PERSISTENT HOMOLOGY AND THE UPPER BOX DIMENSION

BENJAMIN SCHWEINHART

Abstract. We introduce a fractal dimension for a metric space based on the persistent homology of subsets of that space. We exhibit hypotheses under which this dimension is comparable to the upper box dimension; in particular, the dimensions coincide for subsets of $\mathbb{R}^2$ whose upper box dimension exceeds 1.5.

1. Introduction

Several notions of fractal dimensions based on persistent homology have been proposed in the literature, including in the PhD thesis of Vanessa Robins [15], in a paper written by Robert MacPherson and the current author [14], and in unpublished work by the members of the Pattern Analysis Library at CSU [1]. In that work, empirical estimates of the proposed dimensions were compared with classically defined fractal dimensions. Here, we prove the first rigorous analogue of those comparisons.

The motivation for the definition proposed here comes from the literature on minimal spanning trees. The properties of minimal spanning trees of point collections contained in a bounded subset of $\mathbb{R}^m$ have long been of interest [18, 12]. In 2005, Kozma, Lotker, and Stupp [13] proved a relationship between the total lengths of these trees and the upper box dimension. To be precise, let $T(Y)$ denote the minimal spanning tree of a finite point set Y in a metric space and let

$$E_\alpha^0(Y) = \frac{1}{2}\sum_{e\in T(Y)} \|e\|^\alpha$$

where the sum is taken over all edges e in the tree $T(Y)$, and $\|e\|$ denotes the length of the edge. Define $\dim_{\mathrm{MST}}(X)$ to be the infimal exponent β so that $E_\beta^0\left(\left\{x_j\right\}\right)$ is uniformly bounded for all finite point sets $\left\{x_j\right\} \subset X$:

Date: March 2018.

$$\dim_{\mathrm{MST}}(X) = \inf\left\{\beta : E_\beta^0\left(\left\{x_j\right\}\right) < C\ \forall\ \left\{x_j\right\} \subset X\right\}$$

Theorem 1 (Kozma, Lotker, and Stupp)**.** *For any metric space X,*

$$dim_{MST}(X) = dim_{box}(X)$$

where $\dim_{box}(X)$ denotes the upper box dimension of X (defined below).

If Y is a finite point set contained in a triangulable metric space, and $PH_i(Y)$ is the i-dimensional persistent homology of the Čech complex of Y, there is a bijection between the edges of the Euclidean minimal spanning tree of Y and the intervals in the canonical decomposition of $PH_0(Y)$, where the length of an interval in $PH_0(Y)$ is half the length of the corresponding edge. (Note that this is also true for the persistent homology of the Rips complex of Y, without the requirement of an ambient space) This observation suggests a generalization of the previous result to higher-dimensional persistent homology. Namely, let

$$E_\alpha^i(Y) = \sum_{(b,d)\in PH_i(Y)} (d-b)^\alpha$$

where the sum is taken over all bounded PH_i intervals, and

$$\dim_{PH}^i(X) = \inf\left\{\beta : E_\beta^i\left(\left\{x_j\right\}\right) < C\ \forall\ \left\{x_j\right\} \subset X\right\}$$

Also, let $\dim_{\widetilde{PH}}^i(X)$ be defined by replacing the Čech complex with the Rips complex in the previous construction. Then

$$\dim_{PH}^0(X) = \dim_{\widetilde{PH}}^0(X) = \dim_{\mathrm{MST}}(X)$$

Note that $\dim_{\widetilde{PH}}^i(X)$ is defined for all bounded metric spaces X, rather than subsets of a triangulable metric space.

Question 1. *Are there hypotheses on X under which $dim_{PH}^i(X) = dim_{box}(X)$ or $dim_{\widetilde{PH}}^i(X) = dim_{box}(X)$ for $i > 0$?*

Unlike the 0-dimensional case, equality will not always hold. For example, the PH_1-dimension of a line in $\mathbb{R}^m$ will always be 0. Also, there are metric spaces not embeddable in any finite-dimensional Euclidean space whose $\widetilde{PH}_1$-dimension exceeds their upper box dimension, which is why we restrict our attention to subsets of Euclidean space. See Section 3 for such an example. However, we conjecture:

Conjecture 1. *For any $m \in \mathbb{N}, 0 \leq i < m$, there is a constant $\gamma(i,m) < m$ so that for any bounded subset X of $\mathbb{R}^m$ satisfying $dim_{box}(X) > \gamma(i,m)$,*

$$dim^i_{PH}(X) = dim_{box}(X)$$

Cohen-Steiner, Edelsbrunner, Harer, and Mileyko studied a quantity similar to E^i_α in their paper on the L_p−stability of the persistent homology of Lipschitz functions [6]. Their results immediately imply that if $X \subset \mathbb{R}^m$, then $\dim^i_{PH}(X) \leq m$ for all $i \in \mathbb{N}$. If U is a bounded, convex subset of $\mathbb{R}^m$, the corresponding lower bound is easy to show and $\dim^i_{PH}(X) = m$ for $i = 1, \ldots, m-1$. However, our focus here will be to prove results about subsets of fractional box dimension. This task is challenging, and involves difficult combinatorial problems. The main result of this paper is:

Theorem 2. *Let X be a bounded subset of $\mathbb{R}^2$. If $dim_{box}(X) > 1.5$, then*

$$dim^1_{PH}(X) = dim_{box}(X)$$

The upper bound $\dim^1_{PH}(X) \leq \dim_{box}(X)$ is proven in Section 4, and the lower bound in Section 6.

We also prove partial results in more general cases:

Theorem 3. *Let X be a bounded subset of $\mathbb{R}^m$. If $dim_{box}(X) > m - 1/2$, then*

$$dim_{box}(X) \leq dim^1_{PH}(X) \leq m$$

Also,

Theorem 4. *If X is a bounded subset of $\mathbb{R}^m$ then*

$$dim^1_{\widetilde{PH}}(X) \leq dim_{box}(X)$$

In the process of proving this, we also show:

Theorem 5. *If X is a finite subset of $\mathbb{R}^m$ then then the first-dimensional persistent homology of the Rips complex of X contains $O(|X|)$ intervals.*

2. Preliminaries

Let X be a bounded metric space and let $N_\delta(X)$ be the maximal number of disjoint closed δ-balls with centers in X.

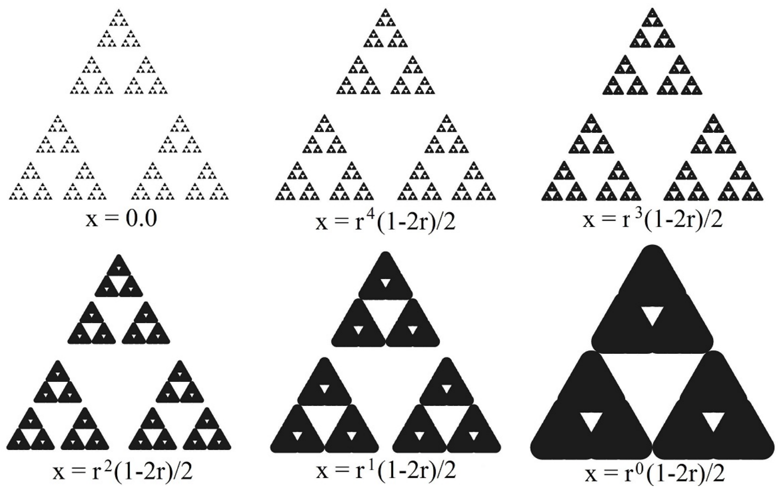


Figure 1. A self-similar fractal and its r-neighborhoods [14]

Definition 1. *The* upper box dimension *of* X *is*

$$dim_{box}(X) = \limsup_{\delta \to 0} \frac{\log(N_\delta(X))}{-\log(\delta)}$$

There are several equivalent definitions. Here, we will also use that $N_\delta(X)$ can be taken to be either the smallest number of δ-balls required to cover X or — if $X \subset \mathbb{R}^m$ — the number of cubes in the grid of mesh δ on $\mathbb{R}^m$ (that is, the grid formed by taking the Voronoi cells of the points of $\frac{1}{\delta}\mathbb{Z}^m$ in $\mathbb{R}^m$) which intersect X. See Falconer [10] for details.

2.1. **Persistent Homology.**

2.1.1. *Filtrations.* If X is a subset of a metric space M and $\epsilon > 0$, let the **ϵ-neighborhood** of X, denoted X_ϵ, be the set of all points within distance ϵ of X :

$$X_\epsilon = \{x \in M : d(x,y) < \epsilon \text{ for some } y \in X\}$$

See Figure 1 for an example. A **filtration** is a sequence of topological spaces $\{S_\alpha\}_{\alpha \in I}$ together with continuous maps $i : S_\alpha \to S_\beta$ for $\alpha, \beta \in I, \alpha < \beta$, where I is an ordered index set — usually the positive real numbers, the natural numbers, or a finite set of real numbers. For example, the ϵ-neighborhoods of a subset of a metric space, together with the inclusion maps between them, form a filtration indexed in the positive e real numbers.

If $\{x_i\}$ is a finite subset of a metric space, the **Čech complex** is a filtration of simplicial complexes $\left\{C_{\alpha_i}\right\}$ indexed by a finite set of real numbers $\{0=\alpha_1,\alpha_2,\ldots,\alpha_n\}$ so C_{α_j} is homotopy equivalent to the ϵ-neighborhood $\{x_i\}_\epsilon$ for all $\alpha_j \leq \epsilon < \alpha_{j+1}$. The **Alpha complex** of a finite subset of Euclidean space is a filtration on the simplices of the Delaunay triangulation that also has this homotopy equivalence property. In general, it is smaller and more amenable to computation than the Čech complex. [9] Several other complexes, such as the **Rips complex** [20, 8] and intrinsic Čech [4] complex, can be defined for finite metric spaces without reference to an ambient metric space.

2.1.2. *Definition of Persistent Homology.* Let $H_i(X)$ denote the reduced homology of a topological space, with coefficients in a field k. If X_α is a filtration of topological spaces, the **persistent homology** of X_α, is the product $\prod_\alpha H_i\left(X_j\right)$, together with the maps $i_{\alpha,\beta}: H_i(X_\alpha) \to H_i\left(X_\beta\right)$ for $\alpha < \beta$. If $H_i(X_\alpha)$ is finite dimensional for all α, the structure of the persistent homology of X_α is captured by a set of intervals [21, 7] that track the birth and death of homology generators through the filtration. Note that if Y is a bounded subset of a triangulable, compact metric space, the ϵ-neighborhood filtration satisfies the required finiteness hypotheses [2, 3]. Slightly abusing notation, we will refer to this set of intervals as the persistent homology of the filtration. In particular, if X is a subset of a triangulable metric space, $PH_i(X)$ will refer to the set of intervals of the i-dimensional persistent homology of the ϵ-neighborhood filtration of X. $PH_i(X)$ depends on the ambient space.

We will repeatedly use a stability theorem for persistent homology [5, 4]. If X and Y are filtrations, let the bottleneck distance between $PH_i(X) = \{(b_i,d_i)\}$ and $PH_i(Y) = \left\{\left(\hat{b}_j,\hat{d}_j\right)\right\}$ be

$$d_B\left(PH_i(X), PH_i(Y)\right) = \inf_\eta \sup_i \max\left(\left|b_i - \hat{b}_{\eta(i)}\right|, \left|d_i - \hat{d}_{\eta(i)}\right|\right)$$

where η ranges over all matchings between the intervals $PH_i(X)$ and $PH_i(Y)$, allowing intervals to be matched to intervals of length zero.

Theorem 6. *(Stability of the Bottleneck Distance) If X and Y are bounded subsets of a metric space, $i \in \mathbb{N}$, and d_H denotes the Hausdorff distance then*

$$d_B\left(PH_i(X), PH_i(Y)\right) \leq d_H(X,Y)$$

In particular, this result implies the following corollary:

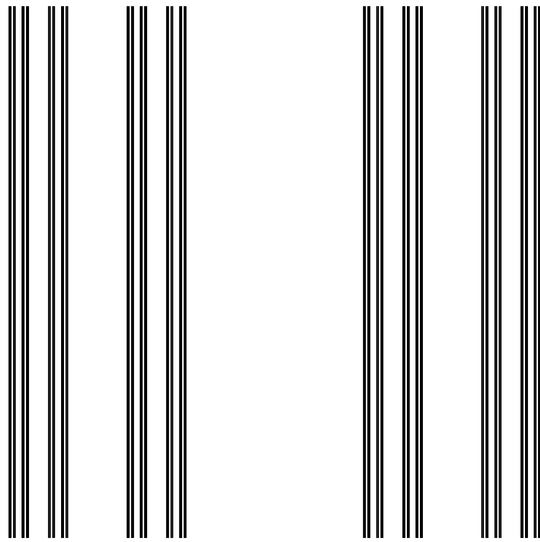


Figure 2. The Cantor set cross an interval. If C is the Cantor set and I is an interval, $\dim^0_{\widehat{PH}}(C \times I) = \dim^0_{\widehat{PH}}(C) = \log_3 2$ and $\dim^1_{\widehat{PH}}(C \times I) = 0$, but $\dim^0_{PH}(C \times I) = \dim^1_{PH}(C \times I) = \dim_{box}(C \times I) = 1 + \log_3 2$.

Corollary 1. *Let X and Y are bounded subsets of a triangulable metric space, $\epsilon, \delta > 0$, and $d_H(X, Y) < \epsilon/2$ then*

$$|\{(b, d) \in PH_i(X) : d - b > \epsilon + \delta\}| \leq |\{(b, d) \in PH_i(Y) : d - b > \delta\}|$$

2.2. **Persistent Homology Dimension.**

Definition 2. *The i-th dimensional persistent homology dimension of bounded subset of a triangulable metric space is*

$$dim^i_{\widehat{PH}}(X) = \inf\left\{\beta : E^i_\beta(X) < \infty\right\}$$

This is equivalent to the definition given in "Measuring Shape with Topology" [14] (Corollary 7), and does not agree with any classically defined fractal dimension in general. It instead measures the complexity of a shape. See Figure 2 for an example.

Definition 3. *The i-th dimensional persistent homology sampling dimension of a bounded subset X of a triangulable metric space is*

$$dim^i_{PH}(X) = \inf\left\{\beta : E^i_\beta\left(\{x_j\}\right) < C \;\forall\; \{x_j\} \subset X\right\}$$

Here, we will refer to $\dim^i_{PH}(X)$ as the PH_i-dimension of X. It is a consequence of the stability of the bottleneck distance that $\dim^i_{\widetilde{PH}}(X) \leq \dim^i_{PH}(X)$. See Section A for a proof. A similar argument shows that we could modify the definition of $\dim^i_{PH}(X)$ to consider all subsets of X, rather than all finite point sets:

Proposition 1. *Let X be a bounded subset of a triangulable metric space. Then*

$$dim^i_{PH}(X) = \inf\left\{\beta : E^i_\beta(Y) < C \;\forall\; Y \subseteq X\right\}$$

We define corresponding quantities in terms of the Rips complex. If Y is a finite point set, let $\widetilde{PH}_i(Y)$ be the i-th dimensional persistent homology of the Rips complex of Y and let

$$\widetilde{E}^i_\alpha(Y) = \sum_{(b,d)\in\widetilde{PH}_i(Y)} (d-b)^\alpha$$

Definition 4. *The i-th dimensional Rips persistent homology sampling dimension of a bounded metric space X is*

$$dim^i_{\widetilde{PH}}(X) = \inf\left\{\beta : \widetilde{E}^i_\beta\left(\{x_j\}\right) < C \;\forall\; \{x_j\} \subset X\right\}$$

Note that this dimension is defined for arbitrary bounded metric spaces, and does not require the specification of an ambient space.

3. An Example

Here, we construct an example of a metric space not embeddable in any Euclidean space whose $\widetilde{PH}_1$-dimension equals 2 but whose upper box dimension equals 1. This contrasts with the $\widetilde{PH}_0$-dimension, which was proven by Kozma, Lotker, and Stupp [13] to agree with the upper box dimension for all metric spaces. The rough idea of the construction is to build a space that includes a copies of the complete bipartite graph on 2^n vertices at different scales.

Proposition 2. *There is a metric space X so that $dim_{\widetilde{PH}_1}(X) = 2$ but $dim_{box}(X) = 1$.*

Proof. We will construct X as the union of finite point sets X_n, whose pairwise intersections contain only a common point x^*. Let X_n be the set consisting of 2^{n+1} points $x^n_1 = x^*, x^n_2, \ldots x^n_{2^n}$ and $y^n_1, y^n_2, \ldots y^n_{2^n}$, and let

$$d(x^n_i, y^n_i) = \frac{1}{2^{n+1}} \qquad d\left(x^n_i, x^n_j\right) = \left(1-\delta_{i,j}\right)\frac{1}{2^n} \qquad d\left(y^n_i, y^n_j\right) = \left(1-\delta_{i,j}\right)\frac{1}{2^n}.$$

Let $X = \cup_{i=0}^{\infty} X_i$. if $i \neq j$, $z \in X_i$ and $w \in X_j$, and $z \neq x^*, w \neq x^*$, set

$$d(z,w) = \frac{1}{2^{\min(i,j)}}.$$

Denote the minimal number of balls of radius δ required to cover X by $M_\delta(X)$. If $\delta > \frac{1}{2^n}$, then X is covered by placing a ball at each point of $\cup_{i=0}^{n} X_n$ so

$$M_\delta(X) < \sum_{i=0}^{n} |X_i| < \sum_{i=0}^{n} 2^{n+1} = 2^{n+2} - 2.$$

Therefore

$$\begin{aligned}\dim_{box}(X) = &\\ &\limsup_{\delta\to 0} \frac{\log(M_\delta(X))}{-\log(\delta)}\\ \leq &\limsup_{n\to\infty} \frac{\log\left(2^{n+2}-2\right)}{\log(2^n)}\\ = & 1.\end{aligned}$$

On the other hand, if $\mathcal{R}_\epsilon(X_n)$ denotes the Rips complex of X_n at parameter ϵ, then $\mathcal{R}_\epsilon(X_n)$ consists of 2^{n+1} 0-cells if $\epsilon < \frac{1}{2^{n+1}}$, contains those 0-cells and 2^{2n} edges between the points x_i^n and y_j^n for $\frac{1}{2^{n+1}} \leq \epsilon < \frac{1}{2^n}$ and is contractible for $\epsilon \geq \frac{1}{2^n}$, containing all possible simplices. As such, all intervals of $\widetilde{PH}_1(X_n)$ are born at $\epsilon = \frac{1}{2^{n+1}}$ and die at $\epsilon = \frac{1}{2^n}$. By an Euler characteristic argument, there are $2^{2n} - 2^{n+1} + 1$ such intervals. Therefore

$$\begin{aligned}E_\alpha^1(X_n) = &\\ &\left(2^{2n} - 2^{n+1} + 1\right)\left(\frac{1}{2^n} - \frac{1}{2^{n+1}}\right)^\alpha\\ = &\left(2^{2n} - 2^{n+1} + 1\right)\frac{1}{2^{\alpha n+\alpha}}\\ = & 2^{-\alpha}\left(2^{(2-\alpha)n} - 2^{(1-\alpha)n+1} + 2^{-\alpha n}\right)\\ \approx & 2^{-\alpha}2^{(2-\alpha)n}\end{aligned}$$

which is bounded as $n \to \infty$ if $\alpha > 2$, and unbounded otherwise. Thus $\dim_{\widetilde{PH}_1}(X) \geq 2$. The inequality $\dim_{\widetilde{PH}_1}(X) \leq 2$ can be shown by an argument similar to the one in the proof of Proposition 3 below. □

4. An Upper Bound

Our strategy for bounding the PH_i-dimension above by the upper box dimension involves bounding the number of PH_i-intervals of of X exceeding a certain length by approximating subsets of X by well-spaced finite point sets whose size is controlled by the box dimension. In the following, d_H is the Hausdorff distance.

Lemma 1. *Let X be a bounded metric space, and suppose $\alpha > dim_{box}(X)$. There is a constant $C > 0$ so that for all $\epsilon > 0$ and any $Y \subseteq X$, there is a finite point set $\{x_j\} \subset X$ satisfying:*

$$d_H\left(Y, \{x_j\}\right) < \epsilon/2$$

$$d\left(x_i, x_j\right) \geq \epsilon\left(1 - \delta_{i,j}\right)$$

$$\left|\{x_j\}\right| \leq C\epsilon^{-\alpha}$$

Proof. There exists a constant C' so that any collection of disjoint balls of radius ϵ centered at points of X has fewer than $C'\epsilon^{-\alpha}$ points. If $\{B_j\}$ is a maximal collection of disjoint balls of radius $\epsilon/4$ centered at points of $Y \subset X$ and x_j is the center of B_j, then $d_H\left(Y, \{x_j\}\right) < \epsilon/2$. Therefore the desired statement holds with $C = C'4^{\alpha}$. □

Let X be a bounded subset of a metric space, and let $\phi_X^i(n)$ be the maximal number of PH_i intervals of a a set of n points in X. For example, the number of PH_i intervals of a finite point set in $\mathbb{R}^m$ is less than or equal to the number of i-faces (or $i+1$-faces) of the Delaunay triangulation of the set of points [9]. As such, the Upper Bound Theorem [17] implies that if $X \subset \mathbb{R}^m$ then

$$\phi_X^i(n) = O\left(f_i(m, n)\right)$$

where $f_i(m, n)$ is the number of i-faces in the cyclic polytope with n vertices in $\mathbb{R}^m$.

Proposition 3. *Let X be a bounded subset of a triangulable metric space, and suppose that $\phi_X^i(v) = O(v^c)$. Then*

$$dim_{PH}^i(X) \leq c\ dim_{box}(X)$$

Proof. Rescale the metric if necessary so that the diameter of X is less than one. Let $\beta > c\dim_{\text{box}}(X)$. We will show that $E_\beta(Y)$ is uniformly bounded for all $Y \subset X$.

Find a real number α so that $\dim_{PH}^{i}(X) < \alpha < \beta/c$, and let D_1 be a constant so that $\phi_i^m(v) < D_1 v^c$ for all $v > 0$. Let

$$I_k = \left\{(b,d) \in PH_i(Y) : d - b > \frac{1}{2^k}\right\}$$

By the preceding lemma, we an find a collection $\left\{x_j\right\}$ of fewer than $C2^{k\alpha}$ points so that $d_H\left(\left\{x_j\right\}, Y\right) < \frac{1}{2^k}$. The stability of the bottleneck distance on persistence diagrams implies that each interval of I_k corresponds to an interval in $PH_i(\left\{x_j\right\})$. Therefore,

$$|I_k| \leq \left|PH_i\left(\left\{x_j\right\}\right)\right| \leq D_1 C^c 2^{ck\alpha}$$

for all $k > 0$. Then

$$\begin{aligned} E_\beta^i(Y) = \\ & \lim_{k\to\infty} \sum_{(b,d)\in I_k} (d-b)^\beta \\ & \leq \sum_{k=0}^{\infty} |I_k| \frac{1}{2^{\beta(k-1)}} \\ & \leq \sum_{k=0}^{\infty} D_1 C^c 2^{c\alpha k} \frac{1}{2^{\beta(k-1)}} \\ & = D_1 C^c 2^\beta \sum_{k=0}^{\infty} \left(2^{c\alpha-\beta}\right)^k \\ & < \frac{D_1 C^c 2^\beta}{1 - 2^{c\alpha-\beta}} \end{aligned}$$

because $\beta > c\alpha$, and $E_\beta^i(Y)$ is uniformly bounded for all $Y \subset X$, as desired. □

The Upper Bound Theorem implies the following:

Corollary 2. *Let X be a bounded subset of $\mathbb{R}^m$. If $i < \lfloor \frac{m}{2} \rfloor$ then*

$$dim_{PH}^{i}(X) \leq (i+1)\, dim_{box}(X)$$

If $i \geq \lfloor \frac{m}{2} \rfloor$

$$dim_{PH}^{i}(X) \leq \lfloor \frac{m+1}{2} \rfloor dim_{box}(X)$$

In particular, if $m = 2$

$$dim_{PH}^{1}(X) \leq dim_{box}(X)$$

In the proof of Proposition 3, we bounded the number of intervals of Y of length greater than ϵ by counting the number of intervals of a finite point set which approximates X. A stronger result can be proven by counting the number of "long" intervals of a "well-spaced" point set, rather than just the total number of intervals. Let X be abounded subset of a metric space, and let $\mathbf{x}_\epsilon$ be the centers of a maximal set of disjoint balls of radius $\epsilon/4$ centered at points of X. Also, let

$$I_{i,\epsilon}\left(\left\{x_j\right\}\right) = \left\{(b,d) : PH_i\left(\left\{x_j\right\}\right)\right\} : d-b > \epsilon$$

and

$$f\left(\epsilon\right) = \sup_{\left\{x_j\right\}\subset \mathbf{x}_\epsilon} I_{i,\epsilon}\left(\left\{x_j\right\}\right)$$

Definition 5. *Let*

$$c_i\left(X\right) = \limsup_{\epsilon\to 0} \frac{\log\left(f\left(\epsilon\right)\right)}{\log\left(1/\epsilon\right)}$$

For example, if $\phi^i_X\left(v\right) = O\left(v^d\right)$, then $c_i\left(X\right) \leq d$. Also, as noted by Cohen-Steiner et. al., $\left|I_{i,\epsilon}\right|$ is less than the number of simplices in a triangulation of mesh ϵ of the ambient metric space. It follows that if $X\subset\mathbb{R}^m$, $c_i\left(X\right) \leq \frac{m}{\dim_{\text{box}(X)}}$. [6]

Proposition 4. *Let X be a bounded subset of a triangulable metric space. Then*

$$dim^i_{PH}\left(X\right) \leq c_i\left(X\right)\ dim_{box}\left(X\right)$$

Proof. The proof is nearly identical to that of the previous proposition, except that the number of intervals of length ϵ is bounded by finding a finite subset $\left\{x_j\right\}$ of X with $d_H\left(X,\left\{x_j\right\}\right) < \epsilon/2$ and $d\left(x_j,x_k\right) \geq \epsilon\left(1-\delta_{i,j}\right)$, so every interval of X of length greater than ϵ corresponds to an interval of $\left\{x_j\right\}$ of length greater than $\epsilon/2$. ☐

Corollary 3. *(Cohen-Steiner, Edelsbrunner, Harer, and Mileyko* [6]*) Let X be a bounded subset of $\mathbb{R}^m$. Then*

$$dim^i_{PH}\left(X\right) \leq m$$

Conjecture 2. $c_i\left(X\right) = 1$ *for any bounded subset X of $\mathbb{R}^m$, and $0\leq i\leq m-1$.*

4.1. **An Example with Many Intervals.** We present an example of a subset of $\mathbb{R}^3$ with subsets of size n that appear to have $\approx n^{1.5}$ *PH* intervals. This example was suggested by Herbert Edelsbrunner.

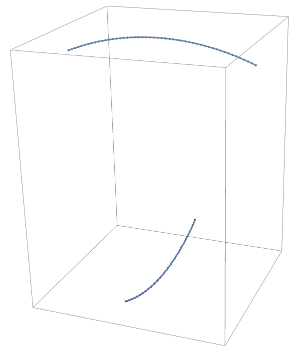
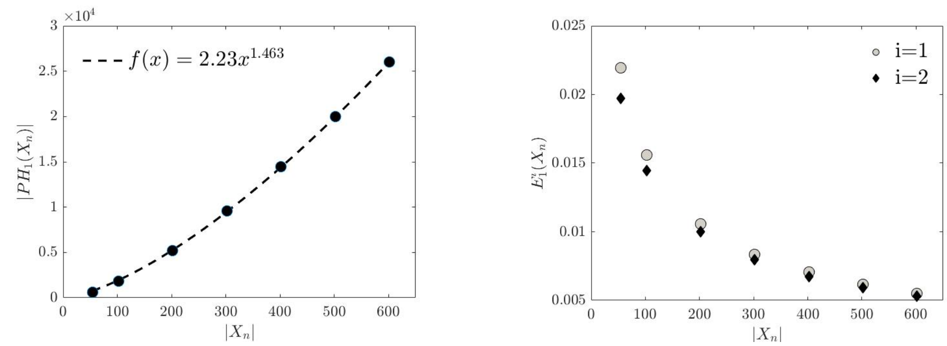


Figure 3. The set X : two opposing arcs from unit circles.



(a) (b)

Figure 4. (a) The number of PH_1 intervals of X_n, which appears to grow approximately as $n^{1.5}$. The data for PH_2 is similar. (b) The quantities $E_1^1(X_n)$ and $E_1^2(X_n)$, which appear to be bounded as a function of n. Persistent homology was computed by calculating the alpha complex with CGAL [19] and passing the resulting filtration to JPLEX [16].

Let C_1 and C_2 be the following arcs from unit circles in $\mathbb{R}^3$:

$$C_1 = \{(\cos(\theta), \sin(\theta), 0) : \theta \in [-\pi/8, \pi/8]\}$$

$$C_2 = \{(1 - \cos(\theta), 0, \sin(\theta)) : \theta \in [-\pi/8, \pi/8]\}$$

and let $X = C_1 \cup C_2$. X is shown in Figure 3. Let $X_n \subset X$ be a point set obtained by placing $\lfloor n/2 \rfloor$ uniformly spaced points on each of the two circular arcs. Computations indicate that $|PH_i(X_n)| \approx |X_n|^{1.5}$ for $i = 1, 2$ (Figure 4(a)), but that $E_1^1(X_n)$ and $E_1^2(X_n)$ are bounded as functions of n (Figure 4(b)). As such, the question remains of whether the PH_i dimension is bounded above by the upper box dimension for any subset of Euclidean space.

5. A Bound for $\dim_{\widetilde{PH}_1}$

We prove an upper bound for the $\tilde{PH}_1$-dimension of a subset of $\mathbb{R}^d$ that does not depend on the ambient Euclidean dimension. Note that the proof of Proposition 3 works for persistent homology defined in terms of either the Čech complex or the Rips complex.

The following argument is a modified version of the proof of Theorem 3.1 in Goff [11], which bounds the (non-persistent) first Betti number of a Rips complex of finitely many points in $\mathbb{R}^m$ in terms of the kissing number K_m of $\mathbb{R}^m$. Let $\mathcal{R}_\epsilon(X)$ be the Rips complex of X at parameter ϵ, let $S_\epsilon(x)$ and $L_\epsilon(x)$ denote the star and link of a vertex of $\mathcal{R}_\epsilon(X)$, respectively:

$$S_\epsilon(x) = \{\sigma \in \mathcal{R}_\epsilon(X) : x \in \sigma\}$$
$$L_\epsilon(x) = \{(z_1, \dots z_k) \in \mathcal{R}_\epsilon(X) : (x, z_1, \dots z_k) \in \mathcal{R}_\epsilon(X)\}$$

Lemma 2. *Let X be a finite subset of $\mathbb{R}^m$, and $x \in X$, and let $\{L_\epsilon(x)\}_{\epsilon\in\mathbb{R}^+}$ be the filtration of the links of x in $\mathcal{R}_\epsilon(X)$. Then*

$$\left|PH_0(\{L_\epsilon(x)\}_{\epsilon\in\mathbb{R}^+})\right| \leq K_m - 1$$

Proof. If necessary, perturb the points of X so all pairwise distances are distinct without reducing the number of PH_0 bars. Let $\{y_1, \dots y_k\}$ be the points of X corresponding to the births of each of the PH_0 intervals of the filtration. That is, $\{y_i\}$ are the points of X so that there is a PH_0-class of the filtration with birth time $d(y_i, x)/2$. Let $\epsilon = d(x, X - x)/2$ and $\hat{y}_i$ be the point on the line segment from x to y_i with $d(x, \hat{y}_i) = 2\epsilon$.

Let y_j and y_k be points in $\{y_i\}$, and assume that $d(y_i, x) > d\left(y_j, x\right)$. Then y_i and y_j are both in $L_{d(y_i,x)}$ but they are in distinct components. Therefore the triangle formed by y_i, y_j, and x enters the Rips filtration at a later value of ϵ and

$$d\left(y_i, y_j\right) > d(y_i, x) > d\left(y_j, x\right)$$

so the angle between the vectors $y_i - x$ and $y_j - x$ is greater than $\pi/3$. It follows that the balls $\{B_\epsilon(\hat{y}_i)\}$ are disjoint, and tangent to $B_\epsilon(x)$. Thus $|\{y_i\}| < K_m - 1$. □

Theorem 7. *If X is a finite subset of $\mathbb{R}^m$ then*

$$\left|\widetilde{PH}_1(X)\right| \leq |X|(K_m - 1)$$

Proof. We will prove this result by induction on the size of X. Let δ be the length of the shortest interval of $\widetilde{PH}_1(X)$. Perturb the points of X by an amount less than $\delta/2$ so that the lengths of all of the pairwise distances between points of X are distinct. The stability of the bottleneck distance implies that no bars of $\widetilde{PH}_1(X)$ are destroyed by this perturbation, so an upper bound for the perturbed set implies one for the original.

Let $\{\epsilon_1 = 0 < \epsilon_2 < \ldots < \epsilon_m\}$ be the values of ϵ at which a simplex is added to the Rips complex of X. For $x \in X$, we have the following commutative diagram of Mayer-Vietoris sequences and inclusion maps:

$$\begin{array}{ccccc}
H_1\left(\mathcal{R}_{\epsilon_{i-1}}(X-x)\right) & \xrightarrow{\alpha_{i-1}} & H_1\left(\mathcal{R}_{\epsilon_{i-1}}(X)\right) & \xrightarrow{\partial_{i-1}} & H_0\left(L_{\epsilon_{i-1}}(x)\right) \\
\downarrow & & \downarrow & & \downarrow \xi_i \\
H_1\left(\mathcal{R}_{\epsilon_i}(X-x)\right) & \xrightarrow{\alpha_i} & H_1\left(\mathcal{R}_{\epsilon_i}(X)\right) & \xrightarrow{\partial_i} & H_0\left(L_{\epsilon_i}(x)\right) \\
\downarrow & & \downarrow & & \downarrow \xi_{i+1} \\
H_1\left(\mathcal{R}_{\epsilon_{i+1}}(X-x)\right) & \xrightarrow{\alpha_{i+1}} & H_1\left(\mathcal{R}_{\epsilon_{i+1}}(X)\right) & \xrightarrow{\partial_{i+1}} & H_0\left(L_{\epsilon_{i+1}}(x)\right)
\end{array}$$

where $L_\epsilon(x)$ is the link of x in $R_\epsilon(x)$. Note that $H_1(S_\epsilon(x)) = 0$ is suppressed in the left column, and that all homology groups are reduced homology groups.

The edge lengths are distinct, and no PH_1 class can be born at the same time a triangle enters the filtration, so at most one PH interval of each of the three columns may be born or die at each row. It follows that

$$\left|\widetilde{PH}_1(X)\right| = \left|\left\{i : \dim H_1\left(\mathcal{R}_{\epsilon_{i+1}}(X)\right) = \dim H_1\left(\mathcal{R}_{\epsilon_i}(X)\right) + 1\right\}\right|$$

Homology is taken with field coefficients so

$$H_1\left(\mathcal{R}_{\epsilon_i}(X)\right) \cong \operatorname{im}\alpha_i \oplus \operatorname{im}\partial_i$$

Therefore, if $H_1\left(\mathcal{R}_{\epsilon_{i+1}}(X)\right) > \dim H_1\left(\mathcal{R}_{\epsilon_i}(X)\right)$ the dimension of either $\operatorname{im}\alpha_i$ or $\operatorname{im}\partial_i$ must also have increased. If the former, commutativity of the diagram implies that

$$\dim H_1\left(\mathcal{R}_{\epsilon_{i+1}}(X-x)\right) = \dim H_1\left(\mathcal{R}_{\epsilon_i}(X-x)\right) + 1$$

so an interval of $\widetilde{PH}_1(X)$ must also be born at time ϵ_{i+1}. Otherwise, commutativity of the diagram implies that there an element of $H_0\left(L_{\epsilon_{i+1}}\right)$ that is not in the image of $\xi_i \circ \xi_{i-1} \ldots \circ \xi_{i-k} \circ \partial_{i-k}$ for any $k \geq 1$ so

$$\left|\{i : \dim \operatorname{im}\partial_{i+1} > \dim\ (\operatorname{im}\partial_i)\}\right| \leq \left|\widetilde{PH}_0(\{L_\epsilon(x)\})\right|$$

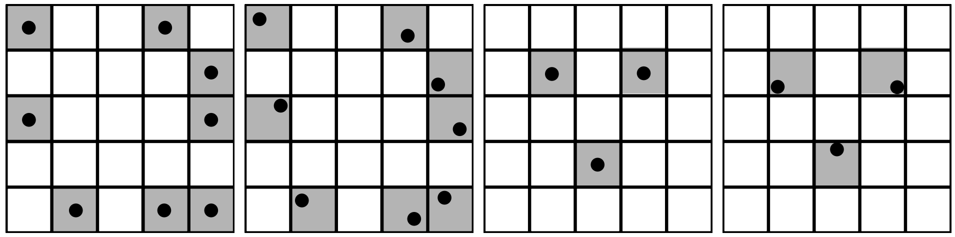


Figure 5. The PH_1 class of the lattice points corresponding to the gray cubes in (a) and (b) is stable — any choice of one point in each cube will result in a set with non-trivial PH_1. The one in (c) and (d) is not.

Therefore,

$$\begin{aligned}\left|\widetilde{PH}_1(X)\right| = & \\ & \left|\{i : \dim \operatorname{im} \partial_i > \dim \operatorname{im} \partial_{i-1}\}\right| + \left|\{i : \dim \operatorname{im} \alpha_i > \dim \dim \alpha_{i-1}\}\right| \\ \leq & \left|\widetilde{PH}_1(X - x)\right| + \left|\widetilde{PH}_0\left(\{L_\epsilon(x)\}_{\epsilon \in \mathbb{R}^+}\right)\right| \\ \leq & \left|\widetilde{PH}_1(X - x)\right| + K_d - 1\end{aligned}$$

and desired statement follows by induction.

□

This, together with Proposition 3, implies Theorem 4.

6. A Lower Bound

Let $[N]$ be the integers $1, 2, \ldots n$ and let $[N]^m \subset \mathbb{Z}^m$ be $[N] \times [N] \times \ldots \times [N]$. Also, let the **cube corresponding to** a point $z \in \mathbb{Z}^m$ be the closed Voronoi cell of $z \in \mathbb{Z}^m$ in $\mathbb{R}^m$. The **grid of mesh** δ in $\mathbb{R}^m$ is the collection of cubes of $\mathbb{Z}^m$ rescaled by a factor δ.

If X is a bounded subset of $\mathbb{R}^m$, the upper box dimension of X controls the number of cubes in the grid of mesh δ that X intersects. Our strategy to prove a lower bound for the PH-dimension of X in terms of the upper box dimension is to show that a sufficiently large collection of points, each in a distinct cube of $[N]^m$, must contain a subset with non-trivial persistent homology.

Definition 6. *$X \subset \mathbb{Z}^m$ has a* ***stable*** *PH_i-class if any point collection Y consisting of exactly one point in each cube corresponding to a point X has a PH_i interval (b, d) so that $d - b > c$ for some constant $c > 0$ (see Figure 5). The supremal such c is called the* ***size*** *of the stable persistence class.*

The stability of the bottleneck distance immediately implies the following:

Proposition 5. *If $X \subset \mathbb{Z}^m$ and there is an interval $(b, d) \in PH_i(X)$ with*

$$d - b > \sqrt{m}$$

then X has a stable PH_i-class.

.

Definition 7. *Let $\xi_i^m(N)$ be the size of the largest subset X of $[N]^m$ so that no subset Y of X has a stable PH_i-class of size greater than 1.*

Proposition 6. *Suppose $\xi_i^m(N) \in O(N^\gamma)$ for some $\gamma > 0$. If X is a bounded subset of $\mathbb{R}^m$ so that and $dim_{box}(X) > \gamma$ then*

$$dim_{PH}^1(X) \geq dim_{box}(X)$$

Proof. Let $d = \dim_{box}(X)$ and $\beta < d$. We will find finite point collections $\{z_j\} \subset X$ so that $E_\beta^i\left(\{z_j\}\right)$ is arbitrarily large. Assume that $d < m$; we will deal with the case $d = m$ separately.

Let $\lambda(N) = \xi_i^m(N) / (N^m)$, and let D be a positive constant so $\lambda(N) < DN^{\gamma - m}$ for all $N \in \mathbb{N}$. For $y \in \mathbb{R}^+$, let B_y be the collection of cubes in the grid of mesh $1/y$ in $\mathbb{R}^m$ which X intersects. Find a positive constant c that is less than $\frac{d-\beta}{m-d}$, and choose $\epsilon > 0$ small enough so

$$\frac{d - \beta - \epsilon/2}{m - d + \epsilon/2} > c > \frac{\epsilon}{\alpha - \gamma - \epsilon/2}$$

Let $\alpha = d + \epsilon/2$. There is a real number Y so that for all $y > Y$,

$$\left|B_y\right| < y^{\alpha - \epsilon}$$

Let $\{y_j\}$ be an increasing sequence of positive integers tending toward ∞ so that $y_j^c > Y + 1$ and

$$y_j^\alpha < \left|B_{y_j}\right| < y_j^{\alpha - \epsilon}$$

Let $k_j = \lfloor y_j^{c/(1+c)} \rfloor$ and $n_j = y_j/k_j$. Note that $n_j \to \infty$ as $y_j \to \infty$ because $c > 0$. The grid of mesh $1/y_j$ in $\mathbb{R}^m$ is obtained by dividing each cube in the grid of mesh $1/n_j$ into k_j^m cubes. Furthermore, each cube in B_{y_j} is contained in exactly one cube of B_{n_j}. A subdivided cube of B_{n_j} can contain at most k_j^m cubes of B_{n_j}, so the number of cubes of B_{n_j} in which the density of B_{y_j} is greater than $\lambda\left(k_j\right)$ is greater than or equal to the smallest integer a satisfying

$$k_j^m a + \left(\lceil \lambda\left(k_j\right) k_j^m \rceil - 1\right)\left(\left|B_{n_j}\right| - a\right) \geq \left|B_{y_j}\right|$$

Rearranging terms, we have

$$\begin{aligned} a &\geq \frac{\left|B_{y_j}\right| - \left|B_{n_j}\right|\left(\lceil \lambda\left(k_j\right) k_j^m \rceil - 1\right)}{k_j^m - \left(\lceil \lambda\left(k_j\right) k_j^m \rceil - 1\right)} \\ &\geq \frac{\left(n_j k_j\right)^{\alpha-\epsilon} - n_j^{\alpha} \lambda\left(k_j\right) k_j^m}{k_j^m - \lambda\left(k_j\right) k_j^m - 1} \\ &\geq n_j^{\alpha} \frac{k_j^{(\alpha-\epsilon-m)} n_j^{-\epsilon} - D k_j^{\gamma-m}}{1 - D k_j^{\gamma-m} - k_j^{-m}} \\ &\approx n_j^{\alpha - c(m-\alpha+\epsilon)-\epsilon} \end{aligned}$$

where we have used that $c > \frac{\epsilon}{\alpha-\gamma-\epsilon}$ and $\gamma - m < 0$. Let $\phi = \alpha - c\left(m - \alpha + \epsilon\right) - \epsilon$.

In each of cube of B_{n_j} for which the density of B_{y_j} exceeds $\lambda\left(k_j\right)$ we may choose a finite set of points $\{x_t\} \subset X$ so that $PH_i(x_t) \neq 0$. There are at least $\lfloor n_j^{\phi} \rfloor$ such cubes of B_{n_j}; let $\{z_l\}_j$ be the union of the point sets chosen for each of these cube. To separate the contributions of each PH_i-class to the persistent homology of X, let $\{z_l\}_j$ be a maximal subset of $\{z_l\}_j$ so that any two occupied cubes of B_{y_j} are separated by at least two empty cubes. Then there is a constant F depending only on the ambient dimension m so that $PH_i\left(\{z_l\}_j\right)$ contains at least $\frac{1}{F}\lfloor n_j^{\phi} \rfloor$ intervals, each of which has birth minus death equal to at least $\frac{1}{n_j}$.

Then, for sufficiently large n_j,

$$E_{\beta}^{i}\left(\{z_l\}_j\right) \geq \frac{1}{F}\lfloor {n_j}^{\phi} \rfloor \left(\frac{1}{n_j}\right)^{\beta} \approx \frac{1}{F} {n_j}^{\phi-\beta}$$

which is unbounded as a function as $n_j \to \infty$ because

$$\begin{aligned}\phi - \beta = & \\ & \alpha - c(m - \alpha + \epsilon) - \epsilon - \beta \\ & > d - \frac{\epsilon}{2} + \frac{d - \beta - \epsilon/2}{m - d + \epsilon/2}(m - d + \epsilon/2) - \beta \\ & = 0\end{aligned}$$

so $\dim_{PH}^{i}(X) \geq \beta$ for all $\beta < d$, and $\dim_{PH}^{i}(X) \geq d$ as desired. □

This leads to the following question:

Question 2. *What is the infimal γ so that $\xi_i^m(N) \in O(N^\gamma)$? That is, what is*

$$\limsup_{N\to\infty} \frac{\log(\xi_i^m(N))}{\log(N)}$$

The corresponding question for stable persistent homology classes of the Rips complex is also of interest.

6.1. **A Bound for ξ_1^m.** We prove an upper bound for ξ_1^m by considering the contribution of small triangles to the persistent homology of a subset of $\mathbb{Z}^m$. Three points in Euclidean space give rise to a PH_1-class if and only if they are the vertices of an acute triangle, whose total persistence (death minus birth) equals the circumradius minus half the length of its longest edge.

Lemma 3.

$$\xi_1^m(N) \in O\left(N^{m-1/2}\right)$$

Proof. Let $X \subset [N]^m$ and $c > \sqrt{2\sqrt{N} + 4}$, and suppose $|X| > 8cN^{m-1/2} + 4^{m-1}$. We will find three points of X that give rise to a stable PH_1-class whose size is greater than one if N is sufficiently large. First, we reduce the problem to a two-dimensional one. If $m > 2$, we may find a two-dimensional slice S of $[N]^m$ of the form $[N]^2 \times w$, so that the density of X in S is greater than or equal to the density of X in $[N]^m$. That is, $X \cap S$ has more than $8cN^{1.5} + 4$ elements. We will treat S as a two-dimensional grid of n-dimensional cubes, with N rows and N columns.

We may remove $4c\sqrt{N} + 4$ elements of $X \cap S$ from each row and column without exhausting $X \cap S$, so there is an $y \in X \cap S$ so that the row of $[N]^2$ containing x contains $2c\sqrt{N} + 1$ additional elements of $X \cap S$ both to the left and to the right of x. Similarly, y may also be chosen so that there are $2c\sqrt{N} + 1$ elements of $X \cap S$

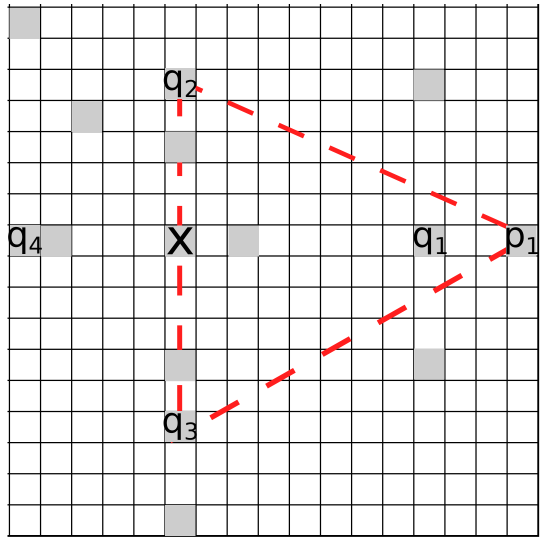


Figure 6. The setup in Lemma 3. The points of X are shown in gray, and the triangle formed by the points p_1, q_2, and q_3 is shown by the red dashed lines. It is an acute triangle, and the q_1, q_2, and q_3 is either acute or right

above and below it in the same column. Let q_1, q_2, q_3 and q_4 be elements of X to the right, above, below, and to the left of x so that there $c\sqrt{N}$ other elements of $X \cap S$ between them and x, and at least $c\sqrt{N}$ additional elements between them and the boundary of $[N]^2$. Translate $[N]^2$ so that y is located at $(0,0)$ (ignoring the higher-dimensional coordinates). q_1, q_2, q_3, and q_4 form a quadrilateral, so there is a non-obtuse angle at at least one of the vertices. Without loss of generality, suppose the angle at q_1 is less than $\pi/2$. Let $q_1 = (\hat{x}, 0)$, $q_2 = (0, y_1)$, and $q_3 = (0, y_2)$, where $\hat{x}$ and y_1 are positive and y_2 is negative. Additionally, let $p_1 = (x)$ be an element of $X \cap S$ in the same row as y and q_1 that is at least $c\sqrt{N}$ cubes further to the right of q_1. Then p_1, q_1, and q_2 form an acute triangle.

Our setup is illustrated in Figure 6. p_1, q_1, and q_2 form an acute triangle, so $PH_1(\{p_1, q_1, q_2\})$ consists of a single interval. We will show that the length of this interval is greater than $1 + \frac{\sqrt{N}}{2}$ for sufficiently large N, which implies that the PH_1 class is a stable class of size greater than one. The length of this interval is equal to the circumradius of the triangle minus half the length of its maximum edge. A quick computation shows that the circumradius equals

$$\frac{\sqrt{\left(x^2 + y_1^2\right)\left(x^2 + y_2^2\right)}}{2x}$$

We consider three cases. In the first case, the edge from q_1 to q_2 is longest. In this case, the total persistence equals

$$TP_1(x, y_1, y_2) = \frac{\sqrt{\left(x^2+y_1^2\right)\left(x^2+y_2^2\right)}}{2x} - \frac{y_1 - y_2}{2}$$

q_1, q_2, q_3 is an acute triangle, so $\hat{x} > -y_1 y_2$ and $x > \sqrt{-y_1 y_2} + c\sqrt{N}$. A tedious computation shows that, subject to the constraints that $c\sqrt{N} \le y_1 \le n$, $-N \le y_2 \le -c\sqrt{N}$, and $x > \sqrt{-y_1 y_2} + c\sqrt{N}$, TP_1 is minimized when $x = c\sqrt{N} + N$, $y_1 = N$, and $y_2 = -N$ for sufficiently large N. Therefore

$$TP_1(x, y_1, y_2) \ge TP_1\left(c\sqrt{N} + N, N, N\right) = \frac{\sqrt{\left(\left(c\sqrt{N}+\sqrt{N^2}\right)^2 + N^2\right)^2}}{2\left(c\sqrt{N}+\sqrt{N^2}\right)} - N$$

which is an increasing function of N whose limit is $\frac{1}{2}c^2$ as N goes to ∞. As $c > \sqrt{2\sqrt{N}+4} > \sqrt{2\sqrt{N}+2}$, the total persistence will be greater than $1 + \frac{\sqrt{N}}{2}$ for sufficiently large N, as desired.

In the second case, the edge between p and q_1 is longest. In this case, the total persistence equals

$$TP_2(x, y_1, y_2) = \frac{\sqrt{\left(x^2+y_1^2\right)\left(x^2+y_2^2\right)}}{2x} - \frac{\sqrt{x^2+y_1^2}}{2}$$

which is minimized when x is as large as possible and the magnitudes of y_1 and y_2 are as small as possible. Thus

$$TP_2(x, y_1, y_2) \ge TP_2\left(N, c\sqrt{N}, -c\sqrt{N}\right) = \frac{1}{2}\left(c^2 + N - \sqrt{c^2 N + N^2}\right)$$

which is a decreasing function of N with limiting value $\frac{1}{4}c^2$. Because $c > \sqrt{2\sqrt{N}+4}$, the total persistence is always greater than $1 + \frac{\sqrt{N}}{2}$, as desired. The argument in the third case is identical to this one. □

The previous result, together with Proposition 6 implies Theorem 3 and completes the proof of Theorem 2.

7. Conclusion

We have taken the first steps toward answering Question 1, which asks for hypotheses under which the PH_i-dimension of a bounded subset of $\mathbb{R}^m$ equals its upper box dimension. However, the question remains open for all cases with $m > 2$ and $i > 0$. We suspect that even the $n = 2, i = 1$ case can be improved, to include sets whose upper box dimension is between 1 and 1.5.

Another interesting question is whether similar results can be shown for a probabilistic version of the PH_i-dimension. That is, if $\{x_j\}$ is a finite point collection drawn from a probability measure on $\mathbb{R}^m$, can the expectation of $E_i\left(\{x_j\}\right)$ be controlled in terms a classically defined fractal dimension for probability measures? This question would perhaps be more interesting for applications, which usually deal with random point collections rather than extremal ones. However, proving a lower bound is already difficult in the extremal case.

8. Acknowledgments

The author would like to thank Henry Adams, Tamal Dey, Herbert Edelsbrunner, Matthew Kahle, Chris Peterson, and Shmuel Weinberger for useful discussions. In particular, he would like to thank Herbert Edelsbrunner for suggesting the example in Section 4.1 and the reference [6]. He would also like to thank the members of the Pattern Analysis Laboratory at Colorado State University, whose computational experiments involving a similar dimension inspired the ideas in this paper [1].

Funding for this research was provided by a NSF Mathematical Sciences Postdoctoral Research Fellowship under award number DMS-1606259.

Appendix A. Properties of $\dim^i_{\widehat{PH}}$

Recall that we defined an another notion of dimension for a subset of a triangulable metric space, based on the persistent homology of the subset itself rather than of finite point sets contained within it:

Definition 8. *The i-th dimensional persistent homology dimension of bounded subset of a triangulable metric space is*

$$dim^i_{\widehat{PH}}\left(X\right) = \inf\left\{\beta : E^i_\beta\left(X\right) < \infty\right\}$$

This definition is equivalent to the one in our previous paper, "Measuring Shape with Topology" [14]. Let X be a bounded subset of a metric space, and let

$$F_i\left(x\right) = \left|\left\{(b,d) \in PH_i\left(X\right) : d-b > x\right\}\right|$$

Define

$$\dim_{F_i}\left(X\right) = \sup\left\{c : \lim_{x\to 0} x^c F_i\left(x\right) = \infty\right\}$$

Proposition 7. *Let X be a bounded subset of a triangulable metric space. Then*

$$dim^i_{\widehat{PH}}\left(X\right) = dim_{F_i}\left(X\right)$$

The proposition is an immediate consequence of the following lemma.

Lemma 4. *Suppose $Y \subset \mathbb{R}^+$ has the property that $\{y \in Y : y > \epsilon\}$ is finite for all $\epsilon > 0$. If*

$$F_Y\left(x\right) = \left|\left\{y \in Y : y > x\right\}\right|$$

then

$$\inf\left\{\alpha : \sum_{y\in Y} y^\alpha < \infty\right\} = \sup\left\{c : \lim_{x\to 0} x^c F_Y\left(x\right) = \infty\right\}$$

Proof. Let $d_1 = \inf\left\{\alpha : \sum_{y\in Y} y^\alpha < \infty\right\}$ and $d_2 = \sup\left\{c : \lim_{x\to 0} x^c F_Y\left(x\right) = \infty\right\}$.

Let $\alpha > \beta > d_2$, $I_k = \left\{y \in Y : 2^{-(k+1)} < y \leq 2^{-k}\right\}$, and $C_0 = \sum_{\{x\in X : x>1\}} x^\alpha$. By definition, there is a constant $C_1 > 0$ so that $F(x) < C_1 x^{-\beta}$ for all $x \in \mathbb{R}$. Then

$$\begin{aligned}
\sum_{y\in Y} y^\alpha \leq & \\
& \sum_{k=1}^{\infty} \left|I_k\right| 2^{-\alpha k} + C_0 \\
< & \sum_{k=1}^{\infty} F_Y\left(2^{-(k+1)}\right) 2^{-\alpha k} + C_0 \\
< & \sum_{k=1}^{\infty} C_1 2^{\beta(k+1)} 2^{-\alpha k} + C_0 \\
= & C_1 2^\beta \sum_{k=1}^{\infty} 2^{(\beta-\alpha)k} + C_0 \\
< & \infty
\end{aligned}$$

because $\beta - \alpha < 0$. Therefore, $d_1 \leq d_2$.

Let $\alpha < d_2$ and $D \in \mathbb{R}$. $\lim_{x\to 0} x^\alpha F_Y(x) = \infty$ so there is an $\hat{x} > 0$ so that $\hat{x}^\alpha F_Y(\hat{x}) > D$. Then

$$\sum_{y\in Y} y^\alpha \geq \sum_{y>\hat{x}} y^\alpha > \hat{x}^\alpha F_Y(\hat{x}) > D$$

so $\sum_{y\in Y} y^\alpha = \infty$ and $d_1 \geq d_2$, as desired. □

We also show that $\dim^i_{\widehat{PH}}$ is bounded above by $\dim^i_{PH}$:

Proposition 8. *Let X be a bounded subset of a triangulable metric space. Then*

$$dim^i_{\widehat{PH}}(X) \leq dim^i_{PH}(X)$$

Proof. Let $\alpha > \dim^i_{\widehat{PH}}(X)$ and $D > 0$. There is a finite collection of intervals $\{b_j, d_j\}_{j=1}^n \subset PH_i(X)$ with

$$\sum_{j=1}^n (d_j - b_j)^\alpha > D + 1$$

The function $f(x) = x^\alpha$ is uniformly continuous on $\left[\min_j (d_j - b_j), \max_j (d_j - b_j)\right]$ so there is a $\epsilon > 0$ so that

$$\left|x^\alpha - (x+\delta)^\alpha\right| < \frac{1}{n}$$

for all $x \in \left[\min_j (d_j - b_j), \max_j (d_j - b_j)\right]$ and $|\delta| < \epsilon$. Let $\{x_k\} \subset X$ be a finite point satisfying $d_H(\{x_k\}, X) < \epsilon$. By Corollary 1, the intervals $\left\{(b_j, d_j)\right\}$ can be paired with intervals $\left\{\left(\hat{b}_j, \hat{d}_j\right)\right\} \in PH_i(\{x_k\})$ so that

$$\left|(d_j - b_j) - \left(\hat{d}_j - \hat{b}_j\right)\right| < \epsilon$$

Then

$$
\begin{aligned}
E_\alpha(\{x_k\}) = & \\
& \sum_{(\hat{b},\hat{d})\in PH_i(\{x_k\})} \left(\hat{d}-\hat{b}\right)^\alpha \\
\geq & \sum_{j=1}^{n} \left(\hat{d}_j-\hat{b}_j\right)^\alpha \\
\geq & \sum_{j=1}^{n} \left(d_j-b_j\right)^\alpha - \frac{1}{n} \\
> & D
\end{aligned}
$$

D and α were arbitrary, so $\dim_{PH}^{i}(X) \geq \dim_{\widehat{PH}}^{i}(X)$, as desired. □

References

[1] The Pattern Analysis Lab at Colorado State University. A fractal dimension for measures via persistent homology. Preprint, 2018.

[2] Francesca Cagliaria and Claudia Landibc. Finiteness of rank invariants of multidimensional persistent homology groups. *Applied Mathematics Letters*, 2011.

[3] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. *The Structure and Stability of Persistence Modules*. Springer, 2016.

[4] Frdric Chazal, Vin de Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 2014.

[5] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103120, 2007.

[6] David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko. Lipschitz functions have l_p-stable persistence. *Foundations of Computational Mathematics*, 2010.

[7] William Crawley-Boevey. Decomposition of pointwise finite-dimensional persistence modules. *Journal of Algebra and its Applications*, 2015.

[8] Vin De Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7(1):339358, 2007.

[9] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplificitaion. *Discrete and Computational Geometry*, 2002.

[10] Kenneth Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley, 2014.

[11] Michael Goff. Extremal Betti numbers of Vietoris-Rips complexes. *Discrete and Computational Geometry*, 2011.

[12] Harry Kesten and Sungchul Lee. The central limit theorem for weighted minimal spanning trees on random points. *Annals of Applied Probability*, 1996.

[13] Gady Kozma, Zvi Lotker, and Gideon Stupp. The minimal spanning tree and the upper box dimension. *Proceedings of the American Mathematical Society*, 2006.

[14] R. D. MacPherson and B. Schweinhart. Measuring shape with topology. *Journal of Mathematical Physics*, 53(7), 2012.
[15] Vanessa Robins. *Computational topology at multiple resolutions: foundations and applications to fractals and dynamics*. PhD thesis, University of Colorado at Boulder, 2000.
[16] Harlan Sexton and Mikael Vejdemo-Johansson. jPlex, December 2008. http://www.math.colostate.edu/ adams/jplex/index.html.
[17] Richard Stalney. The upper bound conjecture and Cohen-Macaulay rings. *Studies in Applied Mathematics*, 1975.
[18] J. Michael Steele. Growth rates of Euclidean minimal spanning trees with power weighted edges. *Annals of Probability*, 1988.
[19] The CGAL Project. *CGAL User and Reference Manual*. CGAL Editorial Board, 4.11 edition, 2017.
[20] Leopold Vietoris. ber den hheren zusammenhang kompakter rume und eine klasse von zusammenhangstreuen abbildungen. *Mathematische Annalen*, 97(1):454472, 1927.
[21] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete and Computational Geometry*, 33:249–274, 2005.