# Some Topics in Computational Topology

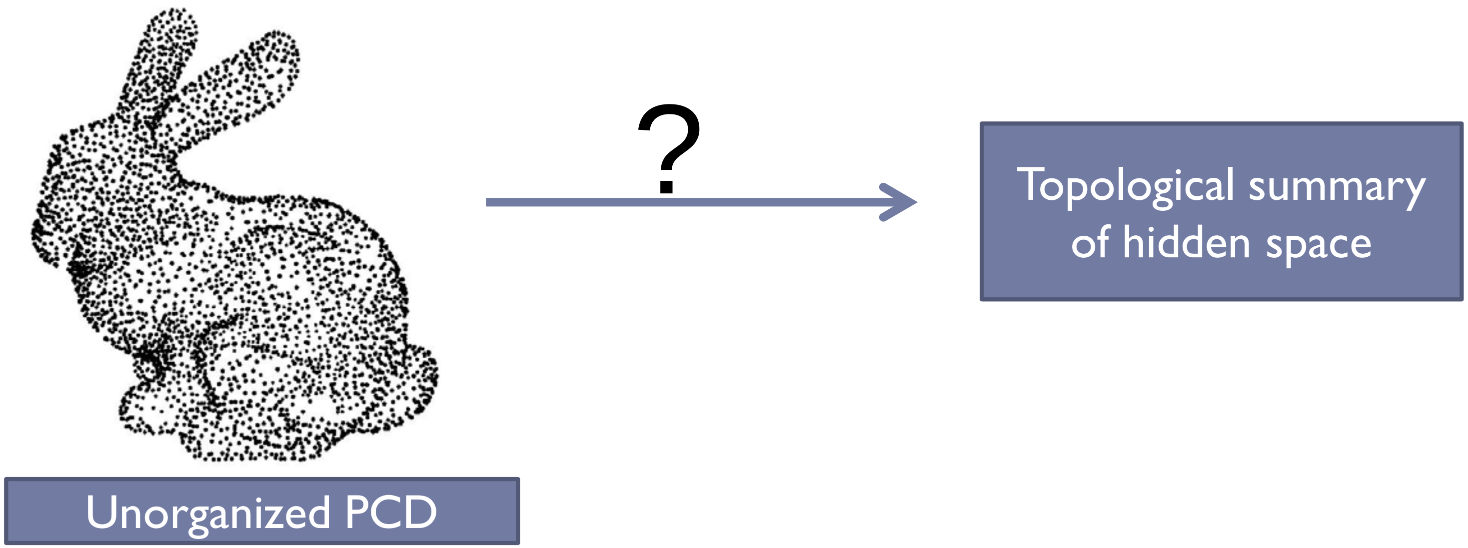
Yusu Wang
*Ohio State University*

*AMS Short Course 2014*

# Introduction

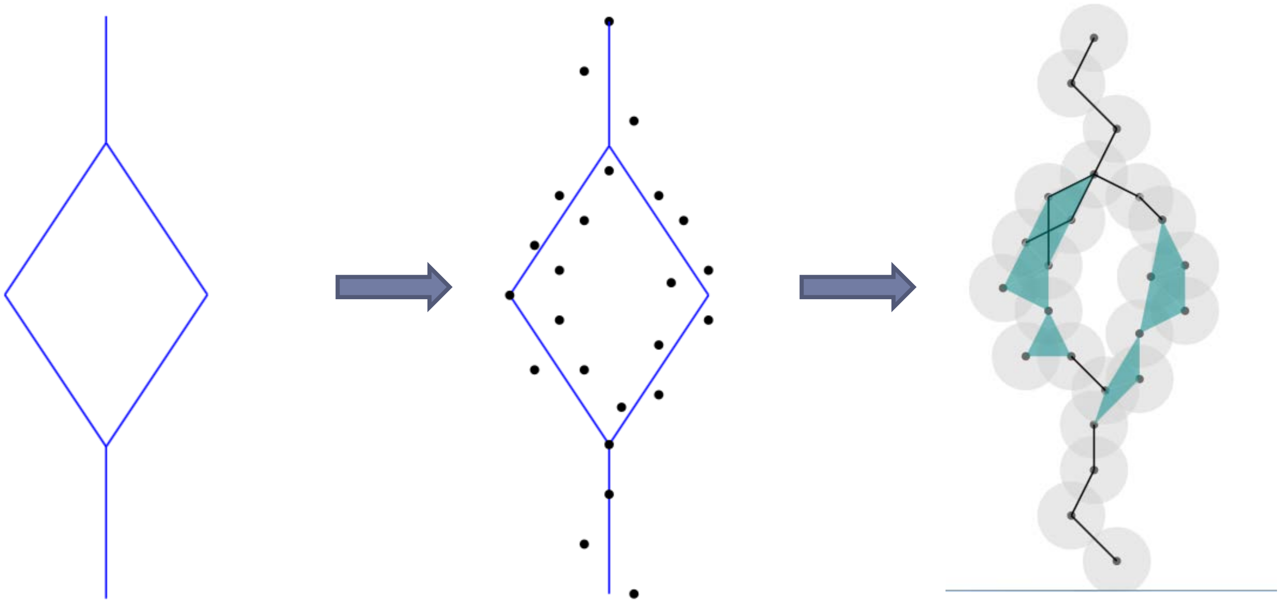
- Much recent developments in computational topology
  - Both in theory and in their applications
  - E.g, the theory of persistence homology
    - *[Edelsbrunner, Letscher, Zomorodian, DCG 2002], [Zomorodian and Carlsson, DCG 2005], [Carlsson and de Silva, FoCM 2010], …*

- This short course:
  - A computational perspective:
    - Estimation and inference of topological information / structure from point clouds data




- Develop discrete analog for the continuous case
- Approximation from discrete samples with theoretical guarantees
- Algorithmic issues

# Main Topics

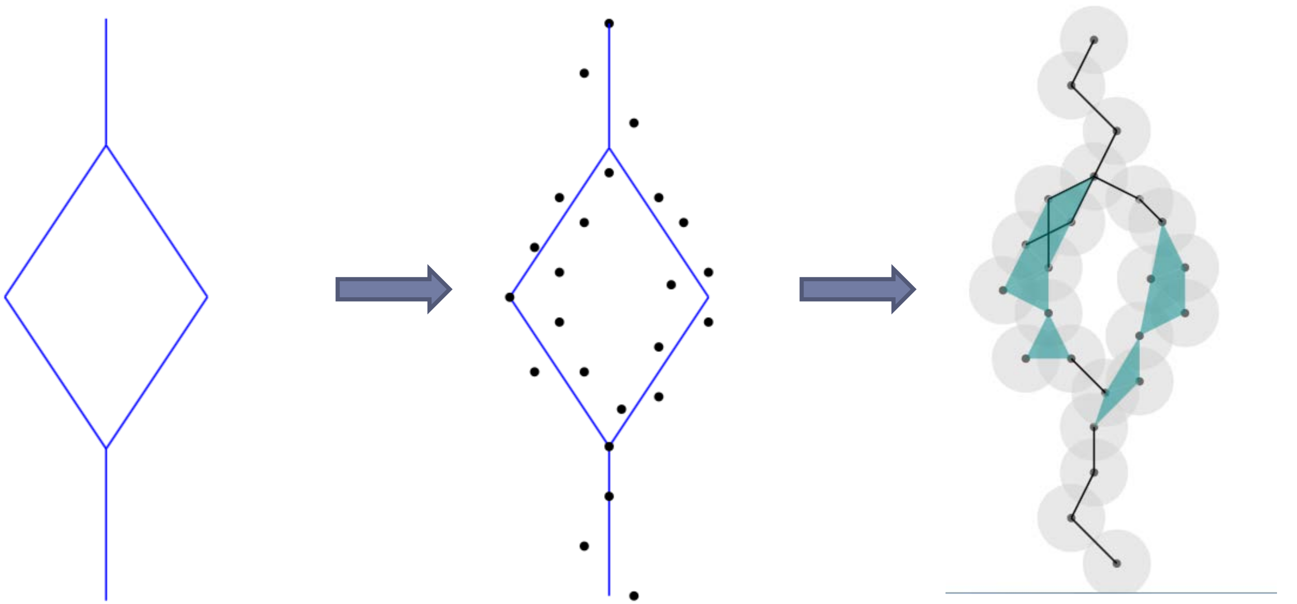


- From PCDs to simplicial complexes
- Sampling conditions
- Topological inferences

# Outline

- From PCDs to simplicial complexes
  - Delaunay, Cech, Vietoris-Rips, witness complexes
  - Graph induced complex
- Sampling conditions
  - Local feature size, and homological feature size
- Topology inference
  - Homology inference
  - Handling noise
  - *Approximating cycles of shortest basis of the first homology group*
  - *Approximating Reeb graph*

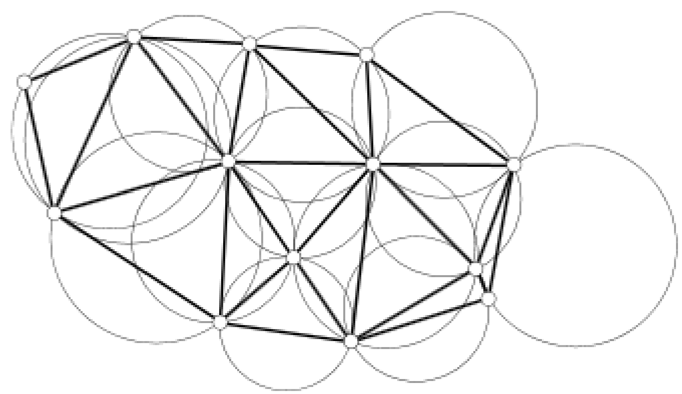
# From PCD to Simplicial Complexes



Choice of Simplicial Complexes
to build on top of point cloud data

# Delaunay Complex

- Given a set of points $P = \{ p_1, p_2, \dots, p_n\} \subset R^d$
- Delaunay complex $Del(P)$
  - A simplex $\sigma = \{p_{i_0}, p_{i_1}, \dots, p_{i_k}\}$ is in $Del(P)$ if and only if
    - There exists a ball $B$ whose boundary contains vertices of $\sigma$, and that the interior of $B$ contains no other point from $P$.



# Delaunay Complex

- Many beautiful properties
  - Connection to Voronoi diagram
- Foundation for surface reconstruction and meshing in 3D
  - *[Dey, Curve and Surface Reconstruction, 2006],*
  - *[Cheng, Dey and Shewchuk, Delaunay Mesh Generation, 2012]*

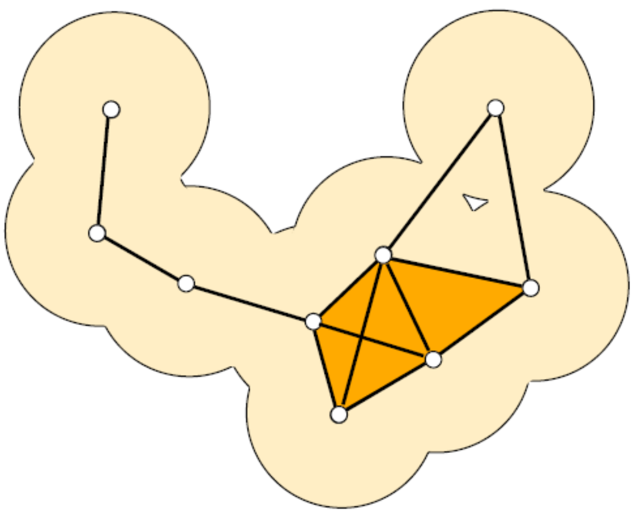
- However,
  - Computationally very expensive in high dimensions

# Čech Complex

- Given a set of points $P = \{ p_1, p_2, \dots, p_n\} \subset R^d$
- Given a real value $r > 0$, the *Čech complex* $C^r(P)$ is the nerve of the set $\{ B(p_i, r)\}_{i\in[1,n]}$
  - where $B_r(p_i) = B(p_i, r) = \{ x \in R^d \mid d(p_i, x) < r \}$
  - I.e, a simplex $\sigma = \{p_{i_0}, p_{i_1}, \dots, p_{i_k}\}$ is in $C^r(P)$ if

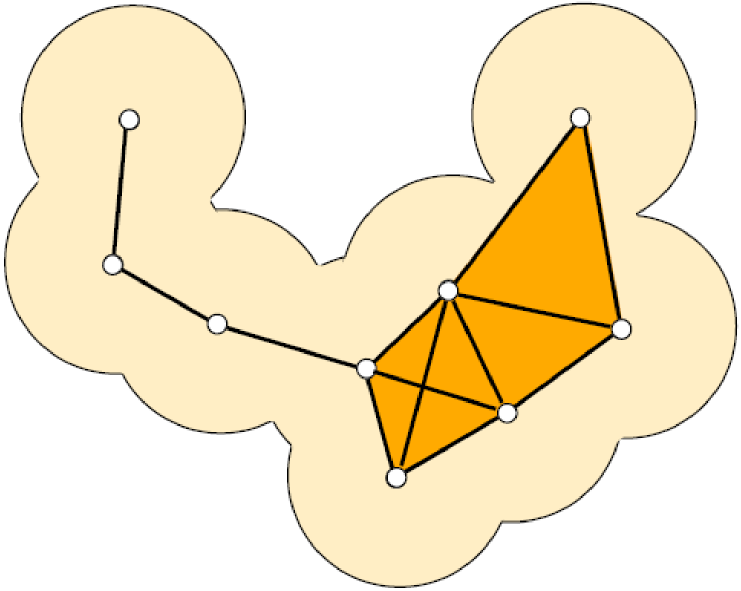
$$\bigcap_{0\le j\le k} B_r\left(p_{i_j}\right) \neq \Phi.$$

- The definition can be extended to a finite sample $P$ of a metric space.



# Rips Complex

- Given a set of points $P = \{ p_1, p_2, \dots, p_n\} \subset R^d$
- Given a real value $r > 0$, the *Vietoris-Rips (Rips) complex* $R^r(P)$ is:
  - $\{ \left(p_{i_0}, p_{i_1}, \dots, p_{i_k}\right) \mid B_r\left(p_{i_l}\right) \cap B_r\left(p_{i_j}\right) \neq \emptyset, \forall\, l, j \in [0, k\,]\}.$



# Rips and Čech Complexes

- Relation in general metric spaces
  - $C^r(P) \subseteq R^r(P) \subseteq C^{2r}(P)$
  - Bounds better in Euclidean space
- Simple to compute
- Able to capture geometry and topology
  - We will make it precise shortly
  - One of the most popular choices for topology inference in recent years

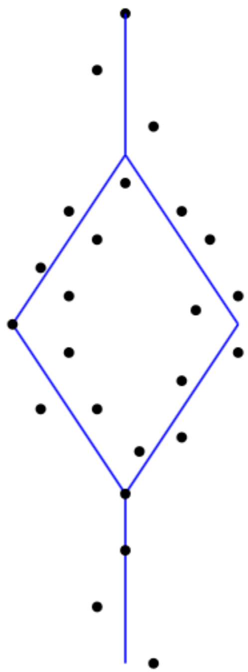
- However:
  - Huge sizes
  - Computation also costly

# Witness Complexes

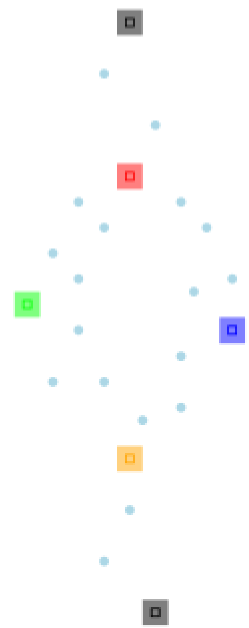
- A simplex $\sigma = \{q_0, \dots, q_k\}$ is *weakly witnessed* by a point x if $d(q_i, x) \leq d(q, x)$ for any $i \in [0,\ k]$ and $q \in Q \setminus \{q_0, \dots, q_k\}$.
  - is *strongly witnessed* if in addition $d(q_i, x) = d\left(q_j, x\right), \forall i, j \in [0, k]$
- Given a set of points $P = \{\, p_1, p_2, \dots, p_n\} \subset R^d$ and a subset $Q \subseteq P$ ,
  - The *witness complex* $W(Q, P)$ is the collection of simplices with vertices from $Q$ whose all subsimplices are weakly witnessed by a point in $P$.
  - *[de Silva and Carlsson, 2004] [de Silva 2003]*
  - Can be defined for a general metric space
  - $P$ does not have to be a finite subset of points

# Intuition

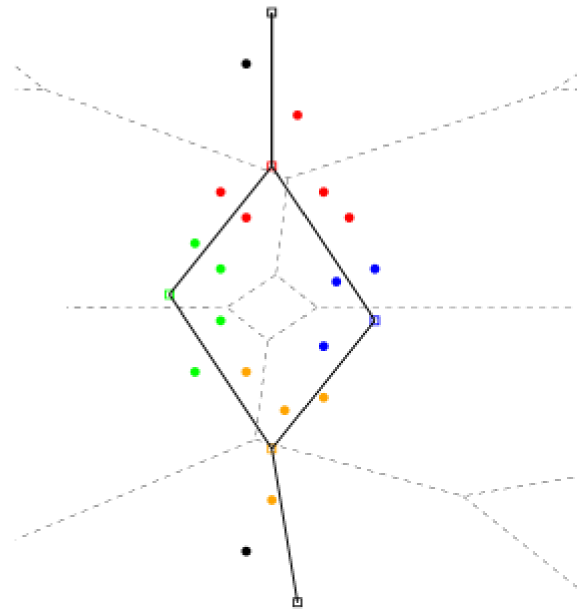
- $L$: landmarks from $P$, a way to subsample.



$P$



$L \subseteq P$



$W(L, P)$

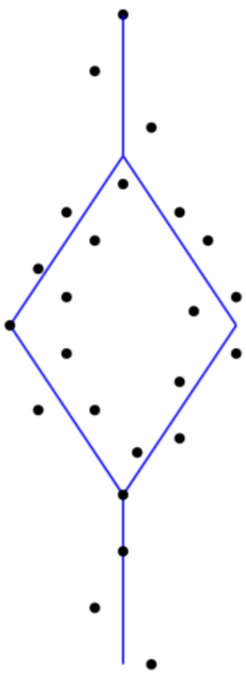
# Witness Complexes

- Greatly reduce size of complex
  - Similar to Delaunay triangulation, remove redundancy
- Relation to Delaunay complex
  - $W(Q,P) \subseteq Del\ Q$ if $Q \subseteq P \subset R^d$
  - $W\left(Q,R^d\right) = Del\ Q$
  - $W(Q,M) = Del|_M\ Q$ if $M \subseteq R^d$ is a smooth 1- or 2-manifold
    - *[Attali et al, 2007]*
- However,
  - Does not capture full topology easily for high-dimensional manifolds

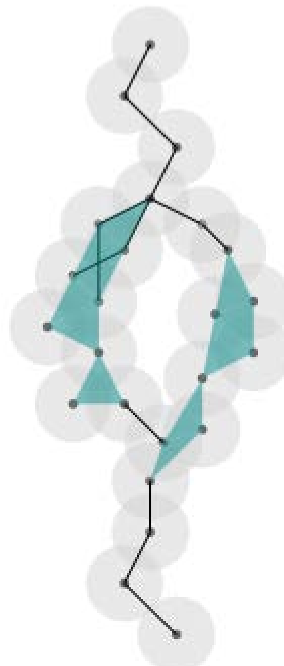
# Remark

- Rips complex
  - Capture homology when input points are sampled dense enough
  - But too large in size
- Witness complex
  - Use a subsampling idea
  - Reduce size tremendously
  - May not be easy to capture topology in high-dimensions
- Combine the two ?
  - Graph induced complex
  - *[Dey, Fan, Wang, SoCG 2013]*

# Subsampling

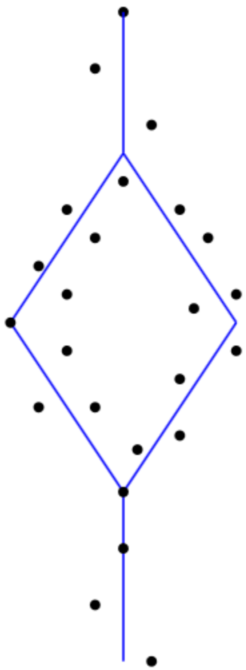


$P$

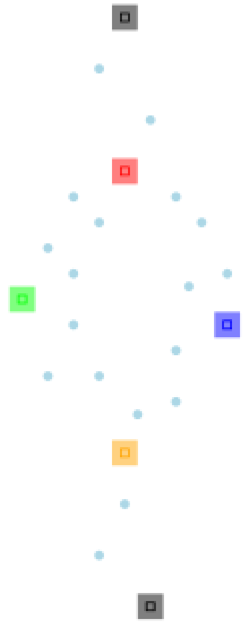


$R^{\epsilon}(P)$

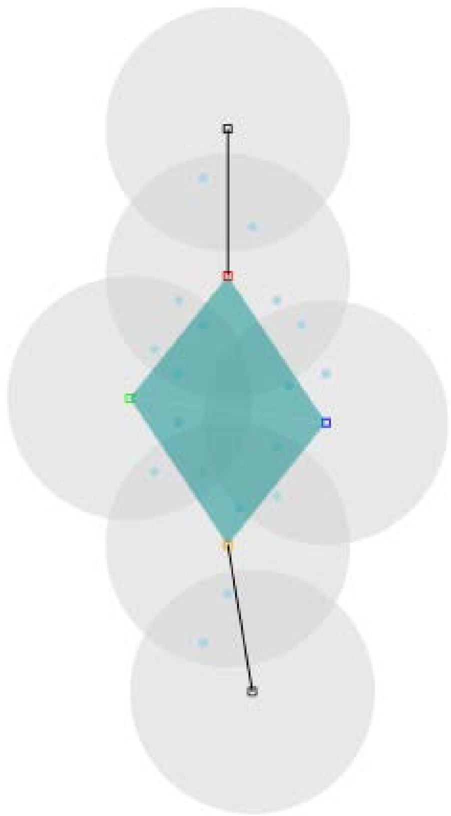
# Subsampling -cont



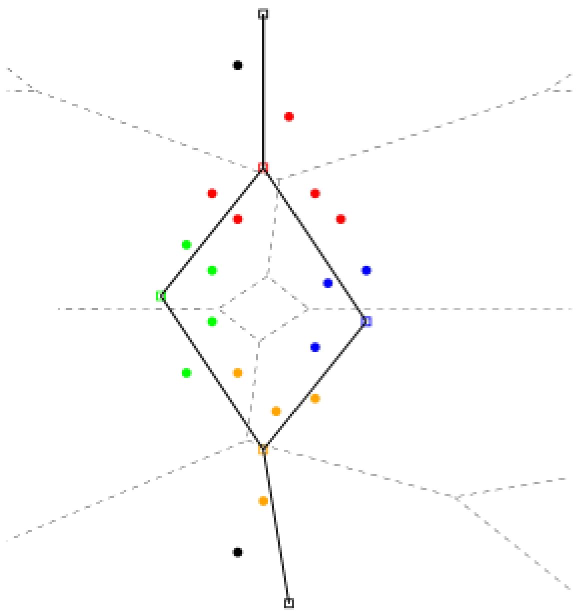
$P$



$Q \subseteq P$

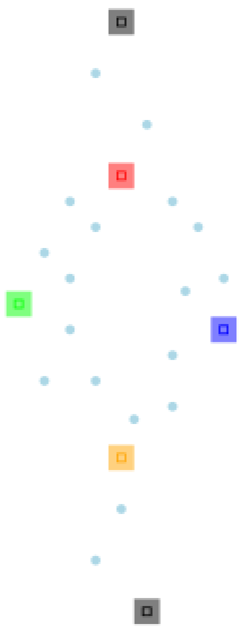


$R^r(Q)$

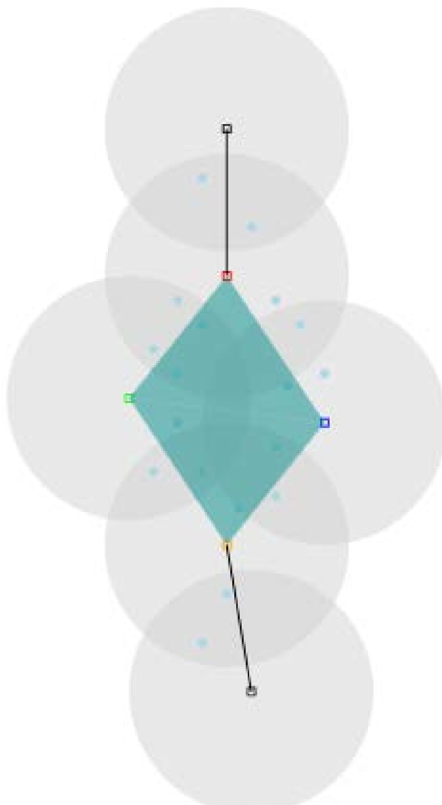


$W(Q, P)$

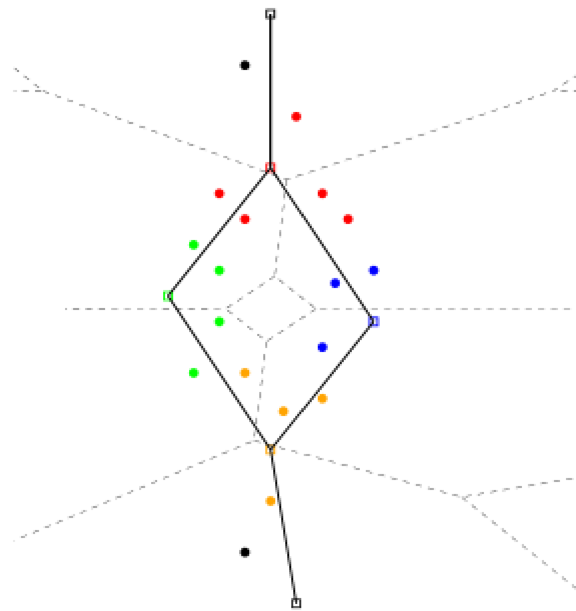
# Subsampling -cont



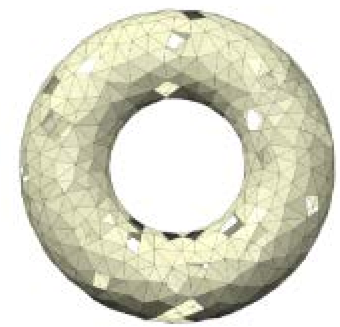
$Q \subseteq P$



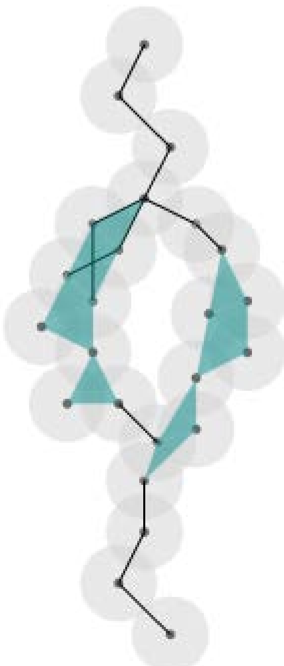
$R^r(Q)$



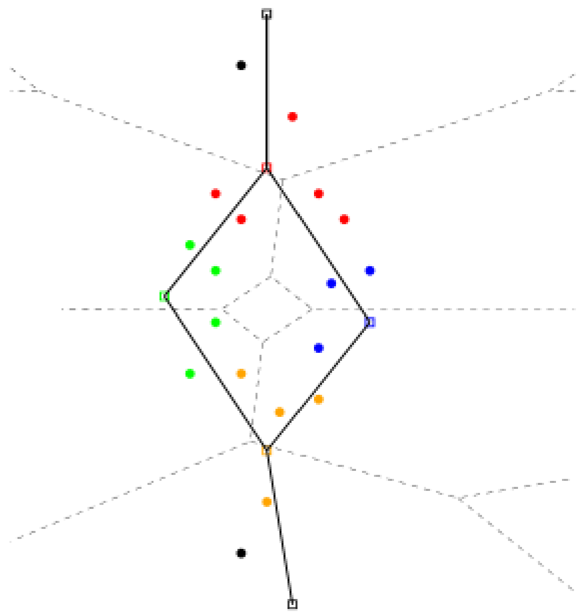
$W(Q, P)$



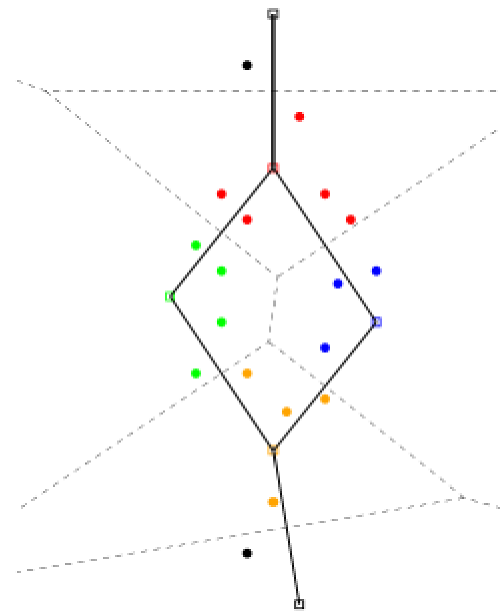
# Subsampling - cont



$R^{\epsilon}(P)$



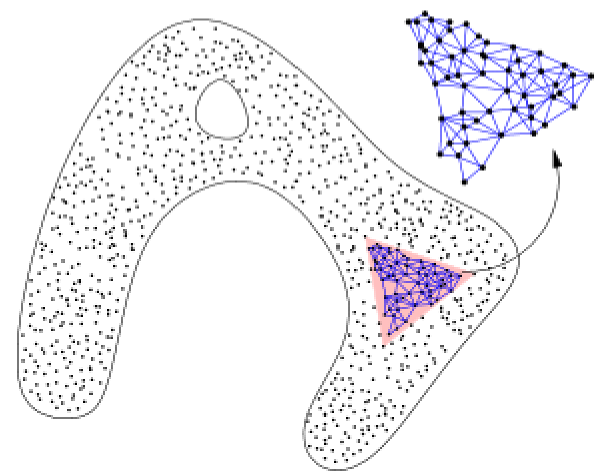
$W(Q,P)$



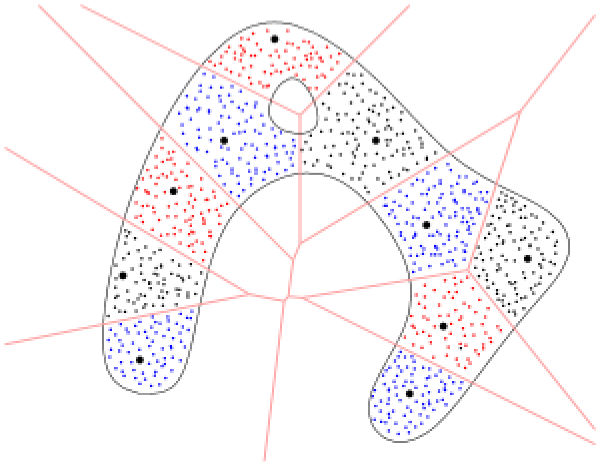
$\mathcal{G}^{r}(Q,P)$

# Graph Induced Complex

- *[Dey, Fan, Wang, SoCG 2013]*
- $P$: finite set of points
- $(P, d)$: metric space
- $G(P)$: a graph

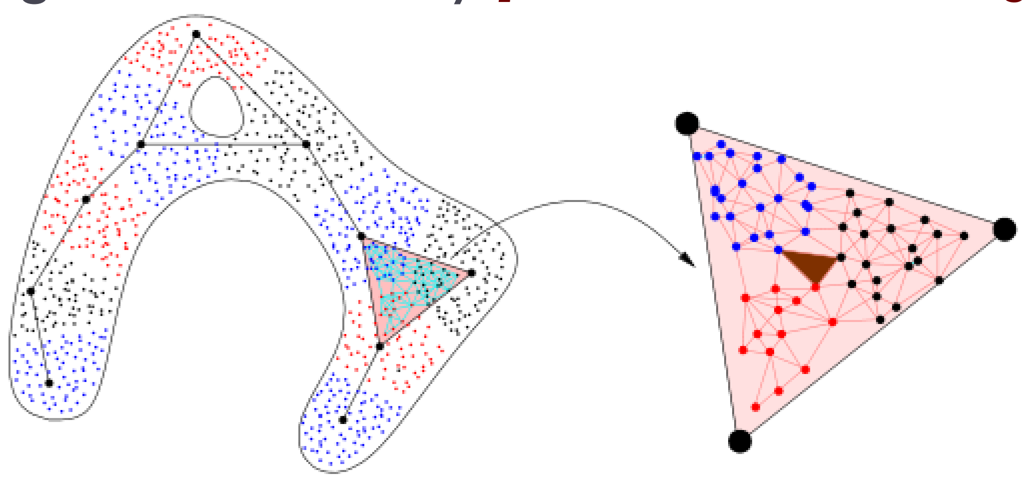


- $Q \subset P$: a subset
- $\pi(p)$: the closest point of $p \in P$ in $Q$



# Graph Induced Complex

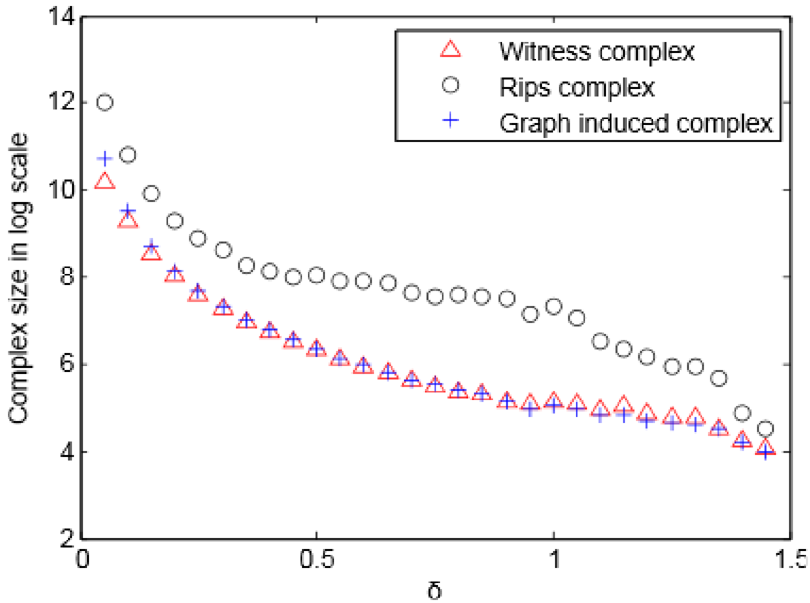
- *Graph induced complex* $\mathcal{G}(P, Q, d)$: $\{q_0, \dots, q_k\} \subseteq Q$
  - if and only if there is a (k+1)-clique in $G(P)$ with vertices $p_0, \dots, p_k$ such that $\pi(p_i) = q_i$, for any $i \in [0, k]$.
  - Similar to geodesic Delaunay *[Oudot, Guibas, Gao, Wang, 2010]*

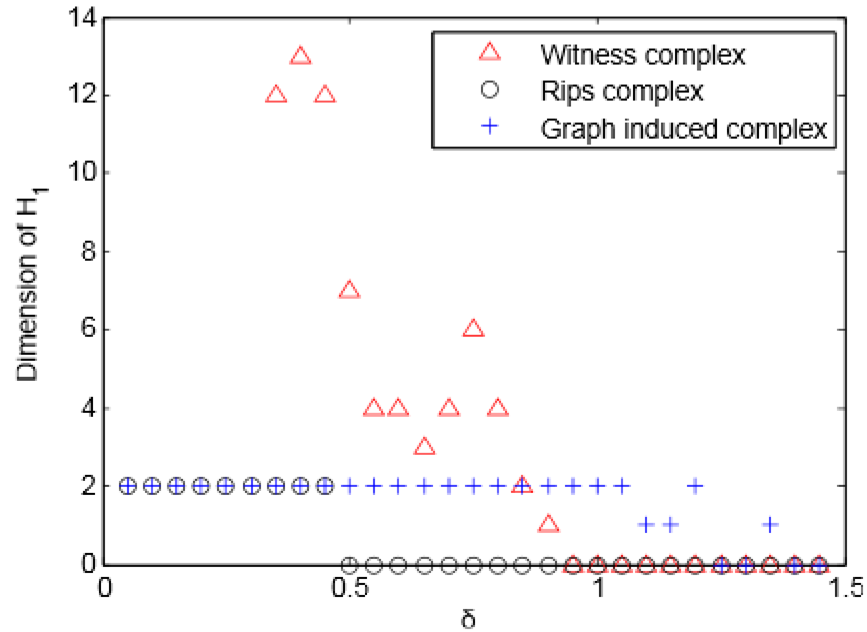


- Graph induced complex depends on the metric $d$:
  - Euclidean metric
  - Graph based distance $d_G$

# Graph Induced Complex

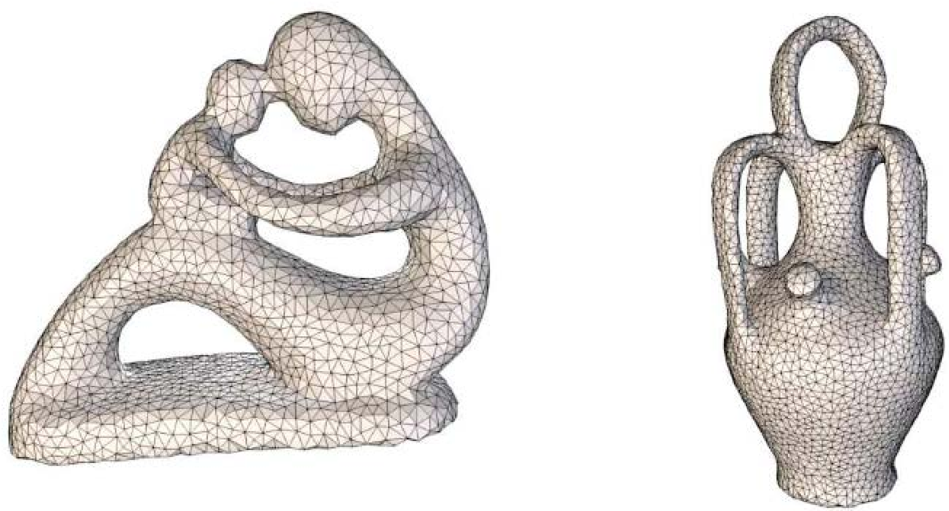
- Small size, but with homology inference guarantees
- In particular:
  - $H_1$ inference from a lean sample



# Graph Induced Complex

- Small size, but with homology inference guarantees
- In particular:
  - $H_1$ inference from a lean sample
  - Surface reconstruction in $R^3$



  - Topological inference for compact sets in $R^d$ using persistence

# Outline

- From PCDs to simplicial complexes
  - Delaunay,  Cech, Vietoris-Rips, witness complexes
  - Graph induced complex
- Sampling conditions
  - Local feature size, and homological feature size
- Topology inference
  - Homology inference
  - *Handling noise*

# Sampling Conditions

# Motivation

- Theoretical guarantees are usually obtained when input points $P$ sampling the hidden domain "well enough".

- Need to quantify the "wellness".
- Two common ones based on:
  - Local feature size
  - Weak feature size

# Distance Function

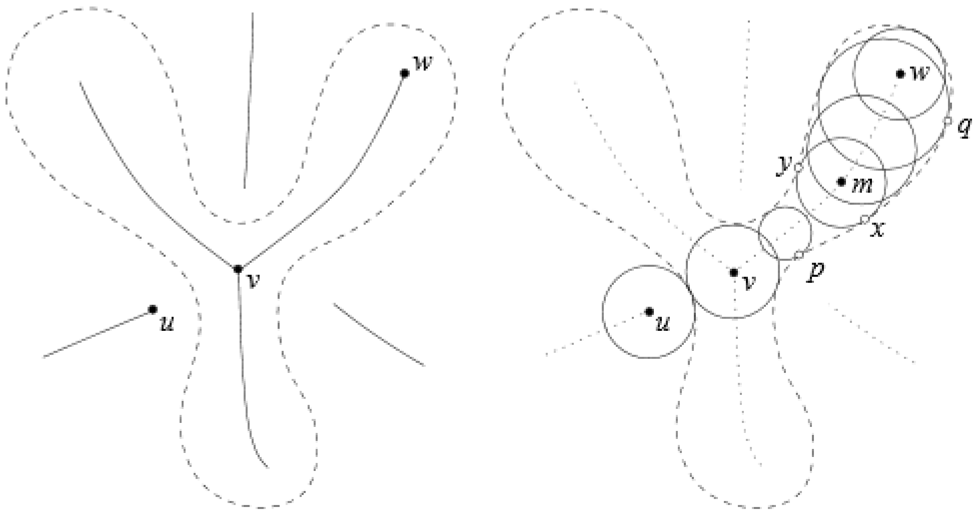
- $X \subset R^d$:  a compact subset of $R^d$
- Distance function $d_X: R^d \to R^+ \cup \{0\}$
  - $d_X(x) = \min_{y \in X} d(x, y)$
  - $d_X$ is a 1-Lipschitz function

- $X^\alpha$:  $\alpha$-offset of $X$
  - $X^\alpha = \{y \in R^d \mid d(y, X) \le \alpha\}$

- Given any point $x \in R^d$
  - $\Gamma(x) := \{\, y \in X \mid d(x, y) = d_X(x) \,\}$

# Medial Axis

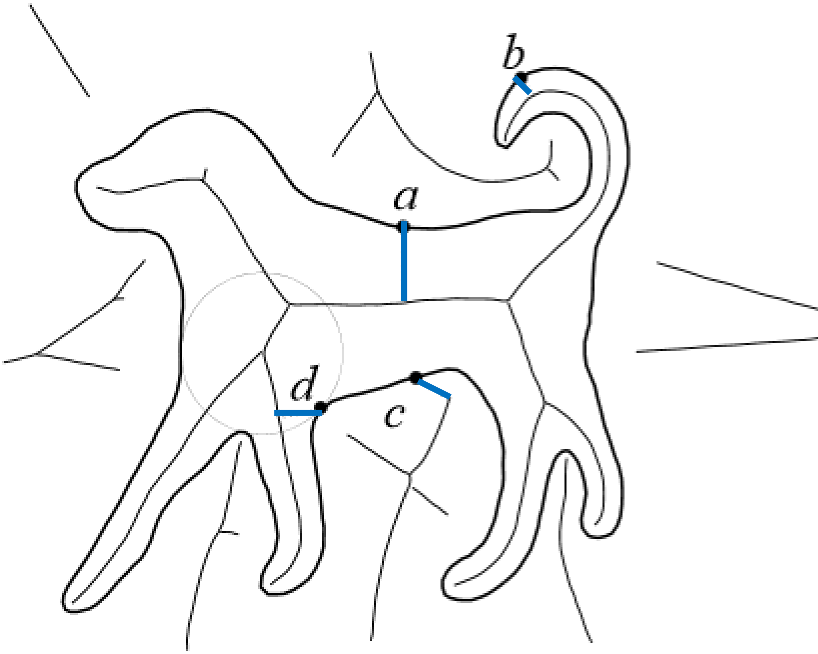
- The *medial axis* $\Sigma$ of $X$ is the closure of the set of points $x \in R^d$ such that $|\Gamma(x)| \geq 2$
  - $|\Gamma(x)| \geq 2$ means that there is a medial ball $B_r(x)$ touching $X$ at more than 1 point and whose interior is empty of points from $X$.




Courtesy of [Dey, 2006]

# Local Feature Size

- The local feature size $lfs(x)$ at a point $x \in X$ is the distance of $x$ to the medial axis $\Sigma$ of $X$
  - That is, $lfs(x) = d(x, \Sigma)$
- This concept is adaptive
  - Large in a place without "features"
- Intuitively:
  - We should sample more densely if local feature size is small.
- The *reach* $\rho(X) = \inf_{x \in X} lfs(x)$




Courtesy of [Dey, 2006]

# Gradient of Distance Function

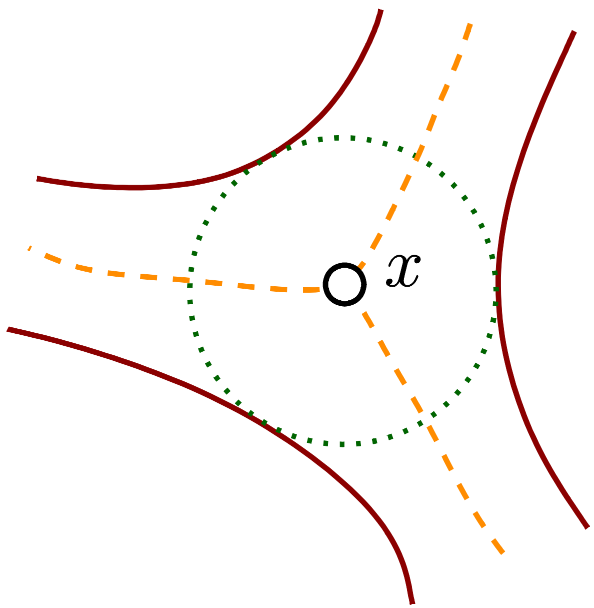
- Distance function not differentiable on the medial axis
- Still can define a generalized concept of gradient
  - *[Lieutier, 2004]*
- For $x \in R^d \setminus X$,
  - Let $c_X(x)$ and $r_X(x)$ be the center and radius of the smallest enclosing ball of point(s) in $\Gamma(x)$
  - The *generalized gradient* of distance function

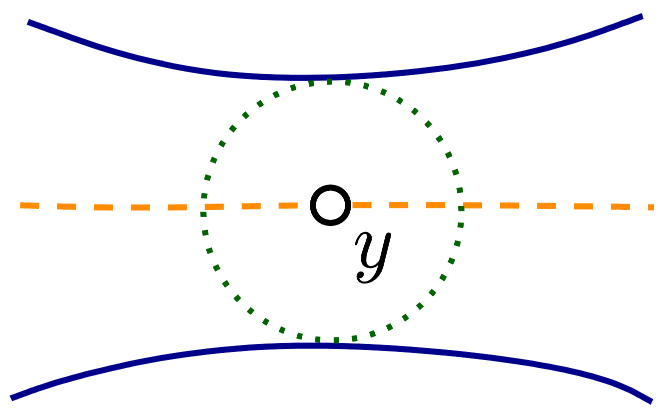
$$\nabla_X(x) = \frac{x - c_X(x)}{r_X(x)}$$

- Flow lines induced by the generalized gradient
- Examples:

# Critical Points

- A critical point of the distance function is a point whose generalized gradient $\nabla_X(x)$ vanishes
- A critical point is either in $X$ or in its medial axis $\Sigma$



# Weak Feature Size

- Given a compact $X \subset R^d$, let $C \subset R^d$ denote
  - the set of critical points of the distance function $d_X$ that are not in $X$

- Given a compact $X \subset R^d$, the *weak feature size* is
  - $wfs(X) = \inf_{x \in X} d(x, C)$
- Equivalently,
  - $wfs(X)$ is the infimum of the positive critical value of $d_X$

- $\rho(X) \leq wfs(X)$

# Why Distance Field?

- Theorem [Offset Homotopy] [*Grove'93*]

  If $0 < \alpha < \alpha'$ are such that there is no critical value of $d_X$ in the closed interval $[\alpha, \alpha']$, then $X^{\alpha'}$ deformation retracts onto $X^{\alpha}$ . In particular, $H(X^{\alpha}) \cong H\left(X^{\alpha'}\right)$.

- Remarks:
  - For the case of compact set $X$, note that it is possible that $X^{\alpha}$, for sufficiently small $\alpha > 0$, may not be homotopy equivalent to $X^0 = X$.
  - Intuitively, by above theorem, we can approximate $H(X^{\alpha})$ for any small positive $\alpha$ from a thickened version (offset) of $X^{\alpha}$.
  - The sampling condition makes sure that the discrete sample is sufficient to recover the offset homology.

# Typical Sampling Conditions

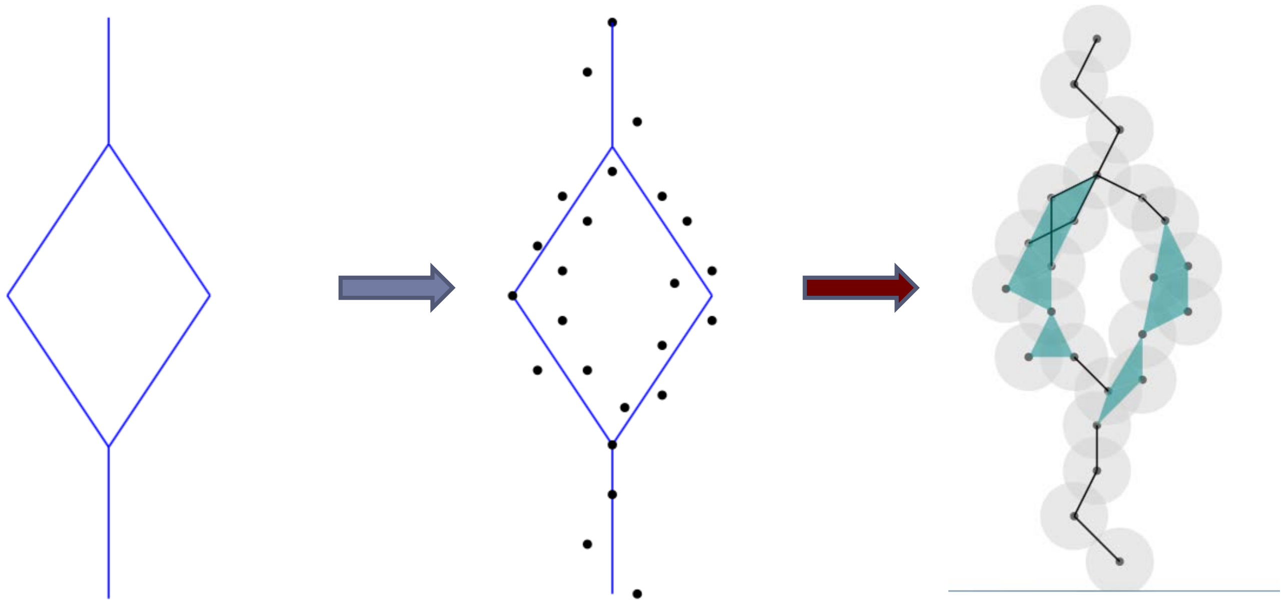
- Hausdorff distance $d_H(A, B)$ between two sets $A$ and $B$
  - infinumum value $\alpha$ such that $A \subseteq B^{\alpha}$ and $B \subseteq A^{\alpha}$
- No noise version:
  - A set of points $P$ is an $\epsilon$-sample of $X$
    if $P \subset X$ and $d_H\ (P, X) \leq \epsilon$
- With noise version:
  - A set of points $P$ is an $\epsilon$-sample of $X$ if $d_H\ (P, X) \leq \epsilon$
- Theoretical guarantees will be achieved when $\epsilon$ is small with respect to local feature size or weak feature size

# Outline

- From PCDs to simplicial complexes
  - Delaunay, Cech, Vietoris-Rips, witness complexes
  - Graph induced complex
- Sampling conditions
  - Local feature size, and homological feature size
- Topology inference
  - Homology inference
  - *Handling noise*

# Homology Inference from PCD

# Problem Setup



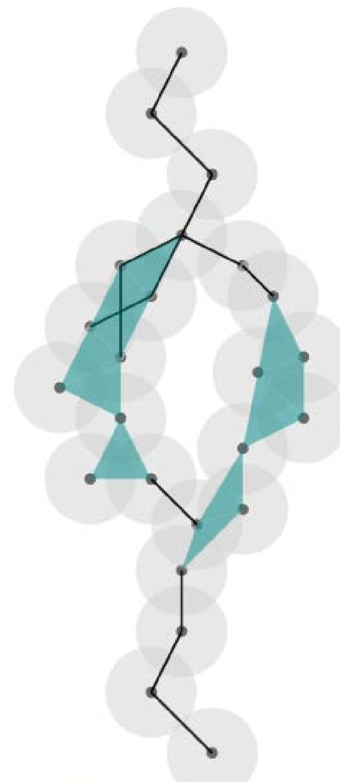
- A hidden compact $X$ (or a manifold $M$)
- An $\epsilon$-sample $P$ of $X$
- Recover homology of $X$ from some complex built on $P$
  - will focus on Čech complex and Rips complex

# Union of Balls

- $X^{\alpha} = \bigcup_{x \in X} B(x, \alpha)$
- $P^{\alpha} = \bigcup_{p \in \mathrm{P}} B(p, \alpha)$

- Intuitively, $P^{\alpha}$ approximates offset $X^{\alpha}$

- The Čech complex $C^{\alpha}(P)$ is the Nerve of $P^{\alpha}$
- By Nerve Lemma, $C^{\alpha}(P)$ is homotopy equivalent to $P^{\alpha}$



# Smooth Manifold Case

- Let $X$ be a smooth manifold embedded in $R^d$

Theorem [Niyogi, Smale, Weinberger]

Let $P \subset X$ be such that $d_H(X, P) \leq \epsilon$. If $2\epsilon \leq \alpha \leq \sqrt{\frac{3}{5}}\rho(X)$, there is a deformation retraction from $P^\alpha$ to $X$.

Corollary A

Under the conditions above, we have

$$\mathrm{H(X)} \cong H(P^\alpha) \cong H\big(C^\alpha(P)\big).$$

- How about using Rips complex instead of Čech complex?
- Recall that

$$C^r(P) \subseteq R^r(P) \subseteq C^{2r}(P)$$

inducing

$$H(C^r) \to H(R^r) \to H(C^{2r})$$

- Idea *[Chazal and Oudot 2008]*:
  - Forming interleaving sequence of homomorphism to connect them with the homology of the input manifold $X$ and its offsets $X^\alpha$

# Convert to Cech Complexes

- Lemma A *[Chazal and Oudot, 2008]*:
- The following diagram commutes:

$$\begin{tikzcd} H(P^\alpha) \arrow[r, "i_*"] \arrow[d, "h_*"'] & H(P^\beta) \arrow[d, "h_*"] \\ H(C^\alpha) \arrow[r, "i_*"] & H(C^\beta) \end{tikzcd}$$

Corollary B

Let $P \subset X$ be s.t. $d_H(X, P) \leq \epsilon$. If $2\epsilon \leq \alpha \leq \alpha' \leq \sqrt{\frac{3}{5}}\rho(X)$, $\mathrm{H(X)} \cong H(C^\alpha) \cong H(C^\beta)$ where the second isomorphism is induced by inclusion.

# From Rips Complex

- Lemma B:

  Given a sequence $A \to B \to C \to D \to E$ of homomorphisms between finite dimensional vector spaces, if $rank(A \to$

- Rips and Čech complexes:

$$C^{\alpha}(P) \subseteq R^{\alpha}(P) \subseteq C^{2\alpha}(P) \subseteq R^{2\alpha}(P) \subseteq C^{4\alpha}(P)$$
$$\Rightarrow H(C^{\alpha}) \to H(R^{\alpha}) \to H(C^{2\alpha}) \to H(R^{2\alpha}) \to H(C^{4\alpha})$$

- Applying Lemma B

$$rank\big(H(R^{\alpha}) \to H(R^{2\alpha})\big) = rank\big(H(C^{2\alpha})\big) = rank\big(H(X)\big)$$

# The Case of Compact

- In contrast to Corollary A, now we have the following (using Lemma B).

- Lemma C *[Chazal and Oudot 2008]*:
  Let $P \subset R^d$ be a finite set such that $d_H(X, P) < \epsilon$ for some $\epsilon < \frac{1}{4} wfs(X)$. Then for all $\alpha, \beta \in [\epsilon, wfs(X) - \epsilon]$ such that $\beta - \alpha \geq 2\epsilon$, and for all $\lambda \in \left(0, wfs(X)\right)$, we have $H\left(X^\lambda\right) \cong image(i_*)$, where $i_*: H(P^\alpha) \to H\left(P^\beta\right)$ is the homomorphism between homology groups induced by the canonical inclusion $i: P^\alpha \to P^\beta$.

# The Case of Compacts

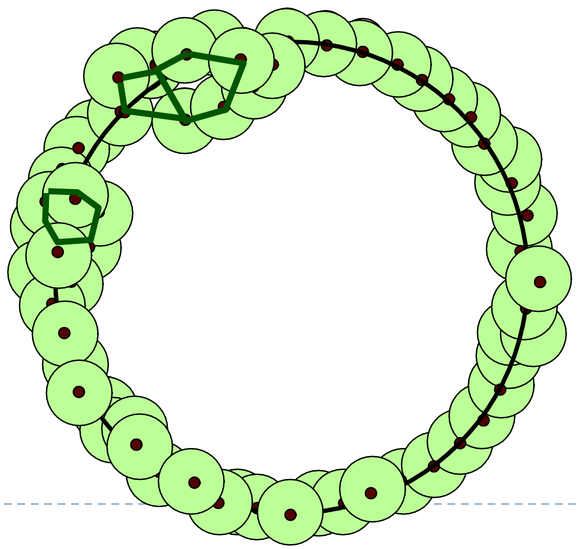
- One more level of interleaving.
- Use the following extension of Lemma B:

  Given a sequence $A \to B \to C \to D \to E \to F$ of homomorphisms between finite dimensional vector spaces, if $rank(A \to F) = rank(C \to D)$, then $rank(B \to E) = rank(C \to D)$.

- Theorem [Homology Inference] *[Chazal and Oudot 2008]*:

  Let $P \subset R^d$ be a finite set such that $d_H(X, P) < \epsilon$ for some $\epsilon < \frac{1}{9} wfs(X)$. Then for all $\alpha \in \left[2\epsilon, \frac{1}{4}(wfs(X) - \epsilon)\right]$ all $\lambda \in (0, wfs(X))$, we have $H(X^\lambda) \cong image(j_*)$, where $j_*$ is the homomorphism between homology groups induced by canonical inclusion $j: R^\alpha \to R^{4\alpha}$.

- Theorem [Homology Inference] *[Chazal and Oudot 2008]*:
  Let $P \subset R^d$ be a finite set such that $d_H(X,P) < \epsilon$ for some $\epsilon < \frac{1}{9} wfs(X)$. Then for all $\alpha \in \left[2\epsilon, \frac{1}{4}(wfs(X) - \epsilon)\right]$ all $\lambda \in \left(0, wfs(X)\right)$, we have $H(X^\lambda) \cong image(j_*)$, where $j_*$ is the homomorphism between homology groups induced by canonical inclusion $j: R^\alpha \to R^{4\alpha}$.



# Summary of Homology Inference

- $X^\alpha$ homotopy equivalent to $X^\beta$
  - Critical points of distance field
- $P^\alpha$ approximates $X^\alpha$ (may be interleaving)
  - E.g, *[Niyogi, Smale, Weinberger, 2006], [Chazal and Oudot 2008]*
- $C^\alpha$ homotopy equivalent to $P^\alpha$
  - Nerve Lemma
- $H(C^\alpha)$ interleaves $H(X^\alpha)$ at homology level
  - *[Chazal and Oudot 2008]*
- $R^\alpha$ and $C^\alpha$ interleave
- Derive homology inference from the interleaving sequence of homomorphisms

# Outline

- From PCDs to simplicial complexes
  - Delaunay,  Cech, Vietoris-Rips, witness complexes
  - Graph induced complex

- Sampling conditions
  - Local feature size, and homological feature size

- Topology inference
  - Homology inference
  - *Handling noise*

# References


- *Weak witnesses for Delaunay triangulations of submanifolds*. D. Attali, H. Edelsbrunner, and Y. Mileyko. Proc. ACM Sympos. on Solid and Physical Modeling, 143–150 (2007)
- *Manifold reconstruction in arbitrary dimensions using witness complexes*. D. Boissonnat and L. J. Guibas and S.Y. Oudot. SoCG 194–203, (2007)
- *Efficient and Robust Topological Data Analysis on Metric Spaces*. M. Buchet, F. Chazal, S.Y. Oudot, D. Sheehy:. CoRR abs/1306.0039 (2013)
- *Geometric Inference*. F. Chazal and D. Cohen-Steiner. Tesselations in the Sciences, Springer-Verlag, (2013, to appear).
- *A sampling theory for compacts in Euclidean space* . F. Chazal, D. Cohen-Steiner, and A. Lieutier. Discrete Comput. Geom. 41, 461--479 (2009).
- *Geometric Inference for Probability Measures*. F. Chazal, D. Cohen-Steiner, Q. Mérigot. Foundations of Computational Mathematics 11(6): 733-751 (2011)
- *Stability and Computation of Topological Invariants of Solids in* $R^n$. F. Chazal, A. Lieutier. Discrete & Computational Geometry 37(4): 601-617 (2007)

# References – cont.


- *Towards persistence-based reconstruction in Euclidean spaces*. F. Chazal and S. Oudot. Proc. Annu. Sympos. Comput. Geom. (2008), 232--241.
- *Proximity of persistence modules and their diagrams*. F. Chazal, D. Cohen-Steiner, M. Glisse, L. Guibas, and S. Oudot. Proc. 25th Annu. Sympos. Comput. Geom. (2009), 237-246.
- *A weak definition of Delaunay triangulation.* Vin de Silva. CoRR cs.CG/0310031 (2003)
- *Topological estimation using witness complexes.* V. de Silva, G. Carlsson, Symposium on Point-Based Graphics,, 2004
- *Topological estimation using witness complexes.* Vin de Silva and Gunnar Carlsson. Eurographics Sympos. Point-Based Graphics (2004).
- *Curve and Surface Reconstruction :Algorithms with Mathematical Analysis*. Tamal. K. Dey. Cambridge University Press, 2006.
- *Graph induced complex on point data.* T. K. Dey, F. Fan, and Y. Wang. to appear in 29th Annual Sympos. Comput. Geom. (SoCG) 2013.

# References – cont.


- *Witnessed k-Distance*. L. J. Guibas, D. Morozov, Q. Mérigot. Discrete & Computational Geometry 49(1): 22-45 (2013)
- *Reconstructing using witness complexes*, L. J. Guibas and S.Y. Oudot. Discrete Comput. Geom. 30, 325–356 (2008).
- *Any open bounded subset of $R^n$ has the same homotopy type as its medial axis*. A. Lieutier. Computer-Aided Design 36(11): 1029-1046 (2004)
- *Finding the homology of submanifolds with high confidence from random samples*. P. Niyogi, S. Smale, and S. Weinberger. Discrete Comput. Geom., 39, 419--441 (2008).
- *Geodesic delaunay triangulations in bounded planar domains*. S. Oudot, L. J. Guibas, J. Gao and Y. Wang. ACM Trans. Alg. 6 (4): 2010.

# References – cont.


- *Minimax rates for homology inference*. S. Balakrishnan, A. Rinaldo, D. Sheehy, A. Singh, L. A. Wasserman. AISTATS 2012: 64-72
- *Local Homology Transfer and Stratification Learning*. P. Bendich, B. Wang and S. Mukherjee. ACM-SIAM Symposium on Discrete Algorithms, (2012).
- *Greedy optimal homotopy and homology generators*. J. Erickson and K. Whittlesey. Proc. Sixteenth ACM-SIAM Sympos. Discrete Algorithms (2005), 1038--1046.
- *Approximating cycles in a shortest basis of the first homology group from point data*. T. K. Dey, J. Sun, and Y. Wang. Inverse Problems, 27 (2011).
- *Reeb graphs: Approximation and Persistence*. T. K. Dey and Y. Wang., Discrete Comput. Geom (DCG) 2013.