma 5.10 *Let $\pi_1 \neq \pi_2$ be any pair of full cells at α such that (π_1, π_2) is an edge in X. There are three possibilities:*

- *There exist adjacent full cells $u, v \in X'$ such that $\theta(u, v) = (\pi_1, \pi_2)$, that is, (π_1, π_2) is the image of an edge from the previous scale. In such a case, we call (π_1, π_2) an* inherited *edge.*
- *There exist full cells $u, v \in X'$ such that $\theta(u) = \pi_1$ and $\theta(v) = \pi_2$, but (u, v) is not an edge in X'. We call (π_1, π_2) an* interactive *edge.*
- *At least one of $\{\pi_1, \pi_2\}$ have no pre-image in X' under θ, that is, there do not exist $u, v \in X'$, such that $\theta(u) = \pi_1$ and $\theta(v) = \pi_2$ both hold. In such a case we call (π_1, π_2) a* split *edge.*

Let (π_1, π_2) be an interactive edge. Then,

- *There exists a pair $(A, B) \in W$ such that α is a critical scale for (A, B).*
- *Let π_3 be the permutahedron containing rep(A) at α. Then, π_1 and π_2 are cells in $\mathcal{NBR}(\mathcal{NBR}(\pi_3))$.*

Proof Since the three edge classes are exhaustive, each edge of X is either an inherited edge, or an interactive edge, or a split edge.

For the second part of the claim,

- Let u, v be distinct full cells at α' such that $\theta(u) = \pi_1$ and $\theta(v) = \pi_2$. Since u and v are full cells, there exist points $p_1, p_2 \in P$ such that $p_1 \in u, p_2 \in v$, and p_1 and p_2 are the closest points to centers of u and v, respectively. At α, $p_1 \in \pi_1$ and $p_2 \in \pi_2$. Let (A, B) be a WSPD pair which covers (p_1, p_2), that is, $p_1 \in P_A, p_2 \in P_B$. Using Lemma 5.3, α is a critical scale for (A, B).
- Using Lemma 5.2, points of P_A lie in $\mathcal{NBR}(\pi_3)$, so $\pi_1 \in \mathcal{NBR}(\pi_3)$. Since $\pi_2 \in \mathcal{NBR}(\pi_1)$, the claim follows. $\qquad\square$

Algorithm 5.11 There are two stages in the algorithm.

Stage 1. We compute a $\frac{1}{6d^2}$-WSPD W on P. For each WSPD pair $(A, B) \in W$, the two critical scales are determined using $\hat{d}(A, B)$. For each scale, we store the WSPD pairs for which the scale is critical.

Stage 2. We construct the complex scale by scale. For this, let $\alpha' < \alpha$ be any two consecutive scales. Suppose we have constructed the complex X' at α'. We choose the origin at α as in Algorithm 5.8. To construct the complex X at α, we start by finding

the full vertices by mapping points of P to their closest lattice point. Then we calculate the vertex map from X' to X which induces the simplicial map $\theta \colon X' \to X$.

The simplices in X are of two kinds: those which are images of θ and those which are not. For simplices of the former kind, we use the vertex map to compute the image under θ, and add it to X. For the latter case, each simplex must contain a new edge, since otherwise the simplex was already in the image of θ. To compute these new edges at α, we use Lemma 5.10: the only new edges at this scale are the interactive and split edges.

Step 1. We process all WSPD pairs which are critical at α. Let (A, B) be the current pair and let π denote the permutahedron which contains rep(A). For each cell $\pi' \in \mathcal{NBR}(\pi)$, we add edges of π' with full cells of $\mathcal{NBR}(\pi') \backslash \pi'$. This amounts to adding edges between all pairs of adjacent full cells in $\mathcal{NBR}(\mathcal{NBR}(\pi))$. By Lemma 5.10, all interactive edges are added by this procedure.

Step 2. We collect the full cells which do not have a pre-image under θ. This is done by excluding the images of the vertices of X' under θ, from the set of vertices of X. For each such full cell π, we go over $\mathcal{NBR}(\pi) \backslash \pi$ and add edges with full cells. This step enumerates all split edges (Lemma 5.10).

Step 3. Steps 1 and 2 generate the new edges of X. With this information, we enumerate the k-skeleton of X, using the technique from Algorithm 5.8.

Theorem 5.12 *Algorithm* 5.11 *takes*

$$\left(O(nd^2) + \mathrm{poly}(d)\right) \log \Delta + n \log n 2^{O(d)} + (M + |W|) 2^{O(d)}$$

time in expectation and $M + O(|W|)$ space, where M is the size of the tower and $|W|$ is the size of the WSPD. Additionally, the expected runtime is upper bounded by

$$\left(O(nd^2) + \mathrm{poly}(d)\right) \log \Delta + n \log n 2^{O(d)} + n 2^{O(d \log d)}$$

and the expected space is upper bounded by $n 2^{O(d \log d)}$.

Proof In Stage 1, we compute a $\frac{1}{6d^2}$-WSPD. This takes time $n \log n 2^{O(d)} + |W|$. For each WSPD pair we calculate two critical scales. This takes $O(1)$ time per pair, so $O(|W|)$ in total. Stage 1, therefore, takes $n \log n 2^{O(d)} + O(|W|)$ time.

In Stage 2, at each scale, we select the origin as in Theorem 5.9, which takes $\mathrm{poly}(d)$ time per scale [24]. Then, we compute the full vertices at each scale. This takes time $O(nd^2)$ per scale. Computing the vertex map which induces θ also takes $O(nd^2)$ per scale. In total, these steps take $(O(nd^2) + \mathrm{poly}(d)) \log \Delta$. Taking the image of simplices of X' takes $O(d)$ time per simplex, as the vertex map is already computed. As argued in Theorem 5.9, this step takes $4MO(d)$ time in expectation. For the remaining simplices of X,

- In Step 1, we add edges between adjacent full cells of $\mathcal{NBR}(\mathcal{NBR}(\pi))$. There are $2^{O(d)}$ such cells, so it takes $2^{O(d)}$ time per WSPD pair per critical scale. Since there are $O(|W|)$ such instances, in total this step takes $2^{O(d)}|W|$ time.

– In Step 2, we inspect the neighbors of full cells which do not have a pre-image under θ. The number of such full cells is the number of vertex inclusions in the tower, which is upper bounded by M. Per cell, this takes $2^{O(d)}$ time, so this step takes no more than $2^{O(d)}M$ time in total.

– In Step 3, the new edges are the inherited and split edges. Each such edge survives four scales in expectation, from Lemma 5.7, so the expected number of new edges in the tower is upper bounded by $4M$. This is also the time required to find the new edges. Completing the k-skeleton has an overhead of $k^2 2^{O(d)}$ per simplex in the tower as in Algorithm 5.8, so it takes $M 2^{O(d)}$ time in total.

In total, Stage 2 takes time $(O(nd^2) + \text{poly}(d)) \log \Delta + (M + |W|)2^{O(d)}$. The time bound follows.

Storing the critical scales for each WSPD pair takes $O(1)$ space per pair. Additionally, we store the tower. The space bound follows.

In the worst case, $|W| = n(6d^2)^{O(d)} = n2^{O(d \log d)}$ and M is upper-bounded by $n2^{O(d \log d)}$ in expectation. The claim follows. □

Algorithm 5.11 can be used in a streaming setting, similar to Algorithm 5.8. In this case, the memory consumption of Algorithm 5.11 is $O(|W|) + M_i$, where M_i is the size of the complex at any scale. Since $|W|$ can be as large as $n2^{O(d \log d)}$ and M_i can be at most $n2^{O(d \log k)}$ (Theorem 5.1), the space requirement is at most $n2^{O(d \log d)}$.

If the spread is a constant, then Algorithm 5.8 has a better runtime, since it does not compute the WSPD. Also, Algorithm 5.8 does not have to store the critical scales of the WSPD, neither in the normal setting nor in the streaming environment, so it is more space-efficient. However, if the spread is large, then Algorithm 5.11 achieves better runtime, since it avoids the $n2^{O(d)} \log \Delta$ factor in the complexity of Algorithm 5.8.

6 Dimension Reduction

For large d, our approximation complex plays nicely together with dimension reduction techniques. We start with noting that interleavings satisfy the triangle inequality. This result is folklore; see [7, Thm. 3.3] for a proof in a generalized context.

Lemma 6.1 *Let* (A_β), (B_β), *and* (C_β) *be persistence modules. If* (A_β) *is a* t_1-*approximation of* (B_β) *and* (B_β) *is a* t_2-*approximation of* (C_β), *then* (A_β) *is a* $(t_1 t_2)$-*approximation of* (C_β).

The following statement is a simple application of interleaving distances from [9].

Lemma 6.2 *Let* $f : P \to \mathbb{R}^m$ *be an injective map such that*

$$\xi_1 \|p - q\| \leq \|f(p) - f(q)\| \leq \xi_2 \|p - q\|$$

for some constants $\xi_1 \leq 1 \leq \xi_2$. *Let* $\overline{\mathcal{R}}_\alpha$ *denote the Rips complex of the point set* $f(P)$. *Then, the persistence module* $(H_*(\overline{\mathcal{R}}_\alpha))_{\alpha \geq 0}$ *is a* $\frac{\xi_2}{\xi_1}$-*approximation of* $(H_*(\mathcal{R}_\alpha))_{\alpha \geq 0}$.

Proof The map f is a bijection between P and $f(P)$. The properties of f ensure that the vertex maps f^{-1} and f, composed with appropriate inclusion maps, induce simplicial maps

$$\overline{\mathcal{R}}_{\frac{\alpha}{\xi_2/\xi_1}} \overset{\phi}{\hookrightarrow} \mathcal{R}_\alpha \overset{\psi}{\hookrightarrow} \overline{\mathcal{R}}_{\alpha\xi_2/\xi_1}.$$

It is straightforward to show that the following diagrams commute on a simplicial level,
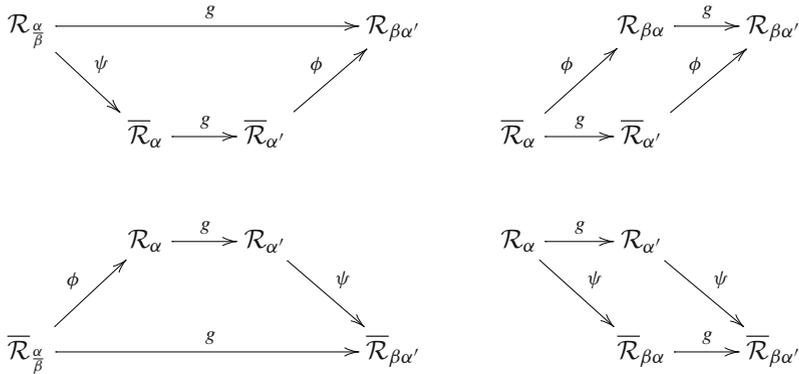
where g is the inclusion map. Hence, the strong interleaving result from [9] implies that both persistence modules are $\frac{\xi_2}{\xi_1}$-approximations of each other. □

As a first application, we show that we can shrink the approximation size from Theorem 5.6 for the case $d \gg \log n$, only worsening the approximation quality by a constant factor.

Theorem 6.3 *Let P be a set of n points in $\mathbb{R}^d$. There exists a constant c and a discrete tower of the form $(\overline{X}_{(c \log n)^{2k}})_{k \in \mathbb{Z}}$ that is $(3c \log n)$-interleaved with the Rips filtration of P and has only $n^{O(\log \log n)}$ simplices in expectation. With high success probability, we can compute the tower in deterministic expected running time $n(\log n)^2 O(\log \Delta) + n^{O(\log \log n)}$ using Algorithm 5.11.*

Proof The famous lemma of Johnson and Lindenstrauss [18] asserts the existence of a map f as in Lemma 6.2 for $m = \lambda \log n / \varepsilon^2$ with some absolute constant λ and $\xi_1 = (1 - \varepsilon)$, $\xi_2 = (1 + \varepsilon)$. Choosing $\varepsilon = 1/2$, we obtain that $m = O(\log n)$ and $\xi_2 / \xi_1 = 3$. With $\overline{\mathcal{R}}_\alpha$ the Rips complex of the Johnson–Lindenstrauss transform, we have therefore that $(H_*(\overline{\mathcal{R}}_\alpha))_{\alpha \geq 0}$ is a 3-approximation of $(H_*(\mathcal{R}_\alpha))_{\alpha \geq 0}$. Moreover, using the approximation scheme from this section, we can define a tower $(\overline{X}_\beta)_{\beta \geq 0}$ whose induced persistence module $(H_*(X_\beta))_{\beta \geq 0}$ is a $6(m + 1)$-approximation of $(H_*(\overline{\mathcal{R}}_\alpha))_{\alpha \geq 0}$, and its expected size is $n 2^{O(\log n \log \log n)} = n^{O(\log \log n)}$. The first half of the result follows using Lemma 6.1.

The Johnson–Lindenstrauss lemma further implies that an orthogonal projection to a randomly chosen subspace of dimension m will yield an f as above, with high probability. Our algorithm picks such a subspace, projects all points into this subspace (this requires $O(dn \log n)$ time) and applies the approximation scheme for the projected point set. The runtime bound follows from Theorem 5.12. □

Note that the approximation complex from the previous theorem has size $n^{O(\log \log n)}$ which is super-polynomial in n. Using a slightly more elaborate dimension reduction result by Matoušek [25], we can get a size bound polynomial in n, at the price of an additional $\log n$-factor in the approximation quality. Let us first state Matoušek's result (whose proof follows a similar strategy as for the Johnson–Lindenstrauss lemma):

Theorem 6.4 *Let P be an n-point set in $\mathbb{R}^d$. Then, a random orthogonal projection into $\mathbb{R}^k$ for $3 \leq k \leq C \log n$ distorts pairwise distances in P by at most $O(n^{2/k} \sqrt{\log n/k})$. The constants in the bound depend only on C.*

By setting $k := \frac{4 \log n}{\log \log n}$ in Matoušek's result, we see that this results in a distortion of at most $O(\sqrt{\log n \log \log n})$.

Theorem 6.5 *Let P be a set of n points in $\mathbb{R}^d$. There exists a constant c and a discrete tower of the form*

$$\left(\overline{X}_{\left(c \log n \left(\frac{\log n}{\log \log n} \right)^{1/2} \right)^{2k}} \right)_{k \in \mathbb{Z}},$$

such that it is $3c \log n \left(\frac{\log n}{\log \log n} \right)^{1/2}$-interleaved with the Rips filtration on P and has $n^{O(1)}$ simplices in expectation. Moreover, we can compute, with high success probability, the tower with this property in deterministic expected running time $n(\log n)^2 O(\log \Delta) + n^{O(1)}$ using Algorithm 5.11.

Proof The proof follows the same pattern of Theorem 6.3 with a few changes. We use Matoušek's dimension reduction result described in Theorem 6.4 with the projection dimension being $m := \frac{4 \log n}{\log \log n}$. Hence, $\xi_2/\xi_1 = O(\sqrt{\log n \log \log n})$ for the Rips construction. The final approximation factor is $6(m + 1)\xi_2/\xi_1$ which simplifies to $O(\log n \left(\frac{\log n}{\log \log n} \right)^{1/2})$. The size and runtime bounds follow by substituting the value of m in the respective bounds. □

Finally, we consider the important generalization that P is not given as an embedding in $\mathbb{R}^d$, but as a point sample from a general metric space. We use the classical result by Bourgain [6] to embed P in Euclidean space with small distortion. In the language of Lemma 6.2, Bourgain's result permits an embedding into $m = O(\log^2 n)$ dimensions with a distortion $\xi_2/\xi_1 = O(\log n)$, where the constants are independent of n. Our strategy for approximating a general metric space consists of first embedding it into $\mathbb{R}^{O(\log^2 n)}$, then reducing the dimension, and finally applying our approximation scheme on the projected embedding. The results are similar to Theorems 6.3 and 6.5,

except that the approximation quality further worsens by a factor of $\log n$ due to Bourgain's embedding. We only state the generalized version of Theorem 6.5, omitting the corresponding generalization of Theorem 6.3. The proof is straightforward with the same techniques as before.

Theorem 6.6 *Let P be a general metric space with n points. There exists a constant c and a discrete tower of the form*

$$\left(\overline{X}_{\left(c \log^2 n \left(\frac{\log n}{\log \log n} \right)^{1/2} \right)^{2k}} \right)_{k \in \mathbb{Z}}$$

that is $3c \log^2 n \left(\frac{\log n}{\log \log n} \right)^{1/2}$-interleaved with the Rips filtration on P and has $n^{O(1)}$ simplices in expectation. Moreover, we can compute, with high success probability the tower with this property in deterministic expected running time $n (\log n)^2 O (\log \Delta) + n^{O(1)}$ using Algorithm 5.11.

7 A Lower Bound for Approximation Schemes

We describe a point configuration for which the Čech filtration gives rise to a large number, say N, of features with "large" persistence, relative to the scale on which the persistence appears. Any ε-approximation of the Čech filtration, for ε small enough, has to contain at least one interval per such feature in its persistent barcode, yielding a barcode of size at least N. This constitutes a lower bound on the size of the approximation itself, at least if the approximation stems from a simplicial tower: in this case, the introduction of a new interval in the barcode requires at least one simplex to be added to the tower; also more generally, it makes sense to assume that any representation of a persistence module is at least as large as the size of the resulting persistence barcode.

To formalize what we mean by a "large" persistent feature, we call an interval (α, α') of $(H_*(C_\alpha))_{\alpha \geq 0}$ δ-*significant* for $0 < \delta < \frac{\alpha' - \alpha}{2\alpha'}$. Our approach from above translates into the following statement:

Lemma 7.1 *For $\delta > 0$, and a point set P, let N denote the number of δ-significant intervals of $(H_*(C_\alpha))_{\alpha \geq 0}$. Then, any persistence module $(X_\alpha)_{\alpha \geq 0}$ that is a $(1 + \delta)$-approximation of $(H_*(C_\alpha))_{\alpha \geq 0}$ has at least N intervals in its barcode.*

Proof If (α, α') is δ-significant, that means that there exist some $\varepsilon > 0$ and $c \in (\alpha, \alpha')$ such that $\alpha/(1 - \varepsilon) \leq c/(1 + \delta) < c(1 + \delta) \leq \alpha'$. Any persistence module that is an $(1 + \delta)$-approximation of $(H_*(C_\alpha))_{\alpha \geq 0}$ needs to represent an approximation of the interval in the range $(c(1 - \varepsilon)/2, c)$; in other words, there is an interval corresponding to (α, α') in the approximation.

We first argue that δ-significance implies the existence of $\varepsilon > 0$ and $c \in (\alpha, \alpha')$ such that $\alpha/(1-\varepsilon) \leq c/(1+\delta) < c(1+\delta) \leq \alpha'$: We choose $c := \alpha'/(1+\delta)$, so that the last inequality is satisfied. For the first inequality, we note first that $(1 - 2\delta) < 1/(1 + \delta)^2$ for all $\delta < 1/2$. By assumption, $\alpha' - \alpha > 2\alpha'\delta$, so $\alpha < \alpha'(1 - 2\delta) < \alpha'/(1 + \delta)^2 = c/(1 + \delta)$. Since the inequality is strict, we can choose some small $\varepsilon > 0$, such that $\alpha/(1 - \varepsilon) \leq c/(1 + \delta)$.

By the definition of $(1 + \delta)$-approximation, we have a commutative diagram

$$
\begin{array}{ccc}
H(\mathcal{C}_{c(1-\varepsilon)/(1+\delta)}) & \xrightarrow{\quad g \quad} & H(\mathcal{C}_{c(1+\delta)}) \\
\phi \searrow & & \nearrow \psi \\
& X_{c(1-\varepsilon)/2} \xrightarrow{\; h \;} X_c &
\end{array}
$$

Let γ be the element in the upper-left vector space, corresponding to the δ-significant interval. By definition, $g(\gamma) \neq 0$. It follows that $h(\phi(\gamma)) \neq 0$ either, so there is a corresponding interval in the approximation. $\qquad\square$

7.1 Setup

We next define our point set for a fixed dimension d. Consider the A^* lattice with origin o. Recall that o has $2^{d+1} - 2$ neighbors in the Delaunay triangulation $\mathcal{D}$ of A_d^*, because its dual Voronoi polytope, the permutahedron Π, has that many facets. We define P as the union of o with all its Delaunay neighbors, yielding a point set of cardinality $2^{d+1} - 1$. As usual, we set $n := |P|$, so that $d = \Theta(\log n)$.

We write $\mathcal{D}_P$ for the Delaunay triangulation of P. Since P contains o and all its neighbors, the Delaunay simplices of $\mathcal{D}_P$ incident to o are the same as the Delaunay simplices of $\mathcal{D}$ incident to o. Thus, according to Proposition 3.5, a $(k-1)$-simplex of $\mathcal{D}_P$ incident to o corresponds to a $(d - k + 1)$-face of Π and thus to an ordered k-partition of $[d + 1]$.

Fix an integer parameter $\ell \geq 3$, to be defined later. We call an ordered k-partition $(S_1, \ldots, S_k)$ *good*, if $|S_i| \geq \ell$ for every $i = 1, \ldots, k$. We define good Delaunay simplices and good permutahedron faces accordingly using Proposition 3.5.

Our proof has two main ingredients: First, we show that a good Delaunay simplex either gives birth to or kills an interval in the Čech module that has a lifetime of at least $\frac{\ell}{8(d+1)^2}$. This justifies our notion of "good", since good k-simplices create features that have to be preserved by a sufficiently precise approximation. Second, we show that there are $2^{\Omega(d \log \ell)}$ good k-partitions, so good faces are abundant in the permutahedron.

7.2 Persistence of Good Simplices

Let us consider our first statement. Recall that α_σ is the tower value of σ in the Čech filtration. It will be convenient to have an upper bound for α_σ. Clearly, such a value is given by the diameter of P. It is not hard to see the following bound (compare Lemma 3.3), which we state for reference:

Lemma 7.2 *The diameter of P is at most $2\sqrt{d}$. Consequently, $\alpha_\sigma \leq 2\sqrt{d}$ for each simplex σ of the Čech filtration.*

Recall that by fixing a simplex-wise tower of the Čech filtration, it makes sense to talk about the persistence of an interval associated to a simplex. Fix a $(k-1)$-simplex σ of $\mathcal{D}_P$ incident to o (which also belongs to the Čech filtration).

Lemma 7.3 *Let f_σ be the $(d-k)$-face of Π dual to σ, and let o_σ denote its barycenter. Then, α_σ is the distance of o_σ from o.*

Proof o_σ is the closest point to o on f_σ because $\mathbf{o}o_\sigma$ is orthogonal to $\mathbf{p}o_\sigma$ for any boundary vertex p of f_α. Since f_σ is dual to σ, all vertices of σ are in same distance to o_σ. ☐

Recall L_σ and L_σ^* from Sect. 2 as the difference of the alpha value of σ and its (co-)facets.

Theorem 7.4 *For a good simplex σ of $\mathcal{D}_P$, both L_σ and L_σ^* are at least $\frac{\ell}{24(d+1)^{3/2}}$.*

Proof We start with L_σ^*. Let σ be a $(k-1)$-simplex and let $S_1, \ldots, S_k$ be the corresponding partition. We obtain a co-facet τ of σ through splitting one S_i into two non-empty parts.

The main step is to bound the quantity $\alpha_\tau^2 - \alpha_\sigma^2$. By Lemma 7.3, the alpha values are the squared norms of the barycenters o_τ of τ and o_σ of σ, respectively. It is possible to derive an explicit expression of the coordinates of o_σ and o_τ. It turns out that almost all coordinates are equal, and thus cancel out in the sum, except at those indices that lie in the split set S_i.

Recall that α_σ^2 is the squared length of the barycenter o_σ, and an analogue statement holds for o_τ. Also, recall that τ is obtained from σ by splitting one S_i in the corresponding partition $(S_1, \ldots, S_k)$ of σ. Assume wlog that S_k is split into S_k' and S_{k+1}' (splitting any other S_i yields the same bound) and that S_k is of size exactly ℓ (a larger cardinality only leads to a larger difference).

Let $s_i := |S_i|$ and $p_i = \sum_{j=1}^{i-1} |s_j|$. Recall that Π is spanned by a permutations of a particular point in $\mathbb{R}^{d+1}$, defined in Sect. 3; we order these coordinates values by size in increasing order. Then, the indices in S_i will contain the coordinate values of order $p_i + 1, \ldots, p_i + s_i$. Writing a_i for their average, the symmetric structure of Π implies that o_σ has value a_i in each coordinate $j \in S_i$. Doing the same construction for τ, we observe that the coordinates of o_σ and o_τ coincide for every coordinate $j \in S_1, \ldots, S_{k-1}$; the only differences appear for coordinate indices of S_k, that is, the partition set that was split to obtain τ from σ. Writing a_k, a_k', a_{k+1}' for the average values of S_k, S_k', S_{k+1}', respectively, and $t := |S_k'|$, we get

$$\alpha_\tau^2 - \alpha_\sigma^2 = \sum_{i=1}^{t} \left((a_k')^2 - a_k^2\right) + \sum_{i=t+1}^{\ell} \left((a_{k+1}')^2 - a_k^2\right)$$
$$= t\left((a_k')^2 - a_k^2\right) + (\ell - t)\left((a_{k+1}')^2 - a_k^2\right)$$

To obtain a_k, a_k', and a_{k+1}', we only need to compute the average of the appropriate coordinate values. A simple calculation shows that $a_k = \frac{(d+1)-\ell}{2(d+1)}$, $a_k' = \frac{(d+1)-i}{2(d+1)}$ and $a_{k+1}' = \frac{(d+1)-\ell-i}{2(d+1)}$. Plugging in these values yields

$$\alpha_\tau^2 - \alpha_\sigma^2 = \frac{(d+1+\ell)t(\ell-t)}{4(d+1)^2},$$

whose minimum is achieved for $t = 1$ (and $t = \ell - 1$). Therefore,

$$\alpha_\tau^2 - \alpha_\sigma^2 \geq \frac{(d+1+\ell)(\ell-1)}{4(d+1)^2} \geq \frac{\ell-1}{4(d+1)}.$$

Moreover, $\alpha_\tau \leq 2\sqrt{d}$ by Lemma 7.2. This yields

$$\alpha_\tau - \alpha_\sigma = \frac{\alpha_\tau^2 - \alpha_\sigma^2}{\alpha_\tau + \alpha_\sigma} \geq \frac{\alpha_\tau^2 - \alpha_\sigma^2}{2\alpha_\tau} \geq \frac{\ell-1}{16(d+1)\sqrt{d}} \geq \frac{\ell}{24(d+1)^{3/2}}$$

for $\ell \geq 3$. The bound on L_σ^* follows. For L_σ, note that $\min_{\tau \text{ facet of } \sigma} L_\tau^* \leq L_\sigma$, so it is enough to bound L_τ^* for all facets of σ. With σ being a $(k-1)$-simplex, all but one of its facets are obtained by merging two consecutive S_i and S_{i+1}. However, the obtained partition is again good (because σ is good), so the first part of the proof yields the lower bound for all these facets. It remains to argue about the facet of σ that is not attached to the origin. For this, we change the origin to any vertex of σ. It can be observed (through the combinatorial properties of Π) that with respect to the new origin, σ has the representation $(S_j, \ldots, S_k, S_1, \ldots, S_{j-1})$, thus the partition is cyclically shifted. In particular, σ is still good with respect to the new origin. We obtain the missing facet by merging the (now consecutive) sets S_k and S_1, which is also a good face, and the first part of the statement implies the result. □

As a consequence of Theorem 7.4, the interval associated with a good simplex has length at least $\frac{\ell}{24(d+1)^{3/2}}$ using Lemmas 2.1 and 2.2. Moreover, the interval cannot persist beyond the scale $2\sqrt{d}$ by Lemma 7.2. It follows

Corollary 7.5 *The interval associated to a good simplex is δ-significant for $\delta <$* $\frac{\ell}{96(d+1)^2}$.

7.3 The Number of Good Simplices

We assume for simplicity that $d + 1$ is divisible by ℓ. We call a good partition $(S_1, \ldots, S_k)$ *uniform*, if each set consists of *exactly* ℓ elements. This implies that $k = (d+1)/\ell$.

Lemma 7.6 *The number of uniform good partitions is exactly $\frac{(d+1)!}{\ell!^{(d+1)/\ell}}$.*

Proof Choose an arbitrary permutation and place the first ℓ entries in the S_1, the second ℓ entries in S_2, and so forth. In each S_i, we can interchange the elements and obtain the same k-simplex. Thus, we have to divide out $\ell!$ choices for each of the $(d+1)/\ell$ bins. □

We use this result to bound the number of good k-simplices in the following theorem. To obtain the bound, we use estimates for the factorials using Stirling's approximation. Moreover, we fix some constant $\rho \in (0, 1)$ and set $\ell = (d+1)^\rho$. After some calculations (see Appendix), we obtain:

Theorem 7.7 *For any constant $\rho \in (0, 1)$, $\ell = (d + 1)^\rho$, $k = (d + 1)/\ell$ and d large enough, there exists a constant $\lambda \in (0, 1)$ that depends only on ρ, such that the number of good k-simplices is at least $(d + 1)^{\lambda(d+1)} = 2^{\Omega(d \log d)}$.*

Putting everything together, we prove our lower bound theorem:

Theorem 7.8 *There exists a point set of n points in $d = \Theta(\log n)$ dimensions, such that any $(1+\delta)$-approximation of its Čech filtration contains $2^{\Omega(d \log d)}$ intervals in its persistent barcode, provided that $\delta < \frac{1}{96(d+1)^{1+\varepsilon}}$ with an arbitrary constant $\varepsilon \in (0, 1)$.*

Proof Setting $\rho := 1 - \varepsilon$, Theorem 7.7 guarantees the existence of $2^{\Omega(d \log d)}$ good simplices, all in a fixed dimension k. In particular, the intervals of the Čech persistence module associated to these intervals are all distinct. Since $\ell = (d+1)^{1-\varepsilon}$, Corollary 7.5 states that all these intervals are significant because $\delta < \frac{1}{96d^{1+\varepsilon}} = \frac{\ell}{96(d+1)^2}$. Therefore, by Lemma 7.1, any $(1+\delta)$-approximation of the Čech filtration has $2^{\Omega(d \log d)}$ intervals in its barcode. □

Replacing d by $\log n$ in the bounds of theorem, we see the number of intervals appearing in any approximation super-polynomial is n if δ is small enough.

8 Conclusion

We presented upper and lower bound results on approximating Rips and Čech filtrations of point sets in arbitrarily high dimensions. For Čech complexes, the major result can be summarized as: for a dimension-independent bound on the complex size, there is no way to avoid a super-polynomial complexity for fine approximations of about $O(\log^{-1} n)$, while polynomial size can be achieved for rough approximation of about $O(\log^2 n)$.

Filling in the large gap between the two approximation factors is an attractive avenue for future work. A possible approach is to look at other lattices. It seems that lattices with good covering properties are correlated with a good approximation quality, and it may be worthwhile to study lattices in higher dimension which improve largely on the covering density of A^* (e.g., the Leech lattice [11]).

Further research directions include approximations using small-sized triangulations of cubes, such as the barycentric subdivision. Since the ratio of the diameter to the shortest distance between non-adjacent cells is less for cubes compared to permutahedra, this approach can yield superior quality approximations of comparable size. Another possibility for approximating Čech filtrations is to approximate the union of balls with small permutahedra and to take its nerve as the approximation complex. This amounts to replacing the original input points with a fine sample of the union of balls. The approach shows potential for $(1 + \varepsilon)$-approximations.

Appendix: Missing Proofs

Details of the Proof of Lemma 3.8

We start with the proof of the second claim. Assume that k facets $f_1, \ldots, f_k$ of Π are pairwise intersecting. For any facet f_i, there is a partition $(S_i, [d+1]\backslash S_i)$ associated to it. By Lemma 3.1, we have that either $S_i \subset S_j$ or $S_j \subset S_i$ for each $i \neq j$. This means that the S_i are totally ordered, that means, there exists an ordering π of $\{1, \ldots, k\}$ such that $S_{\pi(1)} \subset S_{\pi(2)} \subset \cdots \subset S_{\pi(k)}$. Now, the partition

$$\left(S_{\pi(1)}, S_{\pi(2)}\backslash S_{\pi(1)}, S_{\pi(3)}\backslash S_{\pi(2)}, \ldots, S_{\pi(k)}\backslash S_{\pi(k-1)}, [d+1]\backslash S_{\pi(k)}\right)$$

is a common refinement of all partitions, which implies that the corresponding face is incident to all k facets. This proves the claim.

Now, we prove the first claim. Let $(S_1, [d+1]\backslash S_1)$, $(S_2, [d+1]\backslash S_2)$ be the partitions defining facets f_1 and f_2, respectively. Since f_1 and f_2 are disjoint, we have that $S_1 \not\subset S_2$ and $S_2 \not\subset S_1$ by Lemma 3.1. Let us define the sets $T_1 = S_1\backslash S_2$, $T_2 = S_2\backslash S_1$, $T_3 = S_1 \cap S_2$ and $T_4 = [d+1]\backslash S_1 \cup S_2$. Also, let $|T_1| = a$, $|T_2| = b$ and $|T_3| = c$ with $a, b, c \geq 1$. Then, $|T_4| = d+1-(a+b+c)$, $|S_1| = k := a+c$ and $|S_2| = p := b+c$.

Let ℓ_1, ℓ_2 denote the lattice points at the centers of the permutahedra Π_1, Π_2 that are attached to Π on the faces f_1 and f_2, respectively. We can derive the coordinates of ℓ_1 and ℓ_2 easily: an elementary calculation shows that barycenter of the face f_1 has coordinates $\frac{k-(d+1)}{2(d+1)} = \frac{k}{2(d+1)} - \frac{1}{2}$ at indices in S_1 and $\frac{k}{2(d+1)}$ at the rest of the positions. Similarly, the barycenter of f_2 has coordinates $\frac{p-(d+1)}{2(d+1)} = \frac{p}{2(d+1)} - \frac{1}{2}$ at indices in S_2 and $\frac{p}{2(d+1)}$ otherwise. Since Π is centered at the origin, the coordinates of ℓ_1 and ℓ_2 are obtained by multiplying these coordinates with 2. See Table 1 for details.

Let B denote the bisector hyperplane between ℓ_1 and ℓ_2. We show that B is a separating hyperplane for Π_1 and Π_2 with no point of either on the hyperplane, which proves the claim. The vector $n = (n_1, \ldots, n_{d+1}) := \ell_2 - \ell_1$ is a normal vector to B. Then, we define B by $n \cdot (x - m) = 0$ with $m = (\ell_1 + \ell_2)/2$ being the midpoint of ℓ_1 and ℓ_2. See Table 1 for a description of n and m.

Since permutahedra tile space by translation, the vertices of Π_1 are of the form $x_1 = \ell_1 + \pi$ where π is any permutation of $y = \frac{1}{d+1}\left(\frac{d}{2}, \frac{d}{2} - 1, \ldots, \frac{-d}{2}\right)$. Writing $B(x_1) := n \cdot (x_1 - m)$ for the function whose sign determines the halfspace of x_1 with respect to B, we can write $B(x_1) = B(\ell_1 + \pi) = n \cdot (\ell_1 + \pi - m) = n \cdot \ell_1 - n \cdot m + n \cdot \pi$. Similarly, for any vertex $x_2 = \ell_2 + \pi$ of Π_2, $B(x_2) = n \cdot \ell_2 - n \cdot m + n \cdot \pi$. We show that $B(x_1) < 0$ and $B(x_2) > 0$ for all permutations π, which proves the claim. First,

we calculate $n \cdot \ell_1$, $n \cdot \ell_2$ and $n \cdot m$ using Table 1:

$$n \cdot \ell_1 = (\alpha + 1)\left(\frac{k}{d+1} - 1\right)a + (\alpha - 1)\frac{k}{d+1}b + \alpha\left(\frac{k}{d+1} - 1\right)c$$
$$+ \alpha\left(\frac{k}{d+1}\right)\{d + 1 - (a + b + c)\}.$$

Upon simplification, this reduces to $n \cdot \ell_1 = -a + \frac{k}{d+1}(a - b)$. Similarly, one can calculate that $n \cdot \ell_2 = b + \frac{p}{d+1}(a - b)$. Next,

$$n \cdot m = (\alpha + 1)(\beta - 1/2)a + (\alpha - 1)(\beta - 1/2)b + \alpha(\beta - 1)c$$
$$+ \alpha\beta[d + 1 - (a + b + c)]$$

This simplifies to $n \cdot m = -(a - b)\frac{(d+1)-(p+k)}{2(d+1)}$. Subtracting, we get

$$n \cdot \ell_1 - n \cdot m = -a + \frac{k}{d+1}(a - b) + (a - b)\frac{(d + 1) - (p + k)}{2(d + 1)}$$

which reduces to $n \cdot \ell_1 - n \cdot m = -\frac{a+b}{2} + \frac{(b-a)^2}{2(d+1)}$.

Since $\ell_1 - m = -(\ell_2 - m)$, hence $n \cdot (\ell_2 - m) = -n \cdot (\ell_1 - m)$. Also,

$$n \cdot \ell_1 - n \cdot m = -\frac{a + b}{2} + \frac{(b - a)^2}{2(d + 1)} < -\frac{a + b}{2} + \frac{(b + a)^2}{2(d + 1)}$$
$$< -\frac{a + b}{2}\left(1 - \frac{a + b}{d + 1}\right) < 0.$$

Hence, $n \cdot \ell_1 - n \cdot m$ is negative and $n \cdot \ell_2 - n \cdot m$ is positive. Substituting these values in $B(x_1)$ and $B(x_2)$, we get

$$B(x_1) = -\frac{a + b}{2} + \frac{(b - a)^2}{2(d + 1)} + n \cdot \pi, \quad B(x_2) = \frac{a + b}{2} - \frac{(b - a)^2}{2(d + 1)} + n \cdot \pi.$$

We now calculate the maximum absolute value of $n \cdot \pi$ and use it to show that $B(x_1)$ is always negative and $B(x_2)$ is always positive.

Table 1 ℓ_1, ℓ_2, n, m. Here, $e = d + 1 - (a + b + c)$

Indices	ℓ_2	ℓ_1	$n = \ell_2 - \ell_1$	$m = (\ell_2 + \ell_1)/2$	Count
T_1	$\frac{p}{d+1}$	$\frac{k}{d+1} - 1$	$\frac{(p-k)}{d+1} + 1 = \alpha + 1$	$\frac{(p+k)}{2(d+1)} - \frac{1}{2} = \beta - 1/2$	a
T_2	$\frac{p}{d+1} - 1$	$\frac{k}{d+1}$	$\frac{(p-k)}{d+1} - 1 = \alpha - 1$	$\frac{(p+k)}{2(d+1)} - \frac{1}{2} = \beta - 1/2$	b
T_3	$\frac{p}{d+1} - 1$	$\frac{k}{d+1} - 1$	$\frac{p-k}{d+1} = \alpha$	$\frac{(p+k)}{2(d+1)} - 1 = \beta - 1$	c
T_4	$\frac{p}{d+1}$	$\frac{k}{d+1}$	$\frac{p-k}{d+1} = \alpha$	$\frac{p+k}{2(d+1)} = \beta$	e

The dot product $n \cdot \pi$ is obtained by first multiplying each component y_i of the vector $y = \frac{1}{d+1}(\frac{d}{2}, \frac{d}{2} - 1, \ldots, \frac{-d}{2})$ with a component of n, which has one of three values: $\alpha + 1$ for indices in T_1, α for $T_3 \cup T_4$, $\alpha - 1$ for T_2 (refer Table 1); the intermediate products are then added up. The permutation of y maximizing $n \cdot \pi$ follows from a simple arithmetic fact, which can be proved by a simple induction on the dimension of the vector.

Lemma 8.1 *For any natural number $N \geq 2$, let $V = (v_1, \ldots, v_N)$ and $W = (w_1, \ldots, w_N)$ be two vectors in $\mathbb{R}^N$ with $v_1 \leq \ldots \leq v_N$ and $w_1 \leq \ldots \leq w_N$. Let π be a permutation over $[N]$, and let $\pi(W)$ be the vector with the corresponding permuted coordinates of W. Then, $\max_{\pi}\{V \cdot \pi(W)\} = V \cdot W$.*

Let us denote the sum of the q smallest components of y by N_q and the sum of the q largest components of y by M_q. It is easy to verify that $M_q + N_q = 0$, $N_q = N_{d+1-q}$ and $M_q = M_{d+1-q}$. Then,

$$
\begin{aligned}
\max(|n \cdot \pi|) &= (\alpha + 1)M_a + (\alpha - 1)N_b + \alpha(N_{d+1-a} - N_b) \\
&= \alpha(M_a + N_{d+1-a}) + M_a - N_b \\
&= 0 - N_a - N_b = -\left[\frac{a\{a - (d + 1)\}}{2(d + 1)} + \frac{b\{b - (d + 1)\}}{2(d + 1)}\right] \\
&= \frac{a + b}{2} - \frac{a^2 + b^2}{2(d + 1)} < \frac{a + b}{2} - \frac{(b - a)^2}{2(d + 1)}.
\end{aligned}
$$

The last inequality implies that

$$
B(x_1) = -\frac{a + b}{2} + \frac{(b - a)^2}{2(d + 1)} + n \cdot \pi < 0,
$$

and similarly, $B(x_2) > 0$. The claim follows. $\qquad\square$

Proof of Theorem 7.7

Recall from Lemma 7.6 that the number of good simplices is at least

$$
\frac{(d + 1)!}{\ell!^{(d+1)/\ell}}.
$$

Stirling's approximation [31] states that

$$
\sqrt{2\pi} n^{n+1/2} e^{-n+1/(12n+1)} < n! < \sqrt{2\pi} n^{n+1/2} e^{-n+1/(12n)}.
$$

We rephrase the upper bound as

$$
\sqrt{2\pi} n^{n+1/2} e^{-n+1/(12n)} \leq \sqrt{2\pi} e^{1/(12n)} n^{n+1/2} e^{-n} \leq e n^{n+1/2} e^{-n}
$$

for $n \geq 2$ and the lower bound simply as

$$\sqrt{2\pi} n^{n+1/2} e^{-n+1/(12n+1)} \geq n^n e^{-n}.$$

In this way, we can lower bound the number of good simplices as

$$\frac{(d+1)!}{\ell!^{(d+1)/\ell}} \geq \frac{(d+1)^{(d+1)} e^{-(d+1)}}{(e\ell^{\ell+1/2} e^{-\ell})^{(d+1)/\ell}}$$

$$\geq \frac{(d+1)^{(d+1)} e^{-(d+1)}}{e^{(d+1)/\ell} \ell^{(d+1)+(d+1)/(2\ell)} e^{-(d+1)}}$$

$$\geq \exp\left((d+1)\log(d+1) - \frac{(d+1)}{\ell} - (d+1)\log\ell\left(1 + \frac{1}{2\ell}\right)\right).$$

Choose $\ell = (d+1)^\rho$ with some constant $\rho < 1$. The above simplifies to

$$\exp\left((d+1)\log(d+1) - (d+1)^{1-\rho} - \rho(d+1)\log(d+1)\left(1 + \frac{1}{2(d+1)^\rho}\right)\right)$$

$$= \exp\left((d+1)\log(d+1)\left(1 - \rho\left(1 + \frac{1}{2(d+1)^\rho}\right)\right) - (d+1)^{1-\rho}\right).$$

Now, pick some $\lambda \in [0, 1]$ such that $\rho < 1 - 2\lambda < 1$. We have that $\rho\left(1 + \frac{1}{2(d+1)^\rho}\right) < 1 - 2\lambda$ for d large enough. Thus, for d large enough,

$$\exp\left((d+1)\log(d+1)\left(1 - \rho\left(1 + \frac{1}{2(d+1)^\rho}\right)\right) - (d+1)^{1-\rho}\right)$$

$$\geq \exp\left(2\lambda(d+1)\log(d+1) - (d+1)^{1-\rho}\right) \geq \exp\left(\lambda(d+1)\log(d+1)\right). \qquad \square$$

References

1. Baek, J., Adams, A.: Some useful properties of the permutohedral lattice for Gaussian filtering. Stanford University. http://graphics.stanford.edu/papers/permutohedral/permutohedral_techreport.pdf (2009)
2. Bambah, R.P.: On lattice coverings by spheres. Proc. Indian Natl. Sci. Acad. **20**(1), 25–52 (1954)
3. Bernoulli's inequality. https://en.wikipedia.org/wiki/Bernoulli's_inequality
4. Borsuk, K.: On the imbedding of systems of compacta in simplicial complexes. Fundam. Math. **35**, 217–234 (1948)
5. Botnan, M.B., Spreemann, G.: Approximating persistent homology in Euclidean space through collapses. Appl. Algebra Eng. Commun. Comput. **26**(1–2), 73–101 (2015)
6. Bourgain, J.: On Lipschitz embedding of finite metric spaces in Hilbert space. Isr. J. Math. **52**(1–2), 46–52 (1985)
7. Bubenik, P., Scott, J.A.: Categorification of persistent homology. Discrete Comput. Geom. **51**(3), 600–627 (2014)
8. Callahan, P.B., Kosaraju, S.R.: A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. J. ACM **42**(1), 67–90 (1995)
9. Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L.J., Oudot, S.Y.: Proximity of persistence modules and their diagrams. In: Proceedings of the 25th Annual Symposium on Computational Geometry (SoCG'09), pp. 237–246. ACM, New York (2009)
10. Choudhary, A., Kerber, M., Raghvendra, S.: Polynomial-sized topological approximations using the permutahedron. In: Proceedings of the 32nd International Symposium on Computational Geometry

(SoCG'16). Leibniz International Proceedings in Informatics, vol. 51, pp. 1–16. Schloss Dagstuhl, Dagstuhl (2016)

11. Conway, J.H., Sloane, N.J.A.: Sphere Packings, Lattices, and Groups. Grundlehren der Mathematischen Wissenschaften, vol. 290. With additional contributions by Bannai, E. et al. Springer, New York (1988)

12. Dey, T.K., Fan, F., Wang, Y.: Computing topological persistence for simplicial maps. In: Proceedings of the 30th Annual Symposium on Computational Geometry (SoCG'14), pp. 345–354. ACM, New York (2014)

13. Edelsbrunner, H., Harer, J.L.: Computational Topology: An Introduction. American Mathematical Society, Providence (2010)

14. Edelsbrunner, H., Mücke, E.P.: Simulation of simplicity: a technique to cope with degenerate cases in geometric algorithms. ACM Trans. Graph. **9**(1), 66–104 (1990)

15. Hales, T.C.: A proof of the Kepler conjecture. Ann. Math. **162**(3), 1065–1185 (2005)

16. Har-Peled, S.: Geometric Approximation Algorithms. Mathematical Surveys and Monographs, vol. 173. American Mathematical Society, Providence (2011)

17. Hatcher, A.: Algebraic Topology. Cambridge University Press, Cambridge (2002)

18. Johnson, W.B., Lindenstrauss, J., Schechtman, G.: Extensions of Lipschitz maps into Banach spaces. Isr. J. Math. **54**(2), 129–138 (1986)

19. Jung, H.: Über die kleinste Kugel, die eine räumliche Figur einschliesst. J. Reine Angew. Math. **123**, 241–257 (1901)

20. Kerber, M., Raghvendra, S.: Approximation and streaming algorithms for projective clustering via random projections. In: Proceedings of the 27th Canadian Conference on Computational Geometry (CCCG'15), pp. 179–185 (2015)

21. Kerber, M., Schreiber, H.: Barcodes of towers and a streaming algorithm for persistent homology. In: Accepted to Proceedings of the 33rd International Symposium on Computational Geometry (SoCG'17), pp. 57:1–57:15 (2017)

22. Kerber, M., Sharathkumar, R.: Approximate Čech complex in low and high dimensions. In: Cai, L., Cheng, S.-W., Lam, T.-W. (eds.) Algortihms and Computation (ISAAC'13). Lecture Notes in Computer Science, vol. 8283, pp. 666–676. Springer, Heidelberg (2013)

23. Khuller, S., Matias, Y.: A simple randomized sieve algorithm for the closest-pair problem. Inform. and Comput. **118**(1), 34–37 (1995)

24. Lovász, L., Simonovits, M.: Random walks in a convex body and an improved volume algorithm. Random Struct. Algorithms **4**(4), 359–412 (1993)

25. Matoušek, J.: Bi-Lipschitz embeddings into low-dimensional Euclidean spaces. Commentat. Math. Univ. Carol. **31**(3), 589–600 (1990)

26. Munkres, J.R.: Elements of Algebraic Topology. Westview Press, Boulder (1984)

27. Rennie, B.C., Dobson, A.J.: On Stirling numbers of the second kind. J. Comb. Theory **7**(2), 116–121 (1969)

28. Sheehy, D.: Linear-size approximations to the Vietoris–Rips filtration. Discrete Comput. Geom. **49**(4), 778–796 (2013)

29. Sheehy, D.: The persistent homology of distance functions under random projection. In: Proceedings of the 30th Annual Symposium on Computational Geometry (SoCG'14), pp. 328–334. ACM, New York (2014)

30. Smid, M.H.M.: The well-separated pair decomposition and its applications. In: Gonzalez, T.F. (ed.) Handbook of Approximation Algorithms and Metaheuristics, pp. 53-1–53-12. Chapman and Hall/CRC, Boca Raton (2007)

31. Stirling's approximation for factorials. https://en.wikipedia.org/wiki/Stirling's_approximation

32. Ziegler, G.M.: Lectures on Polytopes. Graduate Texts in Mathematics, vol. 152. Springer, New York (1995)