

# Persistent Homology of KDE filtration on Rips complex and Related Work

Jisu KIM

INRIA Saclay

18/Oct/2019

## Introduction

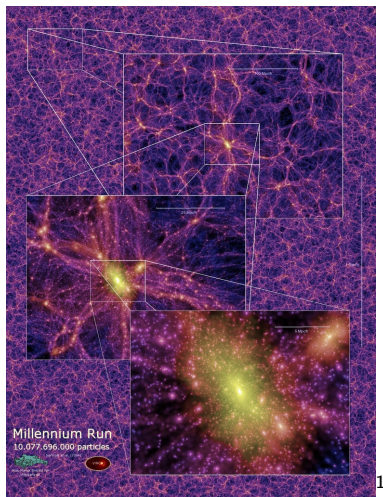
Conclusion

Persistent Homology of KDE filtration on Rips complex

Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension

Homotopy Reconstruction of a positive reach set using Rips Complex

Topological structures in the data provide information.



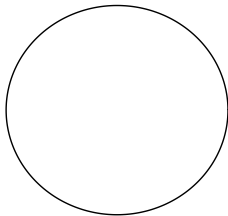
---

<sup>1</sup>[http://www.mpa-garching.mpg.de/galform/virgo/millennium/poster\\_half.jpg](http://www.mpa-garching.mpg.de/galform/virgo/millennium/poster_half.jpg)

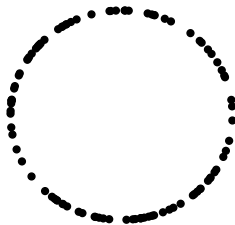
Homology of finite sample is different from homology of underlying manifold, hence it cannot be directly used for the inference.

- ▶ When analyzing data, we prefer robust features where features of the underlying manifold can be inferred from features of finite samples.
- ▶ Homology is not robust:

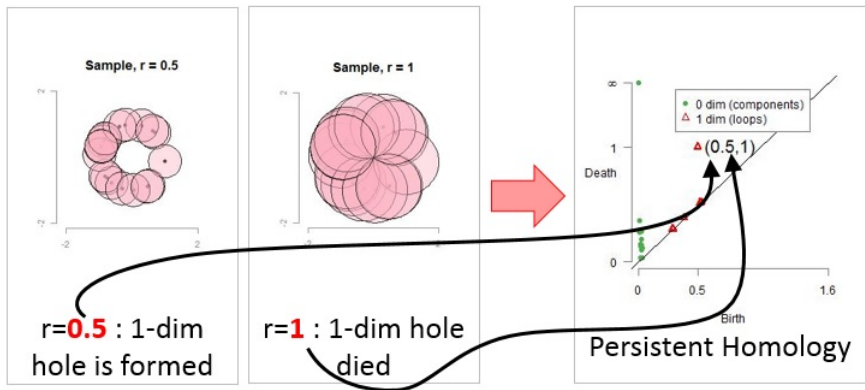
Underlying circle:  $\beta_0 = 1, \beta_1 = 1$



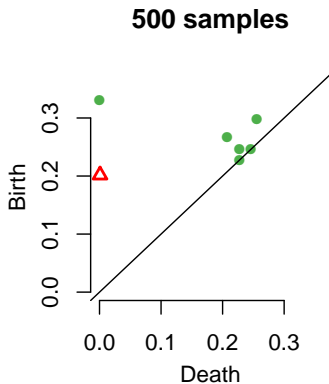
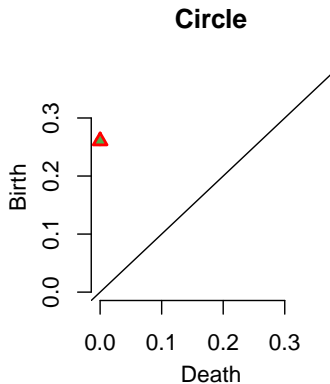
100 samples:  $\beta_0 = 100, \beta_1 = 0$



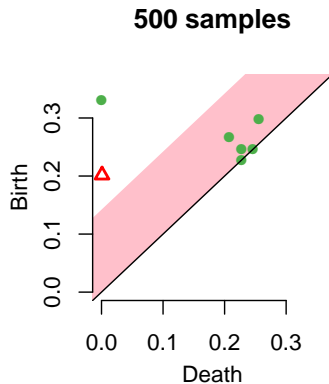
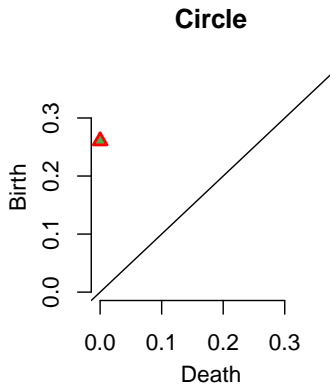
Persistent homology computes homologies on collection of sets, and tracks when topological features are born and when they die.



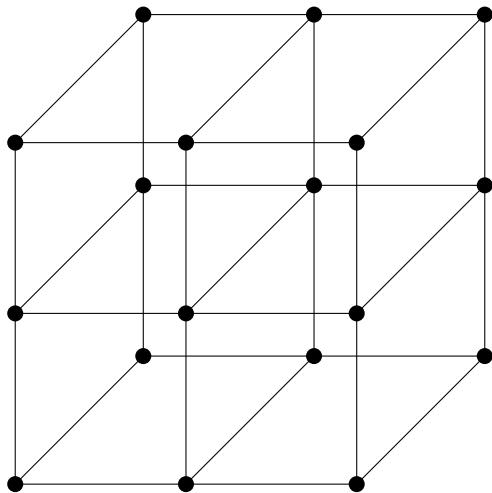
Persistent homology of the underlying manifold can be inferred from persistent homology of finite samples.



Confidence band for persistent homology of density function separates homological signal from homological noise.

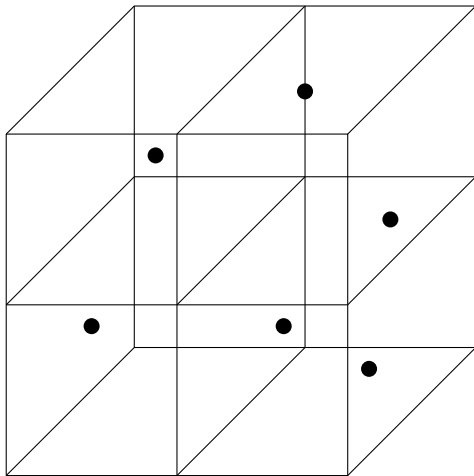


Computing the persistent homology of density function incurs computing on a grid of points, which is infeasible in high dimensional space.



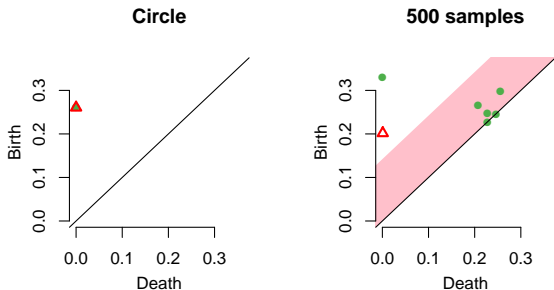


Computing the persistent homology of density function on data points reduces computational complexity.



How can we analyze the convergence and compute a confidence band for the persistent homology of density function with computation on data points?

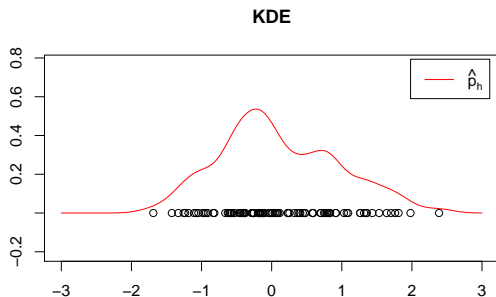
- ▶ (Shin, Kim, Rinaldo, Wasserman, 2019?) : extending work from Fasy et al. [2014], Bobrowski et al. [2014], Chazal et al. [2011].



We rely on the kernel density estimator to compute the persistent homology of density function.

- ▶ The kernel density estimator is

$$\hat{p}_h(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right).$$



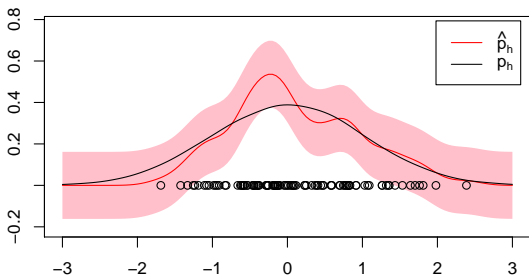
To analyze the convergence of the persistent homology of density function, the uniform convergence of the kernel density estimator needs to be studied.

- ▶ Let  $p_h(x) = \mathbb{E}_P[\hat{p}_h(x)]$ , then we need to study the analytic behavior of

$$\sup_{x \in \text{supp}(P)} |\hat{p}_h(x) - p_h(x)|.$$

- ▶ (Kim, Shin, Rinaldo, Wasserman, 2019)

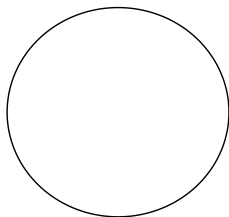
Uniform bound on KDE



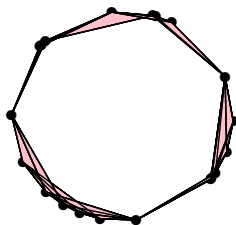
We are also interested in recovering homotopy type of the underlying manifold from finite samples.

- ▶ ??: extending work from Niyogi et al. [2008], Attali et al. [2013]

**Underlying circle**



**Finite samples**



Introduction

Conclusion

Persistent Homology of KDE filtration on Rips complex

Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension

Homotopy Reconstruction of a positive reach set using Rips Complex

# Persistent homology of KDE filtration on rips complex and Related Work

- ▶ Persistent homology of KDE filtration on rips complex (Shin, Kim, Rinaldo, Wasserman, 2019?)
- ▶ Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension (Kim, Shin, Rinaldo, Wasserman, 2019)
- ▶ ??

The persistent homology of the KDE filtration on Rips complexes is consistent.

### Theorem

(Shin, Kim, Rinaldo, Wasserman, 2019?, Theorem 16)

$$d_B \left( PH_*^R(\hat{p}_{h_n}, r_n), PH_*^{supp(P)}(p_{h_n}) \right) = O_P \left( \sqrt{\frac{\log(1/h_n)}{nh_n^d}} + \|r_n\|_\infty \right).$$



# Confidence set for the persistent homology of the KDE filtration.

- ▶ We let the confidence set as the ball centered at  $PH_*^R(\hat{p}_{h_n}, r_n)$  and radius  $\hat{b}_\alpha$ , i.e.

$$\hat{C}_\alpha = \left\{ \mathcal{P} : d_B(\mathcal{P}, PH_*^R(\hat{p}_{h_n}, r_n)) \leq \hat{b}_\alpha \right\}.$$

This is a valid confidence set by the following theorem.

## Theorem

(Shin, Kim, Rinaldo, Wasserman, 2019?, Theorem 20)

$$\mathbb{P} \left( PH_*^{\text{supp}(P)}(p_{h_n}) \in \hat{C}_\alpha \right) \geq 1 - \alpha + o(1).$$

The uniform convergence rate of the Kernel Density Estimator is derived adaptive to the intrinsic dimension.

### Theorem

(Kim, Shin, Rinaldo, Wasserman, 2019, Corollary 13, Corollary 17)  
Suppose  $l_n \rightarrow 0$  and  $nl_n \rightarrow \infty$ . Then with high probability,

$$\sqrt{\frac{1}{nl_n^{2d-d_{\text{vol}}}}} \lesssim \sup_{h \geq l_n, x \in \mathbb{X}} |\hat{p}_h(x) - p_h(x)| \lesssim \sqrt{\frac{\log(1/l_n)}{nl_n^{2d-d_{\text{vol}}}}},$$

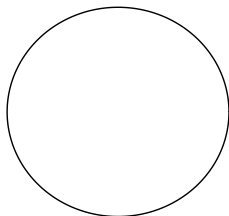
for all large  $n$ .

The homotopy type of a positive reach set can be reconstructed from the Rips complex.

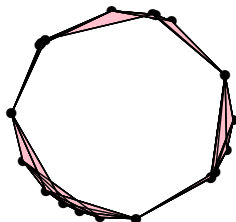
### Theorem

Let  $\mathbb{X}$  be a subset with its reach  $\tau_{\mathbb{X}} > 0$ , and let  $\mathcal{X} \subset \mathbb{X}$  be a finite sample. When  $r \in [d_H(\mathbb{X}, \mathcal{X}), C\tau_{\mathbb{X}}]$ , then the Rips complex  $R(\mathcal{X}, r)$  is homotopic equivalent to  $\mathbb{X}$ , with  $C = \frac{1}{\sqrt{2}} \approx 0.71\dots$

**Underlying circle**



**Rips complex**



Thank you!

Introduction

Conclusion

Persistent Homology of KDE filtration on Rips complex

Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension

Homotopy Reconstruction of a positive reach set using Rips Complex

We rely on the kernel density estimator to extract topological information of the underlying distribution.

- ▶ The kernel density estimator is

$$\hat{p}_h(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right).$$

We are considering the upper level set of the average kernel density estimator on the support.

- ▶ Let  $X_1, \dots, X_n \sim P$ , then the average kernel density estimator is

$$p_h(x) = \mathbb{E}[\hat{p}_h(x)] = \frac{1}{h^d} \mathbb{E} \left[ K \left( \frac{x - X}{h} \right) \right].$$

- ▶ We are considering the upper level sets of the average kernel density estimator

$$\{D_L\}_{L>0}, \text{ where } D_L := \{x \in \text{supp}(P) : p_h(x) \geq L\}.$$

We are considering the upper level set of the average kernel density estimator on the support.

- ▶ We are considering the upper level sets of the average KDE

$$\{D_L\}_{L>0}, \text{ where } D_L := \{x \in \text{supp}(P) : p_h(x) \geq L\}.$$

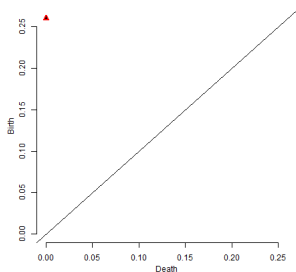


We are targeting the persistent homology of the upper level set of the average kernel density estimator on the support.

- ▶ We are considering the upper level sets of the average KDE

$$\{D_L\}_{L>0}, \text{ where } D_L := \{x \in \text{supp}(P) : p_h(x) \geq L\},$$

and targeting its persistent homology  $PH_*^{\text{supp}(P)}(p_h)$ .

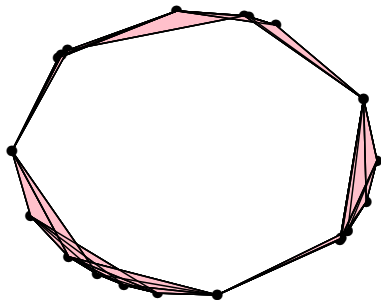


We use the Rips complex to estimate the target persistent homology.

- ▶ For  $\mathcal{X} \subset \mathbb{R}^d$  and  $r > 0$ , the Rips complex  $R(\mathcal{X}, r)$  is defined as

$$R(\mathcal{X}, r) = \{[X_{i_1}, \dots, X_{i_k}] \subset \mathcal{X} : d(X_{i_j}, X_{i_l}) < 2r, 1 \leq \forall j \neq l \leq k\}.$$

### Rips Complex



We estimate the target level set by considering the Rips complex generated from the level set of the KDE.

- ▶ For  $\mathcal{X}_n = \{X_1, \dots, X_n\}$ , we estimate the target level set by the level sets of the KDE on Rips complexes,

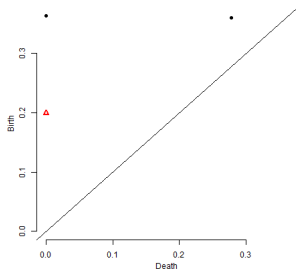
$$\left\{ R \left( \mathcal{X}_{n,L}^{\hat{p}_h}, r \right) \right\}_{L>0}, \text{ where } \mathcal{X}_{n,L}^{\hat{p}_h} := \{X_i \in \mathcal{X}_n : \hat{p}_h(X_i) \geq L\}.$$

We estimate the target persistent homology by the persistent homology of the KDE filtration on Rips complexes.

- ▶ We estimate the target persistent homology by the persistent homology of the level sets of the KDE on Rips complexes,

$$\left\{ R \left( \mathcal{X}_{n,L}^{\hat{p}_h}, r \right) \right\}_{L>0}, \text{ where } \mathcal{X}_{n,L}^{\hat{p}_h} = \{X_i \in \mathcal{X}_n : \hat{p}_h(X_i) \geq L\}.$$

and denote the persistent homology as  $PH_*^R(\hat{p}_h, r)$ .



We estimate the target level set by Rips complexes from the KDE level sets.

- ▶ We approximate the target level set

$$\{D_L\}_{L>0}, \text{ where } D_L := \{x \in \text{supp}(P) : p_h(x) \geq L\},$$

by the level sets of the KDE on Rips complexes,

$$\left\{R\left(\mathcal{X}_{n,L}^{\hat{p}_h}, r\right)\right\}_{L>0}, \text{ where } \mathcal{X}_{n,L}^{\hat{p}_h} = \{X_i \in \mathcal{X}_n : \hat{p}_h(X_i) \geq L\}.$$

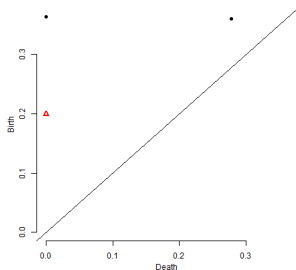
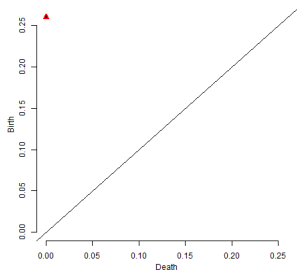
We estimate the target persistent homology by the persistent homology of the KDE filtration on Rips complexes.

- ▶ We estimate the target persistent homology

$$PH_*^{\text{supp}(P)}(\rho_h),$$

by the persistent homology of the KDE filtration on Rips complexes,

$$PH_*^R(\hat{\rho}_h, r).$$



The persistent homology of the KDE filtration on Rips complexes is stable.

### Theorem

(Shin, Kim, Rinaldo, Wasserman, 2019?, Proposition 15) When  $\text{supp}(P) \subset \bigcup_i B(X_i, r)$ , then

$$d_B \left( PH_*^R(\hat{p}_{h_n}, r_n), PH_*^{\text{supp}(P)}(p_{h_n}) \right) \leq \sup_{x \in \text{supp}(P)} |\hat{p}_{h_n}(x) - p_{h_n}(x)| + O(\|r_n\|_\infty).$$

The persistent homology of the KDE filtration on Rips complexes is consistent.

### Theorem

(Shin, Kim, Rinaldo, Wasserman, 2019?, Theorem 16)

$$d_B \left( PH_*^R(\hat{p}_{h_n}, r_n), PH_*^{supp(P)}(p_{h_n}) \right) = O_P \left( \sqrt{\frac{\log(1/h_n)}{nh_n^d}} + \|r_n\|_\infty \right).$$



# Confidence set

- ▶ An asymptotic  $1 - \alpha$  confidence set  $\hat{C}_\alpha$  is a random set of persistent homologies satisfying

$$\mathbb{P}(PH_*^{\text{supp}(P)}(p_{h_n}) \in \hat{C}_\alpha) \geq 1 - \alpha + o(1).$$

# Confidence set for the persistent homology of the KDE filtration.

- ▶ We let the confidence set as the ball centered at  $PH_*^R(\hat{p}_{h_n}, r_n)$  and radius  $\hat{b}_\alpha$ , i.e.

$$\hat{C}_\alpha = \left\{ \mathcal{P} : d_B(\mathcal{P}, PH_*^R(\hat{p}_{h_n}, r_n)) \leq \hat{b}_\alpha \right\}.$$

This is a valid confidence set by the following theorem.

## Theorem

(Shin, Kim, Rinaldo, Wasserman, 2019?, Theorem 20)

$$\mathbb{P} \left( PH_*^{supp(P)}(p_{h_n}) \in \hat{C}_\alpha \right) \geq 1 - \alpha + o(1).$$

Introduction

Conclusion

Persistent Homology of KDE filtration on Rips complex

**Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension**

Homotopy Reconstruction of a positive reach set using Rips Complex

The consistency of the persistent homology of the KDE filtration on Rips complexes is derived from the uniform convergence of the kernel density estimator.

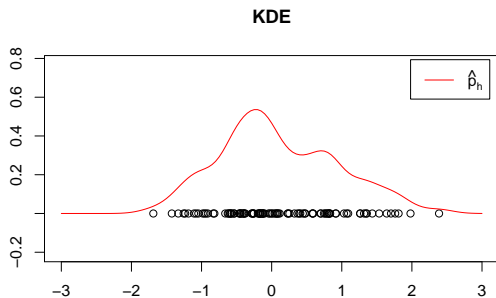
- ▶ We would like to study the analytic behavior of the uniform convergence of the kernel density estimator, i.e.,

$$\sup_{x \in \text{supp}(P)} |\hat{p}_{h_n}(x) - p_{h_n}(x)|.$$

# Kernel Density Estimator

- ▶ For  $X_1, \dots, X_n \sim P$ , a given kernel function  $K$ , and a bandwidth  $h > 0$ , the Kernel Density Estimator (KDE)  $\hat{p}_h : \mathbb{R}^d \rightarrow \mathbb{R}$  is

$$\hat{p}_h(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right).$$

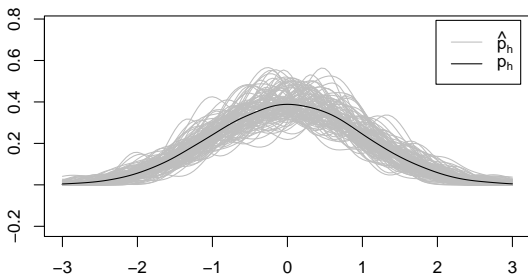


# Average Kernel Density Estimator

- ▶ The Average Kernel Density Estimator (KDE)  $p_h : \mathbb{R}^d \rightarrow \mathbb{R}$  is

$$p_h(x) = \mathbb{E}_P [\hat{p}_h(x)] = \frac{1}{h^d} \mathbb{E}_P \left[ K \left( \frac{x - X}{h} \right) \right].$$

Average KDE

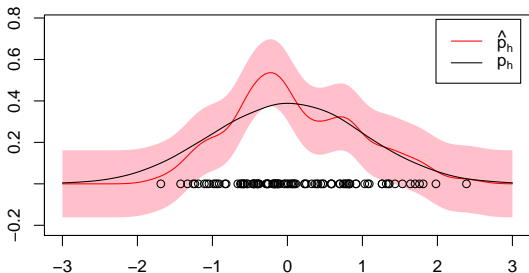


# We get the uniform convergence rate on Kernel Density Estimator.

- ▶ Fix a subset  $\mathbb{X} \subset \mathbb{R}^d$ , we need uniform control of the Kernel Density Estimator over  $\mathbb{X}$ ,  $\sup_{x \in \mathbb{X}} |\hat{p}_h(x) - p_h(x)|$ , for various purposes.
- ▶ We get the concentration inequalities for the Kernel Density Estimator in the supremum norm that hold uniformly over the selection of the bandwidth, i.e.,

$$\sup_{h \geq I_n, x \in \mathbb{X}} |\hat{p}_h(x) - p_h(x)|.$$

Uniform bound on KDE



The volume dimension characterizes the intrinsic dimension of the distribution related to the convergence rate of the Kernel Density Estimator.

- ▶ For a probability distribution  $P$  on  $\mathbb{R}^d$ , the volume dimension is

$$d_{\text{vol}} := \sup \left\{ \nu \geq 0 : \limsup_{r \rightarrow 0} \sup_{x \in \mathbb{X}} \frac{P(\mathbb{B}(x, r))}{r^\nu} < \infty \right\},$$

where  $\mathbb{B}(x, r) = \{y \in \mathbb{R}^d : \|x - y\| < r\}$ .

- ▶ In other words, the volume dimension is the maximum possible exponent rate dominating the probability volume decay on balls.



The uniform convergence rate of the Kernel Density Estimator is derived in terms of the volume dimension.

### Theorem

(Kim, Shin, Rinaldo, Wasserman, 2019, Corollary 13, Corollary 17) Let  $P$  be a probability distribution on  $\mathbb{R}^d$  satisfying weak assumptions and  $K$  be a kernel function satisfying weak assumptions. Suppose  $l_n \rightarrow 0$  and  $nl_n \rightarrow \infty$ . Then with high probability,

$$\sqrt{\frac{1}{nl_n^{2d-d_{\text{vol}}}}} \lesssim \sup_{h \geq l_n, x \in \mathbb{X}} |\hat{p}_h(x) - p_h(x)| \lesssim \sqrt{\frac{\log(1/l_n)}{nl_n^{2d-d_{\text{vol}}}}},$$

for all large  $n$ .

Introduction

Conclusion

Persistent Homology of KDE filtration on Rips complex

Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension

Homotopy Reconstruction of a positive reach set using Rips Complex

Cech complex is constructed by finding nonempty intersections of balls.

- ▶ For  $\mathcal{X}, \mathbb{X} \subset \mathbb{R}^d$  and  $r > 0$ , the intrinsic Cech complex  $\check{C}ech_{\mathbb{X}}(\mathcal{X}, r)$  is defined as

$$\check{C}ech_{\mathbb{X}}(\mathcal{X}, r) = \left\{ [X_{i_1}, \dots, X_{i_k}] \subset \mathcal{X} : \bigcap_{j=1}^k \mathbb{B}_{\mathbb{X}}(X_{i_j}, r) \neq \emptyset \right\},$$

where  $\mathbb{B}_{\mathbb{X}}(x, r) = \{y \in \mathbb{X} : \|y - x\| < r\}$  is the intrinsic ball of radius  $r$  centered at  $x$ .

- ▶ The ambient Cech complex  $\check{C}ech_{\mathbb{X}}(\mathcal{X}, r)$  is when  $\mathbb{X} = \mathbb{R}^d$ .

The reach of  $\mathbb{X}$ , denoted by  $\text{reach}(\mathbb{X})$ , is the minimum distance from the medial axis to the set.

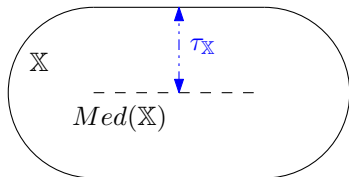
s

- ▶ The Medial Axis is the set of points that have at least two nearest neighbors, i.e.

$$\text{Med}(\mathbb{X}) = \{z \in \mathbb{R}^d : \text{there exists } p \neq q \in \mathbb{X} \text{ with} \\ \|p - z\| = \|q - z\| = d(z, \mathbb{X})\}.$$

- ▶ The reach is the minimum distance from the medial axis to the set, i.e.

$$\text{reach}(\mathbb{X}) = \inf_{x \in \text{Med}(\mathbb{X}), y \in \mathbb{X}} \|x - y\|.$$



The homotopy type of a positive reach set is reconstructed from the intrinsic Čech complex.

### Theorem

*Let  $\mathbb{X} \subset \mathbb{R}^d$  be a subset with reach  $\tau_{\mathbb{X}} > 0$  and  $\mathcal{X} \subset \mathbb{X}$  be a finite sample. When  $r \in [d_H(\mathbb{X}, \mathcal{X}), \sqrt{2}\tau_{\mathbb{X}}]$ , then the intrinsic Čech complex  $\check{C}ech_{\mathbb{X}}(\mathcal{X}, r)$  is homotopic equivalent to  $\mathbb{X}$ .*

- ▶ The condition on  $r$  is optimal.

The homotopy type of a positive reach set can be reconstructed from the ambient Čech complex.

### Theorem

Let  $\mathbb{X} \subset \mathbb{R}^d$  be a subset with reach  $\tau_{\mathbb{X}} > 0$  and  $\mathcal{X} \subset \mathbb{X}$  be a finite sample. When  $r \in [d_H(\mathbb{X}, \mathcal{X}), C\tau_{\mathbb{X}}]$ , then the ambient Čech complex  $\check{C}ech_{\mathbb{R}^m}(\mathcal{X}, r)$  is homotopic equivalent to  $\mathbb{X}$ .

- ▶ Previous result:  $C = 3 - \sqrt{8} \approx 0.17$  in Niyogi et al. [2008]
- ▶ Our (tentative) result:  $C = 1/\sqrt{2} \approx 0.71$

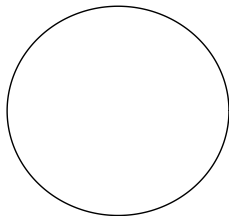
The homotopy type of a positive reach set can be reconstructed from the Rips complex.

### Theorem

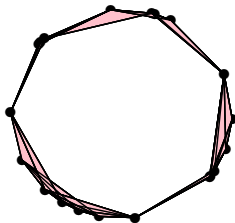
Let  $\mathbb{X} \subset \mathbb{R}^d$  be a subset with its reach  $\tau_{\mathbb{X}} > 0$ , and let  $\mathcal{X} \subset \mathbb{X}$  be a finite sample. When  $r \in [d_H(\mathbb{X}, \mathcal{X}), C\tau_{\mathbb{X}}]$ , then the Rips complex  $R(\mathcal{X}, r)$  is homotopic equivalent to  $\mathbb{X}$ .

- ▶ Previous result:  $C = \frac{2\sqrt{2-\sqrt{2}}-\sqrt{2}}{2+\sqrt{2}} \approx 0.034$  in Attali et al. [2013]
- ▶ Our (tentative) result:  $C = 1/\sqrt{2} \approx 0.71$

**Underlying circle**



**Rips complex**



Thank you!



## Stability and Statistical Inference for Persistent Homology

Reference

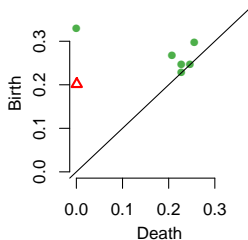
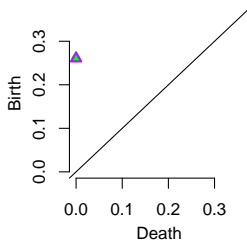
Bottleneck distance gives a metric on the space of persistent homology.

### Definition

Let  $D_1, D_2$  be multiset of points. Bottleneck distance is defined as

$$d_B(D_1, D_2) = \inf_{\gamma} \sup_{x \in D_1} \|x - \gamma(x)\|_{\infty},$$

where  $\gamma$  ranges over all bijections from  $D_1$  to  $D_2$ .



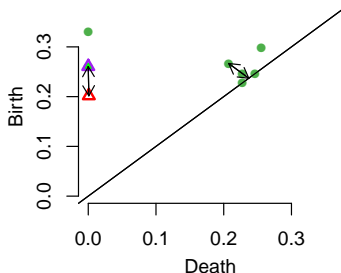
Bottleneck distance gives a metric on the space of persistent homology.

### Definition

Let  $D_1, D_2$  be multiset of points. Bottleneck distance is defined as

$$d_B(D_1, D_2) = \inf_{\gamma} \sup_{x \in D_1} \|x - \gamma(x)\|_{\infty},$$

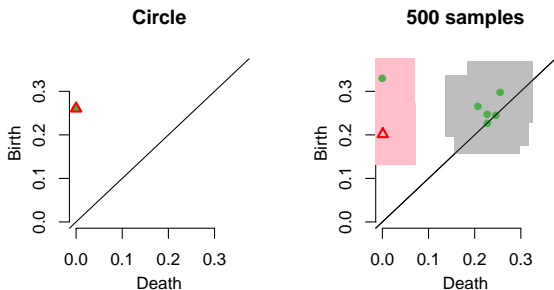
where  $\gamma$  ranges over all bijections from  $D_1$  to  $D_2$ .



Confidence band for the persistent homology is a random quantity containing the persistent homology with high probability.

Let  $M$  be a compact manifold, and  $X = \{X_1, \dots, X_n\}$  be  $n$  samples. Let  $f_M$  and  $f_X$  be corresponding functions whose persistent homology is of interest. Given the significance level  $\alpha \in (0, 1)$ ,  $(1 - \alpha)$  confidence band  $c_n = c_n(X)$  is a random variable satisfying

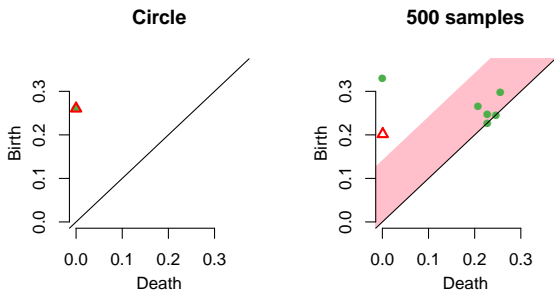
$$\mathbb{P}(d_B(Dgm(f_M), Dgm(f_X)) \leq c_n) \geq 1 - \alpha.$$



Confidence band for the persistent homology is a random quantity containing the persistent homology with high probability.

Let  $M$  be a compact manifold, and  $X = \{X_1, \dots, X_n\}$  be  $n$  samples. Let  $f_M$  and  $f_X$  be corresponding functions whose persistent homology is of interest. Given the significance level  $\alpha \in (0, 1)$ ,  $(1 - \alpha)$  confidence band  $c_n = c_n(X)$  is a random variable satisfying

$$\mathbb{P}(d_B(Dgm(f_M), Dgm(f_X)) \leq c_n) \geq 1 - \alpha.$$



Confidence band for the persistent homology can be computed using the bootstrap algorithm.

1. Given a sample  $X = \{x_1, \dots, x_n\}$ , compute the kernel density estimator  $\hat{\rho}_h$ .
2. Draw  $X^* = \{x_1^*, \dots, x_n^*\}$  from  $X = \{x_1, \dots, x_n\}$  (with replacement), and compute  $\theta^* = \sqrt{n} \|\hat{\rho}_h^*(x) - \hat{\rho}_h(x)\|_\infty$ , where  $\hat{\rho}_h^*$  is the density estimator computed using  $X^*$ .
3. Repeat the previous step  $B$  times to obtain  $\theta_1^*, \dots, \theta_B^*$
4. Compute  $q_\alpha = \inf \left\{ q : \frac{1}{B} \sum_{j=1}^B I(\theta_j^* \geq q) \leq \alpha \right\}$
5. The  $(1 - \alpha)$  confidence band for  $\mathbb{E}[\hat{\rho}_h]$  is  $\left[ \hat{\rho}_h - \frac{q_\alpha}{\sqrt{n}}, \hat{\rho}_h + \frac{q_\alpha}{\sqrt{n}} \right]$ .

## Stability and Statistical Inference for Persistent Homology

### Reference

## Reference

- Dominique Attali, André Lieutier, and David Salinas. Vietoris-Rips complexes also provide topologically correct reconstructions of sampled shapes. *Comput. Geom.*, 46(4):448–465, 2013. ISSN 0925-7721. doi: 10.1016/j.comgeo.2012.02.009. URL <https://doi.org/10.1016/j.comgeo.2012.02.009>.
- O. Bobrowski, S. Mukherjee, and J. E. Taylor. Topological consistency via kernel estimation. *ArXiv e-prints*, July 2014.
- Frédéric Chazal, Leonidas J Guibas, Steve Y Oudot, and Primoz Skraba. Scalar field analysis over point cloud data. *Discrete & Computational Geometry*, 46(4):743–775, 2011.
- Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan, and Aarti Singh. Confidence sets for persistence diagrams. *Ann. Statist.*, 42(6):2301–2339, 12 2014. doi: 10.1214/14-AOS1252. URL <http://dx.doi.org/10.1214/14-AOS1252>.
- Jisu Kim, Jaehyeok Shin, Alessandro Rinaldo, and Larry Wasserman. Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension. *arXiv e-prints*, art. arXiv:1810.05935, 2019.
- Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples