ϵ-net Induced Lazy Witness Complex for Efficient Topological Data Analysis

Naheed Anjum Arafat

Post-doctoral Research Fellow,
Nanyang Technological University[1]

May 25, 2022

[1] This work was done at School of Computing, National University of Singapore (NUS)

What is this Talk About?

Can we construct an approximate representation which

- is a good approximation of the underlying space,

What is this Talk About?

Can we construct an approximate representation which

- is a good approximation of the underlying space,
- can be constructed not only for point clouds but also non-euclidean datasets such as graphs,

What is this Talk About?

Can we construct an approximate representation which

- is a good approximation of the underlying space,
- can be constructed not only for point clouds but also non-euclidean datasets such as graphs,
- maintains a bounded approximation ratio for persistence homology computation, and

What is this Talk About?

Can we construct an approximate representation which

- is a good approximation of the underlying space,
- can be constructed not only for point clouds but also non-euclidean datasets such as graphs,
- maintains a bounded approximation ratio for persistence homology computation, and
- helps compute persistent homology faster?

What is this Talk About?

Can we construct an approximate representation which

- is a good approximation of the underlying space,
- can be constructed not only for point clouds but also non-euclidean datasets such as graphs,
- maintains a bounded approximation ratio for persistence homology computation, and
- helps compute persistent homology faster?

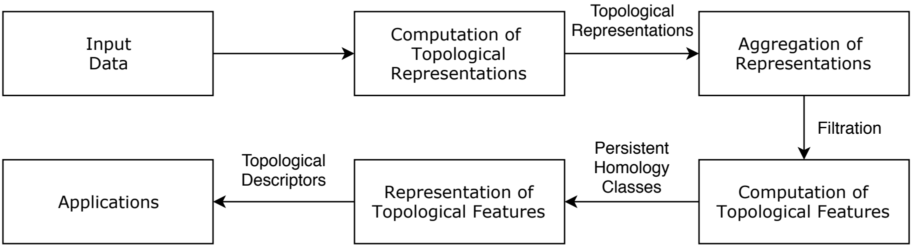
YES

Outline

- Background on Topological data analysis (TDA)
- Computational Issues in TDA and approximate simplicial representations
- ϵ-net: Old approach in new application.
- Our proposal: ϵ-net induced lazy witness complex and approximation guarantees.
- Our algorithms.
- Questions.

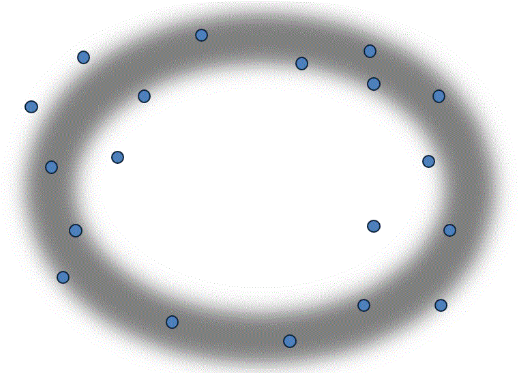
Background

Topological data analysis: the framework



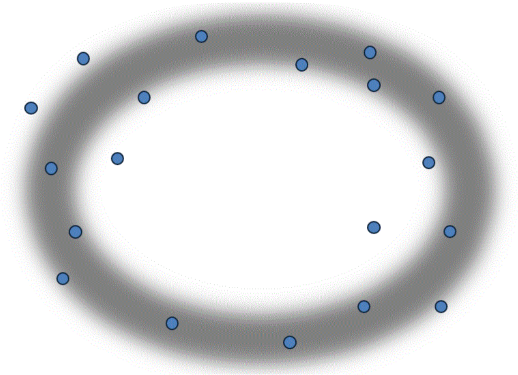
Topological features

The underlying space: A Ring in $\mathbb{R}^2$



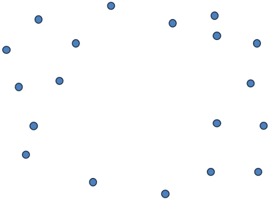
Topological features

The underlying space: A Ring in $\mathbb{R}^2$

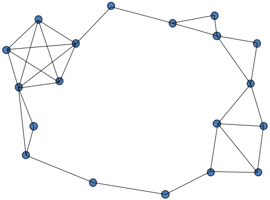


Top. feat. at dim. 0: Connected components (1)
Top. feat. at dim. 1: cycles, Inner-cycle $\sim$ Outer-cycle (1)
Top. feat. at dim. 2: voids (0)
(Formal) Top. feat. at dim. k : Generators of the dim. k homology group.

Input Data: Point cloud, graph



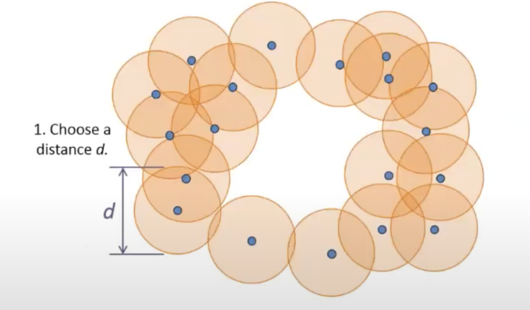
Point cloud



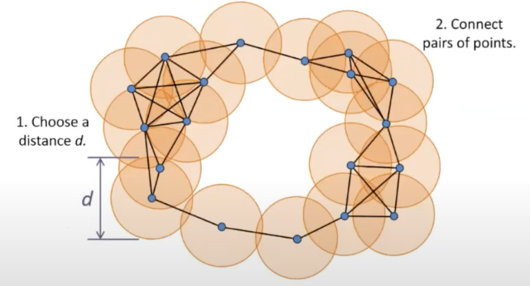
Graph[2]

[2] simple, connected, unweighted, undirected graph throughout this talk.

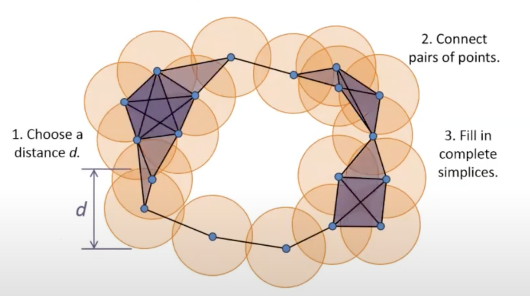
Representation: Simplicial complex at an offset



Representation: Simplicial complex at an offset



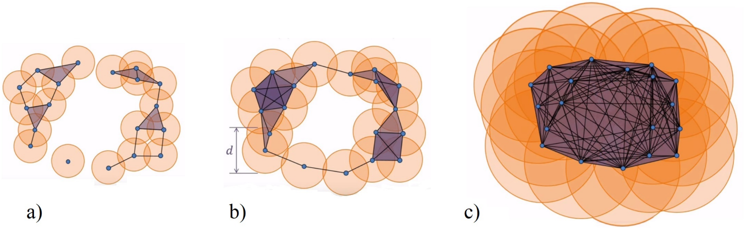
Representation: Simplicial complex at an offset



Simplicial representation at offset $F_\alpha \triangleq \{\sigma(\cup_{x \in P} B_{\alpha/2}(x))\}$

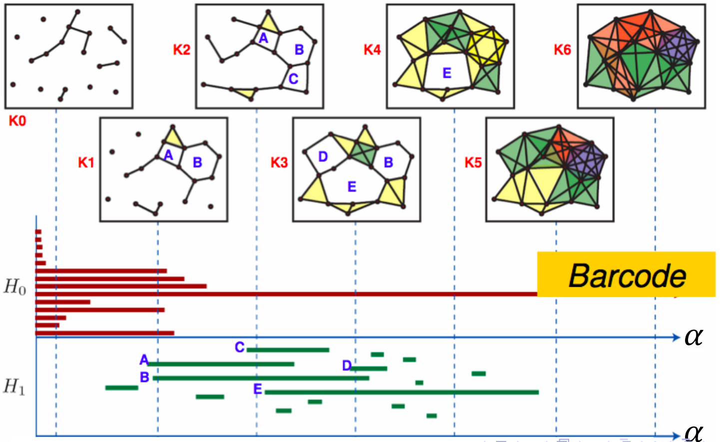
Aggregation of Representations: Filtration

A filtration is a sequence of simplicial representations $\mathcal{F} \triangleq (F_{\alpha_1}, F_{\alpha_2}, \cdots,)$ where $\alpha_1 \leq \alpha_2 \leq \cdots$.



Topological Features and Persistence

- As higher-order simplices appear (born) in the filtration,
  - new cycles may appear (new homology classes being born)
  - some cycles may vanish by becoming boundaries of higher-order simplices. (existing homology classes being merged with others)
- The birth and deaths of homology classes are represented in different ways: barcode, persistence diagrams etc.



TDA pipeline: Summary

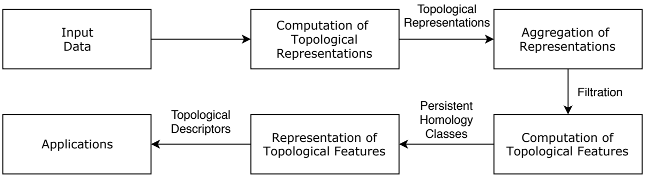


Figure 1: TDA pipeline

- *Input Data:* Point cloud, Graph.
- *Topological Representations:-* Simplicial complex at offset. Mathematically: A vector space
- *Filtration:-* Sequence of offset simplicial complexes. Mathematically: A sequence of vector spaces with canonical inclusion map defined by boundary operator.
- *Persistent Homology Class:-* Inclusion maps induces linear maps in the homology vector spaces in the sequence. The image of these maps charecterizes the persistence (birth,death) of the homology classes in the sequence.
- *Topological descriptors:* Multiset of points in $\mathbb{R}^2$ called persistence diagrams.

Computational Issues with simplicial Representations and approximate representations.

Approximate Simplicial Representations

- **Čech complex** captures the actual topology of the underlying space.
  - *Issue:* computationally challenging $\rightarrow$ at most $O(n^k)$ simplices of dim. up to k.

- **Vietoris-Rips Complex** (R_α) for a given offset α considers a simplex σ to be in R_α if the distance between every pair of points in σ is at most $\frac{\alpha}{2}$.
  - *Good news:* 2-approximation of the Čech complex in arbitrary metric space.
  - *Bad news:* at most $O(n^k)$ simplices of dimension up to k.

Computational bottleneck

Need to enumerate large number of simplices.

Approximate Simplicial Representations

- **Čech complex** captures the actual topology of the underlying space.
  - *Issue:* computionally challenging $\rightarrow$ at most $O(n^k)$ simplices of dim. up to k.

- **Vietoris-Rips Complex** (R_α) for a given offset α considers a simplex σ to be in R_α if the distance between every pair of points in σ is at most $\frac{\alpha}{2}$.
  - *Good news:* 2-approximation of the Čech complex in arbitrary metric space.
  - *Bad news:* at most $O(n^k)$ simplices of dimension up to k.

Computational bottleneck

Need to enumerate large number of simplices.

Research Question

Can we have approximate simplicial representations which are computable in reasonable time, yet good approximations to Vietoris-Rips or Čech complex?

A Computationally Faster Approximation: Lazy Witness Complex and The Question to Solve

Lazy witness Complex $LW_\alpha(P, L)$

Lazy witness Complex $LW_\alpha(P, L)$ of a point-cloud P is a simplicial complex over a landmark set L containing simplices σ such that $\forall v_i, v_j \in \sigma$, $\exists w \in P$ with the following property:

$$\max\{d(w, v_i), d(w, v_j)\} \leq \alpha + d(w, L)$$

here, $d(w, L)$ is the distance from w to its closest point in L.

- *Good news:* Vietoris-Rips complex on landmarks is Lazy witness complex. The number of simplices in Lazy witness complex is $O(|L|^k) << O(|P|^k)$ for $|L| << |P|$
- *Bad news:* There is no approximation guarantee available for $|L| \neq |P|$. There is no algorithm that selects L with any approximation guarantee.

A Computationally Faster Approximation: Lazy Witness Complex and The Question to Solve

Lazy witness Complex $LW_\alpha(P, L)$

Lazy witness Complex $LW_\alpha(P, L)$ of a point-cloud P is a simplicial complex over a landmark set L containing simplices σ such that $\forall v_i, v_j \in \sigma$, $\exists w \in P$ with the following property:

$$\max\{d(w, v_i), d(w, v_j)\} \leq \alpha + d(w, L)$$

here, $d(w, L)$ is the distance from w to its closest point in L.

- *Good news:* Vietoris-Rips complex on landmarks is Lazy witness complex. The number of simplices in Lazy witness complex is $O(|L|^k) << O(|P|^k)$ for $|L| << |P|$
- *Bad news:* There is no approximation guarantee available for $|L| \neq |P|$. There is no algorithm that selects L with any approximation guarantee.

New Questions

How to select the landmarks?
How good are the landmarks selected by an algorithm?
Can we obtain any approximation guarantee for the lazy witness complex?

Our Contributions

The Central Concept

We respond to all these questions by reincarnating the idea of ϵ-net in TDA.

Q. Can we obtain any approximation guarantee for the lazy witness complex?

-> **Lazy witness complex induced by an ϵ-net is a 3-approximation to the Vietoris-Rips complex (point cloud and connected unweighted graph).**

Q. How good are the landmarks selected by an algorithm?

-> **ϵ-net is an ϵ-approximation of the point cloud and graph vertices in Hausdorff distance.**

Q. How many landmarks are there in an ϵ-net?

-> **For a connected unweighted graph of diameter Δ, there exists an ϵ-net of size at most $(\frac{\Delta}{\epsilon})^{O(\log \frac{|V|}{\epsilon})}$.** (For point cloud it is known[3] to be $(\frac{\Delta}{\epsilon})^{\Theta(D)}$)

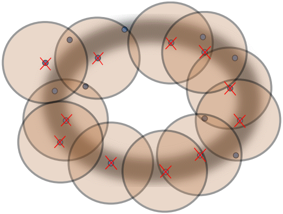
Q. How to select the landmarks?

-> **polynomial-time algorithms to construct ϵ-net on point cloud and on graphs.**

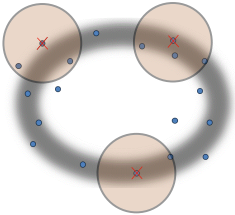
[3] Robert Krauthgamer and James R Lee. "Navigating nets: simple algorithms for proximity search". In: *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms.* 2004.

ϵ-net: old approach in new application

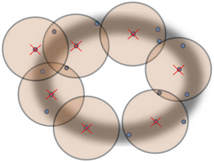
ϵ-net for point cloud



ϵ-sample (Each blue point is within ϵ of some red point)



ϵ-sparse (Each pair of red points are ϵ-far from each other



ϵ-net

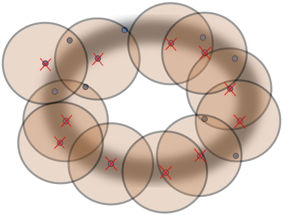
Definition (ϵ-sample)

A set $L \subseteq P$ is an ϵ-sample of P if the collection $\{B_\epsilon(x) : x \in L\}$ of ϵ-balls of radius ϵ covers P , i.e. $P = \cup_{x \in L} B_\epsilon(x)$.

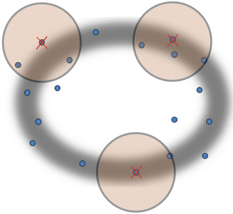
Definition (ϵ-sparse)

A set $L \subset P$ is ϵ-sparse if for all $x, y \in L$, $d(x, y) > \epsilon$.

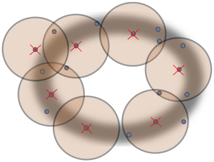
ϵ-net for point cloud



ϵ-sample (Each blue point is within ϵ of some red point)



ϵ-sparse (Each pair of red points are ϵ-far from each other

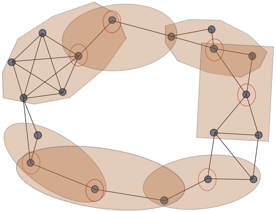


ϵ-net

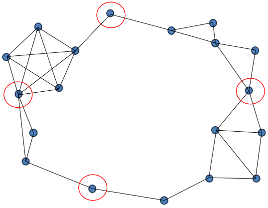
ϵ-net

A subset L of points which is ϵ-sparse and ϵ-sample of point cloud P.

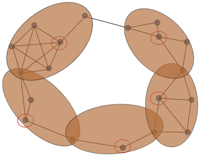
ϵ-net for undirected graphs



ϵ-sample



ϵ-sparse



ϵ-net

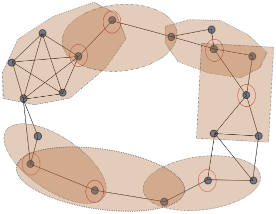
Definition (ϵ-sample)

A set $L = \{u_1, u_2, \ldots, u_{|L|}\} \subseteq V$ is an ϵ-sample of graph $G = (V, d_G)$ if the collection $\{\mathcal{N}_\epsilon(u_i) : u_i \in L\}$ of ϵ-neighbourhoods covers G i.e. $\cup_i \mathcal{N}_\epsilon(u_i) = V$.

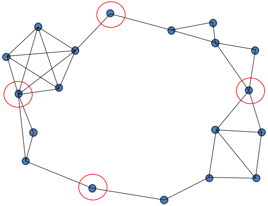
Definition (ϵ-sparse)

A set $L = \{u_1, u_2, \ldots, u_{|L|}\} \subset V$ is ϵ-sparse if for any distinct $u_i, u_j \in L$, $d_G(u_i, u_j) > \epsilon$ in graph G.

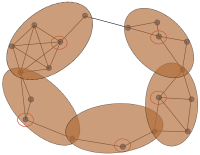
A graph $G = (V, E)$



ϵ-sample



ϵ-sparse



ϵ-net

Definition (ϵ-net)

A subset $L \subset V$ of vertices which is ϵ-sparse and an ϵ-sample of graph $G = (V, d_G)$.

ϵ-net induced Lazy witness complex: Theoretical guarantees

The meaning of approximation guarantee

Suppose $\mathcal{F} = (F_\alpha)_{\alpha>0}$ and $\mathcal{G} = (G_\alpha)_{\alpha>0}$ be two filtrations and their associated persistence diagrams at dimension k be $Dgm(\mathcal{F})$ and $Dgm(\mathcal{G})$. *Approximation is defined in terms of distance between persistence diagrams.*

Definition (Bottlneck distance between diagrams)

$$d_B(Dgm(\mathcal{F}), Dgm(\mathcal{G})) \triangleq \inf_{\phi: Dgm(\mathcal{F}) \to Dgm(\mathcal{G})} \sup_{x \in Dgm(\mathcal{F})} \|x - \phi(x)\|_\infty$$

where ϕ is a bijection between points in the diagrams $Dgm(\mathcal{F})$ and $Dgm(\mathcal{G})$

Definition (δ-approximation between diagrams)

Let $(\log \mathcal{F})_A$ be the re-parameterisation of $(\mathcal{F})_A$ on the natural logarithm scale:

$$\log \mathcal{F}_A \triangleq \{\mathcal{F}_{e^\alpha}\}, \text{ for any } \alpha \in A$$

A persistence diagram $Dgm(\mathcal{F}_A)$ is defined to be δ-approximation to diagram $Dgm(\mathcal{G}_B)$ if the following holds

$$d_B(Dgm(\log \mathcal{F}_A), Dgm(\log \mathcal{G}_B)) \leq \log(\delta)$$

.

The meaning of approximation guarantee (contd.)

We will use the following lemma[4] for our final result.

Lemma (Persistence Approximation Lemma)

If there exist $\delta > 0$ such that the two filtrations $(\mathcal{F})_{\alpha \geq 0}$ and $(\mathcal{G})_{\alpha \geq 0}$ satisfy $\mathcal{F}_{\alpha/\delta} \subseteq \mathcal{G}_\alpha \subseteq \mathcal{F}_{\delta\alpha}$ for all $\alpha \geq 0$, the persistence diagrams $Dgm(\mathcal{F}_\alpha)$ and $Dgm(\mathcal{G}_\alpha)$ are δ-approximations of each other.

This lemma suggests that *proving suitable interleaving between simplicial representations is sufficient to show approximation guarantee between the associated persistence diagrams.*

[4] Donald R Sheehy. "Linear-size approximations to the Vietoris–Rips filtration". In: *Discrete Computational Geometry* 49.4 (2013), pp. 778–796.

Main results.

Lemma (Interleaving)

If the landmark set L is an ϵ-net of the point cloud P, the following interleaving of lazy witness complex at α and Vietoris-Rips complex of L holds, for any $\epsilon \in \mathbb{R}^+$ and $\alpha \geq 2\epsilon$.

$$R_{\alpha/3}(L) \subseteq LW_\alpha(P, L) \subseteq R_{3\alpha}(L)$$

Theorem (Persistent diagram approximation)

If L is an ϵ-net of the point cloud P the persistence diagram $Dgm(\mathcal{LW}_{c+2\epsilon}(L))$ of the filtration $(\mathcal{LW}(L))_{c+2\epsilon}$ induced by L is a 3-approximation to the diagram $Dgm(\mathcal{R}_{c+2\epsilon}(L))$ of the Vietoris-Rips filtration $(\mathcal{R}(L))_{c+2\epsilon}$ induced by L, for $\epsilon \in \mathbb{R}^+$ and $c \geq 0$.

Corollary

If L is an ϵ-net of the point cloud P, the bottleneck distance between the logarithm-scale persistence diagrams $Dgm(\log \mathcal{LW}_{c+2\epsilon}(L))$ and $Dgm(\log \mathcal{R}_{c+2\epsilon}(L))$ is at most $\log(3)$, for $\epsilon \in \mathbb{R}^+$ and $c > 0$.

Main results. (Graph)

Lemma (Interleaving)

If L is an ϵ-net of the vertex set V, the following interleaving of lazy witness complex at α and Vietoris-Rips complex of L holds, for any $\epsilon \in \mathbb{R}^+$ and $\alpha \geq 2\epsilon$

$$R_{\alpha/3}(L) \subseteq LW_\alpha(V, L) \subseteq R_{3\alpha}(L)$$

The main theorem and its corollary follows.

Secondary results: Properties of ϵ-Net

How good are the landmarks? Does the subspace (point cloud sample,subgraph) induced by ϵ-net representative of the actual point cloud or graph data?

Point cloud

The Hausdorff distance between the point cloud P and its ϵ-net $L \subseteq P$ is at most ϵ.

Graph

The Hausdorff distance between (V, d_G) and its ϵ-net induced subspace (L, d_L) is at most ϵ.

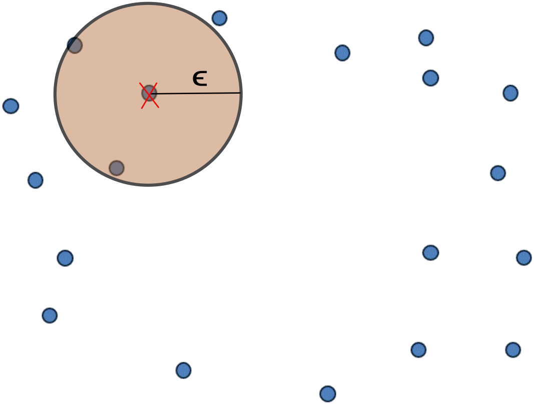
Number of landmarks: For a given ϵ, how many points (vertex) are there in the ϵ-net?

Graph

For a connected unweighted graph of diameter Δ, there exists an ϵ-net of size at most $(\frac{\Delta}{\epsilon})^{O(\log \frac{|V|}{\epsilon})}$

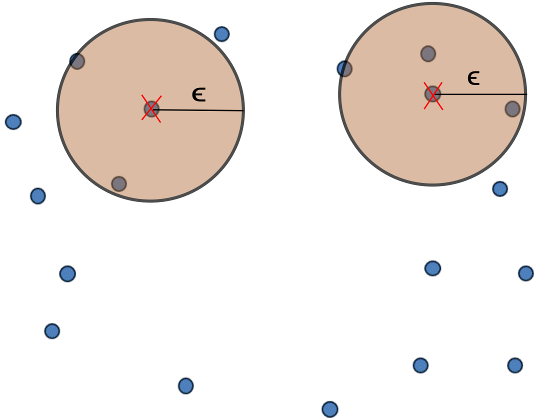
Algorithms

Algorithm for point cloud: ϵ-net-rand



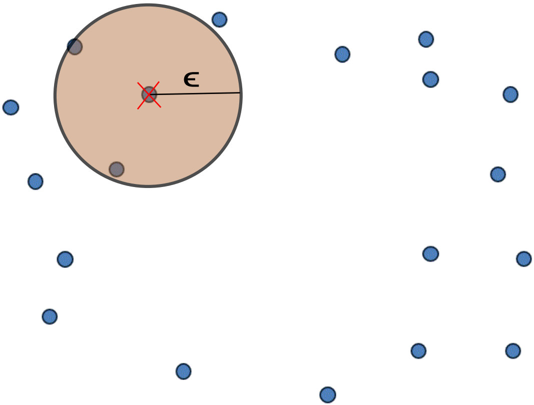
First landmark: Select uniformly at random from the point-cloud. Mark points in its ϵ-ball.

Algorithm for point cloud: ϵ-net-rand



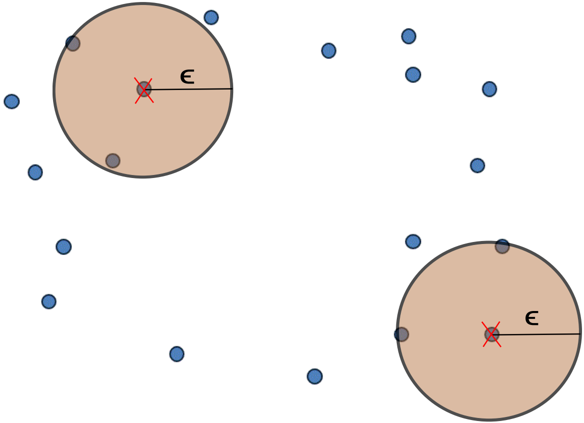
Next landmark: Select u.a.r from the set of unmarked points. Mark points in its ϵ-ball. And so on.

Algorithm for point cloud:: ϵ-net-maxmin



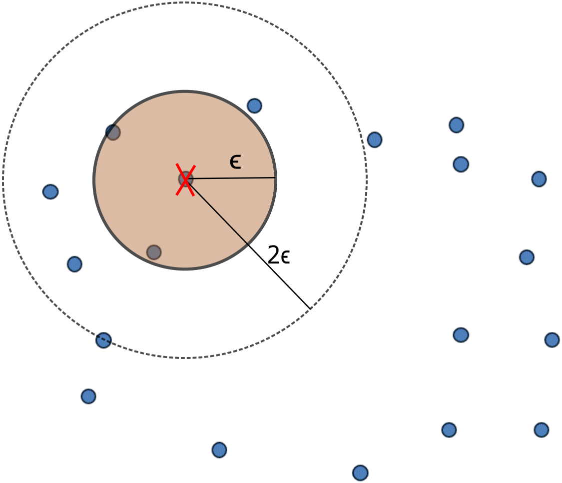
First landmark: Select u.a.r. from the point-cloud. Mark points in its ϵ-ball.

Algorithm for point cloud:: ϵ-net-maxmin



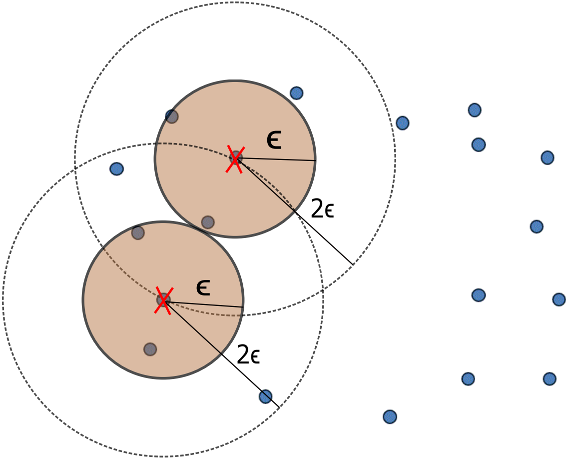
Next landmark: Select the point that is farthest from the current set of landmarks. And so on.

Algorithm for point cloud:: $(\epsilon, 2\epsilon)$-net



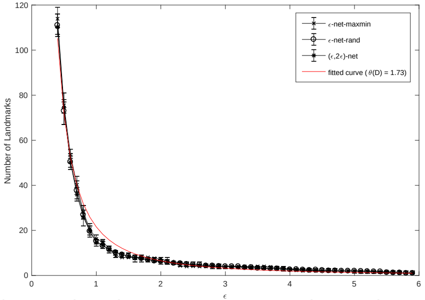
First landmark: Select u.a.r from the point-cloud. Mark points in its ϵ-ball.

Algorithm for point cloud:: $(\epsilon, 2\epsilon)$-net



Second landmark: Select u.a.r from the unmarked points in the $(\epsilon, 2\epsilon)$ envelope of the current set of landmarks. And so on.

Practicality of the algorithms



The size of ϵ-net generated by our algorithms are consistent with Krauthgamer's theoretical upper-bound (dataset: A sample from Torus surface).

Computational Complexity: For n points in $\mathbb{R}^D$

- ϵ-net-rand: $O(\frac{n}{\epsilon^D})$
- ϵ-net-maxmin: $O(\frac{n^2}{\epsilon^D})$
- $(\epsilon, 2\epsilon)$-net: $O(\frac{n^2}{\epsilon^D})$

Algorithm for Graph: Greedy ϵ-net

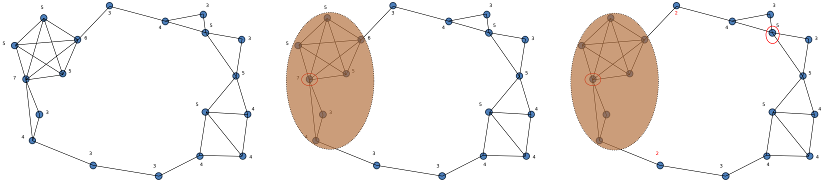


Figure 2: Greedy-ϵ-Net

Compute the ϵ-cover sizes for each vertex.
At each step, select the vertex with the largest ϵ-cover and mark the vertices covered by it, finally, update the ϵ-cover sizes of non-marked vertices. Continue until all vertices are marked as covered.
Time-Complexity: $O(n\hat{u}_\epsilon + (\frac{\Delta}{\epsilon})^{\log \frac{n}{\epsilon}})$ where $\hat{u}_\epsilon$ is the cover-size of the vertex with the largest ϵ-cover.

Algorithm for Graph: Iterative ϵ-net

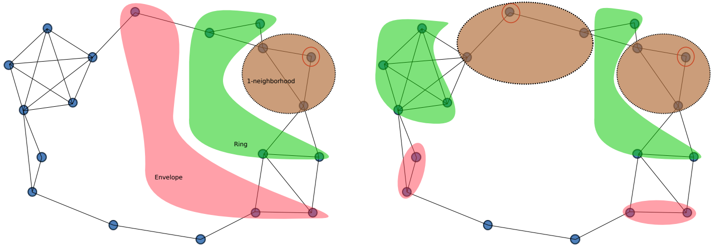


Figure 3: Iterative-ϵ-Net

Remarks: For a given ϵ, we empirically found Greedy ϵ-net to produce the least number of landmarks in general. However Greedy ϵ-net is significantly inefficient compared to Iterative ϵ-net.

Comparison with Other approximate representations.

There are simplicial representations with better approximation guarantee than ours, but

- some of them are limited to point clouds.
- some uses simplicial maps which may not induce a filtration with canonical inclusion maps. As a result, one can not use well-studied matrix reduction algorithm to compute persistence any more.

Comparison with Other approximate representations.

Existing approximations.

- Sparse-Čech filtration[5]: $(1+\epsilon)$ $\epsilon > 0$ w.r.t to Čech. (Does not induce a filtration)
- Sparse-Rips filtration[6]: $(1+2\epsilon)$ $\epsilon < \frac{1}{3}$ w.r.t Vietoris-Rips.
- Simplicial Batch-collapse (Simba)[7]: $(3+\frac{2}{\epsilon-1})$ $\epsilon > 1$ w.r.t Vietoris-Rips (for point cloud only).
- ϵ-net induced lazy witness: $3+2\epsilon$ $\epsilon > 0$ w.r.t Vietoris-Rips on the whole dataset. Works for point cloud and graphs.

[5] M. Kerber and R. Sharathkumar. "Approximate Čech complex in low and high dimensions". In: *International Symposium on Algorithms and Computation*. 2013.

[6] Sheehy, "Linear-size approximations to the Vietoris–Rips filtration".

[7] Tamal K Dey, D. Shi, and Y. Wang. "Simba: An efficient tool for approximating rips-filtration persistence via simplicial batch collapse". In: *Journal of Experimental Algorithmics (JEA)* (2019).

Future works:

We want to use ϵ-net

- to obtain better approximate guarantees.
- to investigate, compare and unify approximation schemes for Vietoris-Rips representations such as Sparse-Rips and Simba.
- to design scalable, fast algorithms for any metric space.
- to efficiently apply persistent homology to machine learning problems.

Thank You!

This is a joint work [8] with Debabrota Basu (INRIA) and Stëphane Bressan (NUS)

[8] The work on point cloud was published in Database and Expert systems proceedings (DEXA) 2019.
The extension for graphs was presented in Applied Topological Data Analysis workshop (ATDA), ECML-PKDD 2019.

Supplementary Slides

Proof of Interleaving lemma:

To prove the first inclusion $R_{\alpha/3}(L) \subseteq LW_\alpha(P, L)$,

1. consider a k-simplex $\sigma_k = [x_0 x_1 \cdots x_k] \in R_{\alpha/3}(L)$.
2. For any edge $[x_i x_j] \in \sigma_k$, let w' be the point in P that is nearest to the vertices of $[x_i x_j]$ and wlog, let that vertex be x_j.
3. Since w' is the nearest neighbour of x_j, $d(w', x_j) \leq \epsilon \leq \frac{\alpha}{2}$ (as $d(L, P) \leq \epsilon$).
4. Since $[x_i x_j] \in R_{\alpha/3}$, $d(x_i, x_j) \leq \frac{\alpha}{3} < \frac{\alpha}{2}$. By triangle inequality, $d(w', x_i) \leq \frac{\alpha}{2} + \frac{\alpha}{2} \leq \alpha$.
5. Therefore x_i is within distance α from w'. The α-neighbourhood of point w' contains both x_i and x_j. Since $d(w', L) \geq 0$, the $(d(w', L) + \alpha)$-neighbourhood of w' also contains x_i, x_j. Therefore, *$[x_i x_j]$ is an edge in $LW_\alpha(P, L)$*.
6. Since the argument is true for any $x_i, x_j \in \sigma_k$, the k-simplex $\sigma_k \in LW_\alpha(P, L)$.

Proof of Interleaving lemma(contd.):

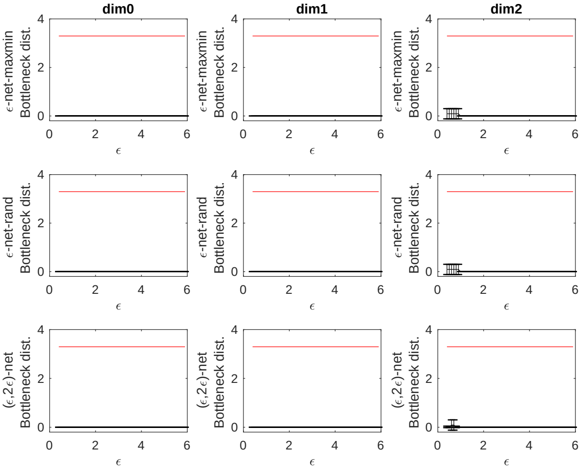
To prove the second inclusion $LW_{\alpha}(P, L, 1) \subseteq R_{3\alpha}(L)$

1. consider a k-simplex $\sigma_k = [x_0 x_1 \cdots x_k] \in LW_{\alpha}(P, L)$.
2. By definition of lazy witness complex, $\forall [x_i x_j] \in \sigma_k$ there is a witness $w \in P$ such that, the $(d(w, L) + \alpha)$-neighbourhood of w contains both x_i and x_j.
3. Hence, $d(w, x_i) \leq d(w, L) + \alpha \leq \epsilon + \alpha \leq 3\alpha/2$. By the same argument, $d(w, x_j) \leq 3\alpha/2$.
4. By triangle inequality, $d(x_i, x_j) \leq 3\alpha$. Therefore, *$[x_i x_j]$ is an edge in $R_{3\alpha}(L)$*.
5. Since the argument is true for any $x_i, x_j \in \sigma_k$, the k-simplex $\sigma_k \in R_{3\alpha}(L)$.

On computational infeasiblity of Čech complex: miniball algorithm

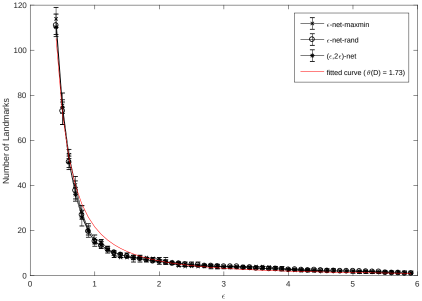
- One can run minimal algorithm to find the minimum enclosing radius of a simplex σ. For a given σ, that radius will give the offset α where σ appears first.
- Fastest miniball algorithm runs in $O(|\sigma|)$ time. (note that, the offset for σ can be found in constant time for Vietoris-rips complex if the complex is constructed dimension by dimension)
- However to run miniball algorithm one still needs to enumerate σ first, rendering the Čech filtration computation to be at least twice as slow as corresponding Vietoris-Rips filtration.

Experimental Validation of Approximation Guarantee

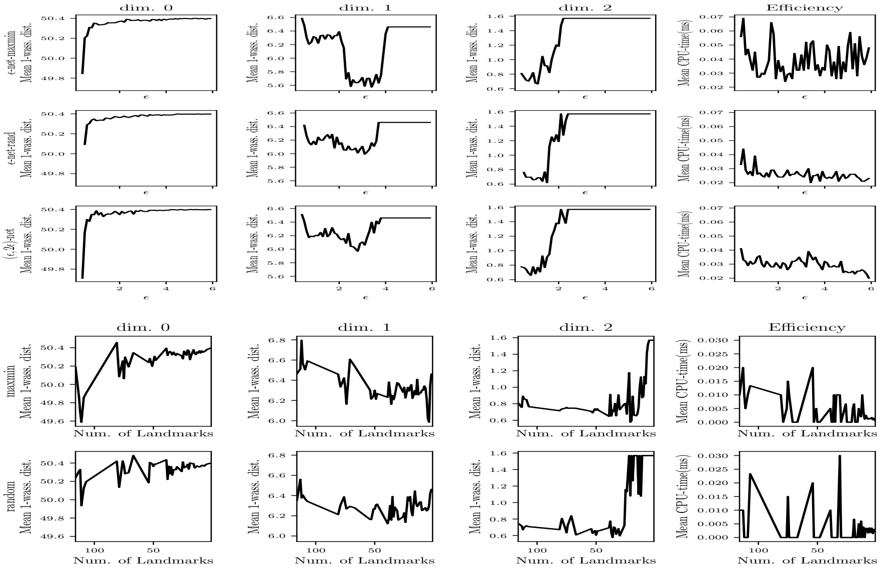


Validation of the bound on Point-cloud sampled from Torus (red-line = log(3) upper bound).

Comparing our algorithm-output ϵ-net with Krauthgamer's guarantee



Experimental Evaluation: Effectiveness-Efficiency



Effectiveness and Efficiency of the algorithms on Torus dataset.

Experiments: Datasets

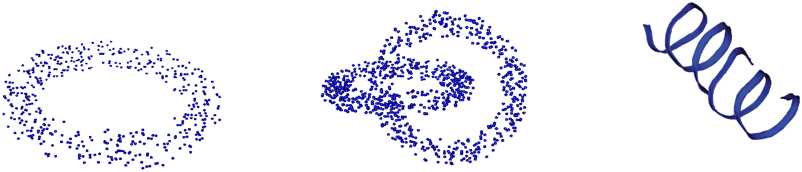


Figure 4: (left) Torus, (middle) Tangled-torus, and (right) 1grm Dataset

Relation to Maxmin and Random Landmark Selection Algorithms

- Given the number of landmarks $K > 1$, the set of landmarks selected by the algorithm random/maxmin is δ-sparse where δ is the minimum of the pairwise distances among the landmarks.
- The choice of K may not necessarily make the landmarks a δ-sample of the point cloud.

Algorithm complexity

- ϵ-net-rand: $O(\frac{n}{\epsilon^D})$
- ϵ-net-maxmin: $O(\frac{n^2}{\epsilon^D})$
- $(\epsilon, 2\epsilon)$-net: $O(\frac{n^2}{\epsilon^D})$

Experimental Evaluation: Stability

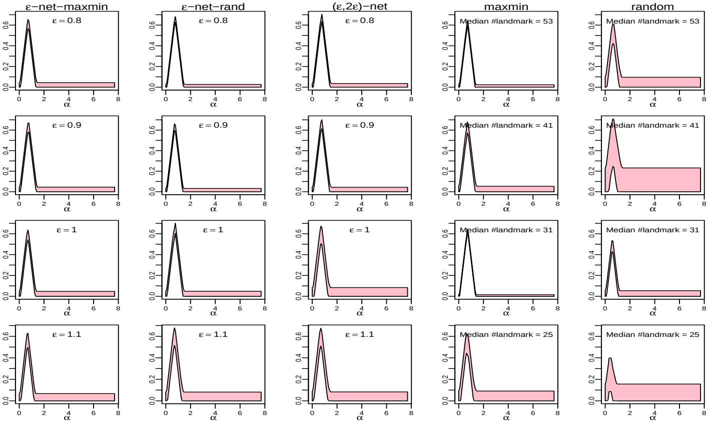


Figure 5: 95% confidence band of the rank one persistence landscape at dimension 1 of the lazy witness filtration induced by the landmark selection algorithms on Tangled-torus dataset.

Whats Next?

The topological approximation guarantee is

- with respect to the Vietoris-Rips complex on ϵ-**net landmarks** chosen from a **point-cloud input**.

Next up -

- Better guarantee w.r.t Vietoris-Rips complex on point-cloud:-
  - Improved the approximation guarantee from 3-approximation of $R_\alpha(L)$ to $\frac{3log(c)}{2}$-approximation of $R_\alpha(P)$ for $c \geq 2$.
- Graph data [9]:-
  - Defined ϵ-net for graphs.
  - Devised algorithms for computing ϵ-net of graphs.
  - Potential applications: Graph clustering, Graph visualization, Graph classification.
- Comparison with Sparse-Rips and Graph Induced filtration (A weakness!).

---

[1] To appear at ECML-PKDD'19 workshop on Applied Topological Data Analysis