Topological Data Analysis and Neuroscience

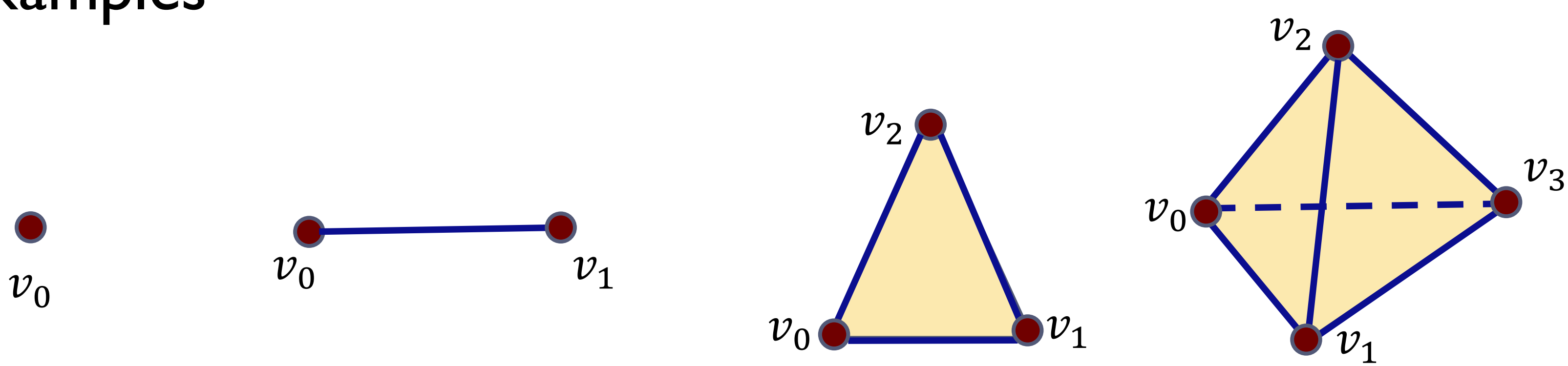
Chapter 2: Simplicial Complex

Instructor: Alex McCleary

Section 1:
Simplicial Complex

A (Geometric) Simplex

- Points $\{p_0, p_1, \dots, p_d\} \subset R^N$ are linearly independent
  - if vectors $v_i = p_i - p_0, i \in [0, d]$, are linearly independent
- Geometric p-simplex $\sigma = \{ v_0, v_1, \dots, v_p \}$
  - Convex combination of $p + 1$ *linearly-independent* points in R^N
    - $\sigma = \{ \sum_{i=0}^{p} a_i v_i \mid a_i \geq 0, \sum a_i = 1 \}$
- Examples



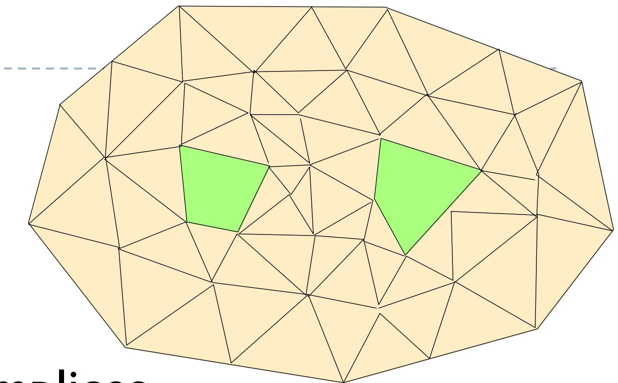
0-simplex 1-simplex 2-simplex 3-simplex

A (Geometric) Simplex

- Points $\{p_0, p_1, \dots, p_d\} \subset R^N$ are linearly independent
  - if vectors $v_i = p_i - p_0, i \in [0, d]$, are linearly independent
- Geometric *p-simplex* $\sigma = \{\, v_0, v_1, \dots, v_p\}$
  - Convex combination of $p + 1$ *linearly-independent* points in R^N
    - $\sigma = \{\, \sum_{i=0}^{p} a_i v_i \mid a_i \geq 0, \sum a_i = 1 \,\}$

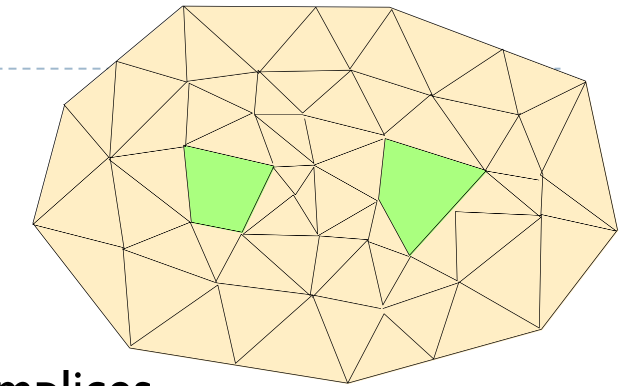
- Simplex τ formed by a subset of $\{\, v_0, v_1, \dots, v_p\}$ is called a face of σ, denoted by $\tau \subseteq \sigma$
  - τ is a proper face of σ if $\dim(\tau) = \dim(\sigma) - 1$
  - $bd\ (\sigma) =$ collection of all proper faces of σ
- For a d-simplex σ
  - $\sigma \cong R^{d+1},\ bd(\sigma) \cong S^d,\ int(\sigma) \cong R^d$

Simplicial complex



- A geometric simplicial complex K
  - A collection of simplices such that
    - If $\sigma \in K$, then any fact $\tau \subseteq \sigma$ is also in K
    - If $\sigma \cap \sigma' \neq \emptyset$, then $\sigma \cap \sigma'$ is a face of both simplices.
  - $\dim(K)$ = highest dim of any simplex in K

Simplicial complex



- A geometric simplicial complex K
  - A collection of simplices such that
    - If $\sigma \in K$, then any fact $\tau \subseteq \sigma$ is also in K
    - If $\sigma \cap \sigma' \neq \emptyset$, then $\sigma \cap \sigma'$ is a face of both simplices.
  - $\dim(K)$ = highest dim of any simplex in K
- Subcomplex
  - $L \subseteq K$ and L is a complex
- Underlying space $|K|$ of K
  - is the pointwise union of all points in all simplices of K,
  - i.e, $|K| = \bigcup_{\sigma \in K} \{x \mid x \in \sigma\}$

Abstract simplicial complex

- An (abstract) p-simplex $\sigma = \{ v_0, v_1, \dots, v_p\}$
  - a set of cardinality $p + 1$
  - A subset $\tau \subseteq \sigma$ is a face of σ
- An (abstract) simplicial complex K
  - A collection of simplices such that
    - If $\sigma \in K$, then any fact $\tau \subseteq \sigma$ is also in K
- Geometric realization of an abstract simplicial complex S
  - is a geometric simplicial complex K such that there is an isomorphism between $Vert(K)$ and $Vert(S)$ inducing an isomorphism between all simplices in K and in S

Geometric realization

- Geometric realization of S in the standard simplex $\Delta \subset R^N$ with $N = |\mathrm{Vert}(S)|$

- Theorem:
  - Any abstract simplicial complex S of dimension d has a geometric realization $K \subset R^{2d+1}$
  - Sketch of proof

- Underlying space $|S|$ of an abstract simplicial complex
  - is the underlying space of its geometric realization into the standard simplex Δ

Star and links

- Given a simplex $\tau \in K$
  - Star: $St(\tau) = \{\sigma \in K \mid \tau \subset \sigma\}$
  - Closed star: $clSt(\tau) = \bigcup_{\sigma \in St(\tau)} \{ \sigma' \mid \sigma' \subset \sigma \}$
  - Link: $Lk(\tau) = \{ \sigma \in clSt(\tau) \mid \sigma \cap \tau = \emptyset \}$

- Intuition and examples

Simplicial map

- Intuitively, analogous to continuous maps between topological spaces

- Given simplicial complexes K and L
  - a function $f: K \to L$ is a simplicial map if
    - $f\big(Vert(K)\big) \subseteq Vert(L)$
    - for any $\sigma = \{p_0, \dots, p_d\}$, $f(\sigma) = \{f(p_0), \dots, f(p_d)\}$ spans a simplex in L

- A function $f: K \to L$ is an isomorphism
  - if f is a simplicial map and it is bijective

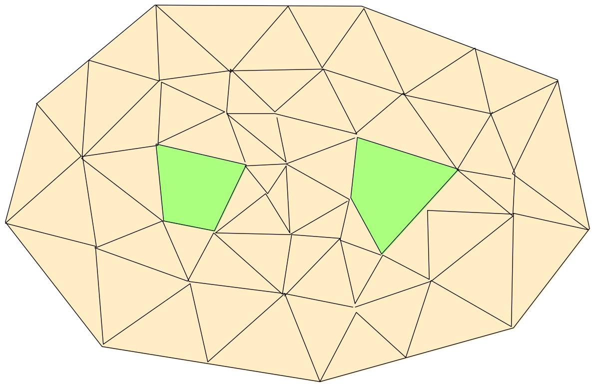
Simplicial map

- A simplicial map $f: K \to L$ induces a natural continuous function $f': |K| \to |L|$
  - s.t $f'(x) = \sum_{i\in[0,d]} a_i f(p_i)$ for $x = \sum_{i\in[0,d]} a_i p_i \in \sigma = \{p_0, \ldots, p_d\}$

- Theorem:
  - An isomorphism $f: K \to L$ induces a homeomorphism $f': |K| \to |L|$

Triangulation of a manifold

- Given a manifold (with or without boundary) M, a simplicial complex K is a triangulation of M
  - if the underlying space $|K|$ of K is homeomorphic to M



Triangulation of a manifold

- Given a manifold (with or without boundary) M, a simplicial complex K is a triangulation of M
  - if the underlying space $|K|$ of K is homeomorphic to M

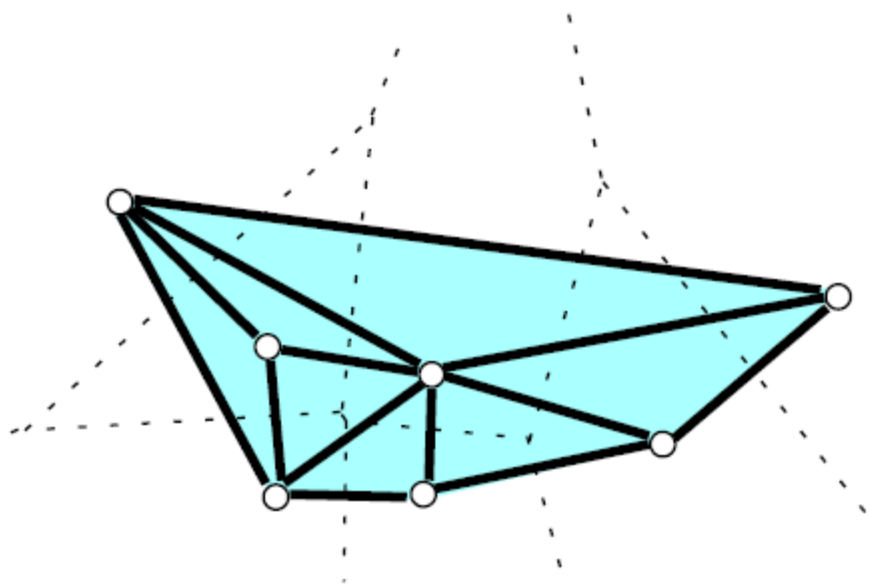
- If K is a triangulation of d-manifold M
  - then the dimension of K is also d
  - for any vertex $v \in Vert(K), St(v) \cong B_d^o \cong R^d$

- Examples of 2-manifolds / 2-complex

Section 2:
Common Complexes

Delaunay Complex

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

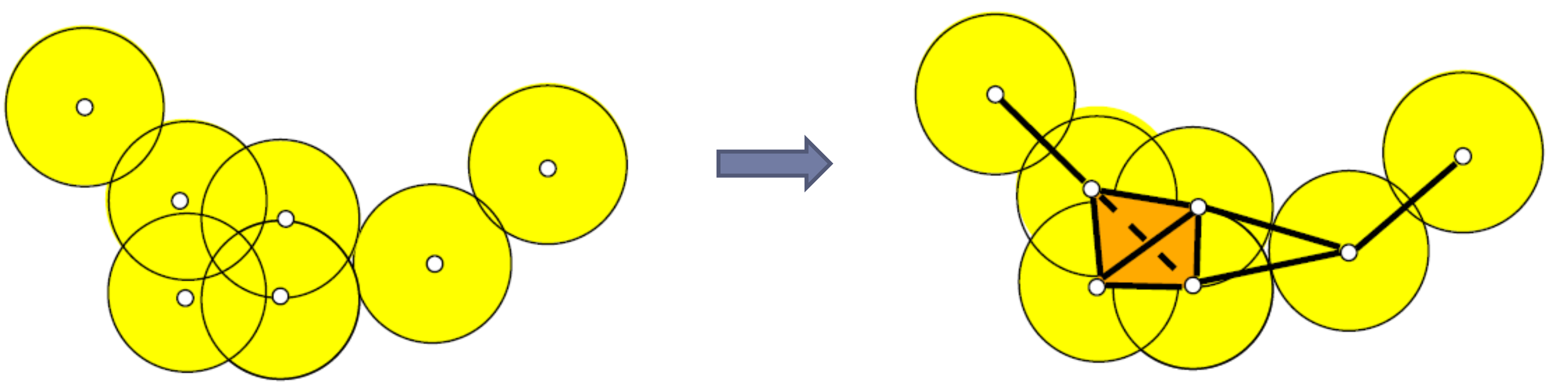


Delaunay Complex

- 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 $P = \{p_1, p_2, \ldots, p_n\} \subset R^d$
- Given a real value $r > 0$, the *Čech complex* $C^r(P)$ is the nerve of the set $\{B(p_i, r)\}_{i\in[1,n]}$
  - i.e, $\sigma = \{p_{i_0}, \ldots, p_{i_s}\} \in C\text{^}r(P)$ iff $\bigcap_{j\in[0,s]} B\left(p_{i_j}, r\right) \neq \emptyset$
- The definition can be extended to a finite sample P of a metric space.



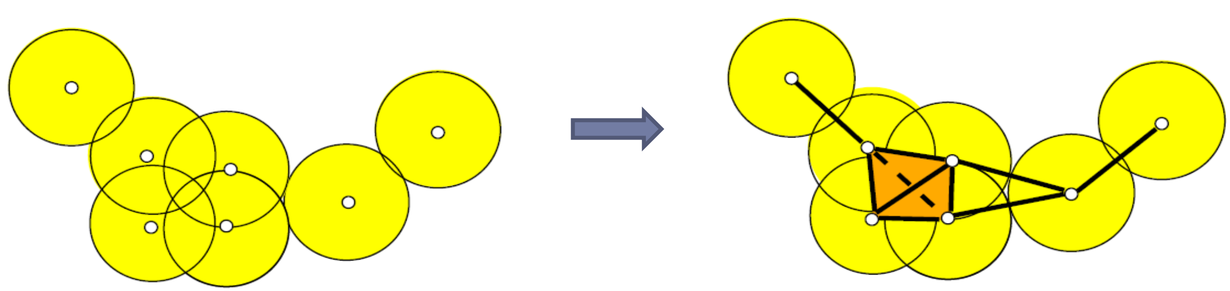
Nerves

- Given a finite set F, its nerve complex $Nrv(F)$ is
  - defined as all non-empty subset of F with non-empty common intersection
  - i.e, $Nrv(F) = \{ X \subseteq F \mid \bigcap_{\sigma \in X} \sigma \neq \emptyset \}$
- Hence Čech complex $C^r(P)$
  - is the nerve of $F = \{ B(p, r) \mid p \in P \}$
  - i.e, $C^r(P) = Nrv(F)$

Nerve Lemma

- **Nerve Lemma:**
  - Let F be a finite set of closed, convex set in R^d. Then $Nrv(F) \simeq |F|$, that is, $Nrv(F)$ is homotopy equivalent to $|F|$.

- **Corollary:**
  - $C^r(P) \simeq \bigcup_{p \in P} B(p, r)$, i.e, $C^r(P)$ is homotopy equivalent to the union of r-balls around points in P

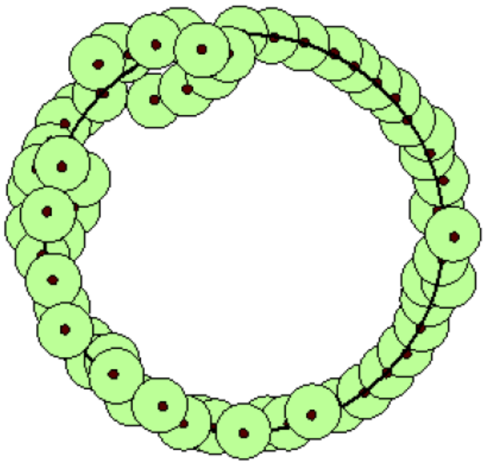


Nerve Lemma

- Nerve Lemma:
  - Let F be a finite set of closed, convex set in R^d. Then $Nrv(F) \simeq |F|$, that is, $Nrv(F)$ is homotopy equivalent to $|F|$.

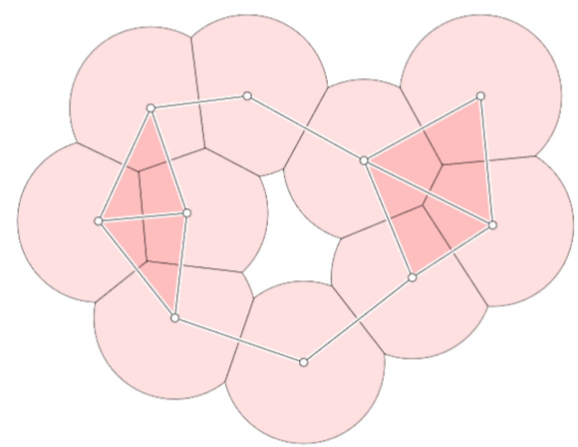
- Corollary:
  - $C^r(P) \simeq \bigcup_{p \in P} B(p, r)$, i.e, $C^r(P)$ is homotopy equivalent to the union of r-balls around points in P

- Given a set of points P
  - approximating a hidden domain M
  - $U^r(P) = \bigcup_{p \in P} B(p, r)$ approximates M
  - $C^r(P)$ approximates $U^r(P)$



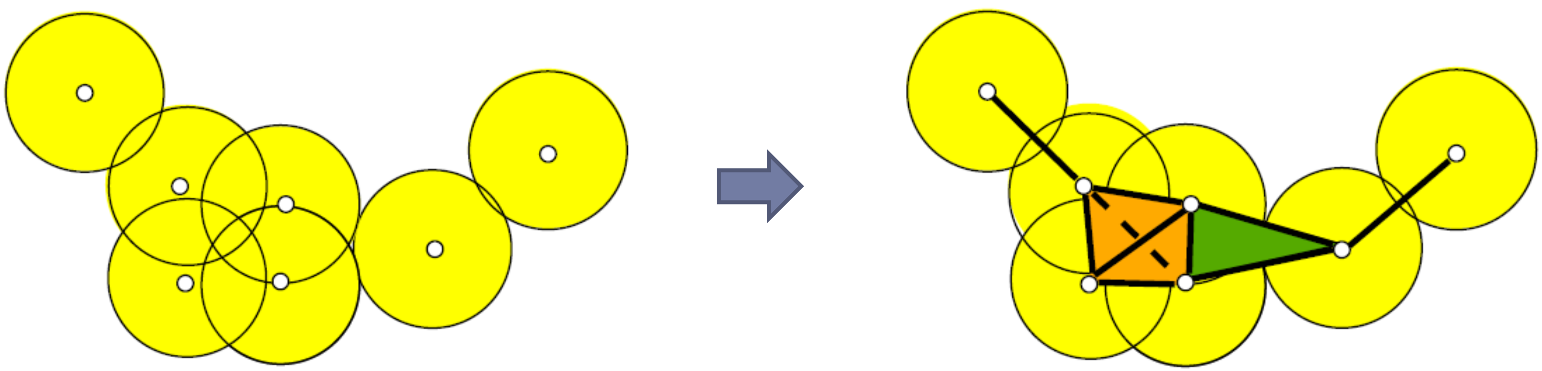
More on Čech

- Given a set of points $P \subset R^d$
  - $C^r(P)$ could have simplex of dimension larger than d
  - often only d-skeleton of $C^r(P)$ is needed
    - as $U^r(P)$ has trivial topology beyond dimension d
- Čech and Delaunay
  - Delaunay complex: $Del(P) = Nrv(\{Vor(p) \mid p \in P\})$
  - α-complex: $Del^r(P) = Nrv(Vor(p) \cap B(p,r) \mid p \in P \})$
  - $Del^{r(P)} \subseteq C^r(P)$
  - $C^r(P)$ typically has much larger size



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:
  - $\{ (p_{i_0}, p_{i_1}, \dots, p_{i_k}) \mid B_r(p_{i_l}) \cap B_r\left(p_{i_j}\right) \neq \emptyset, \forall\, l, j \in [0, k\,] \}$.



- Rips complex shares the same edge set as the Cech complex w.r.t same r.
- It is the clique complex induced by its edge set.

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
  - One of the most popular choices for topology inference in recent years

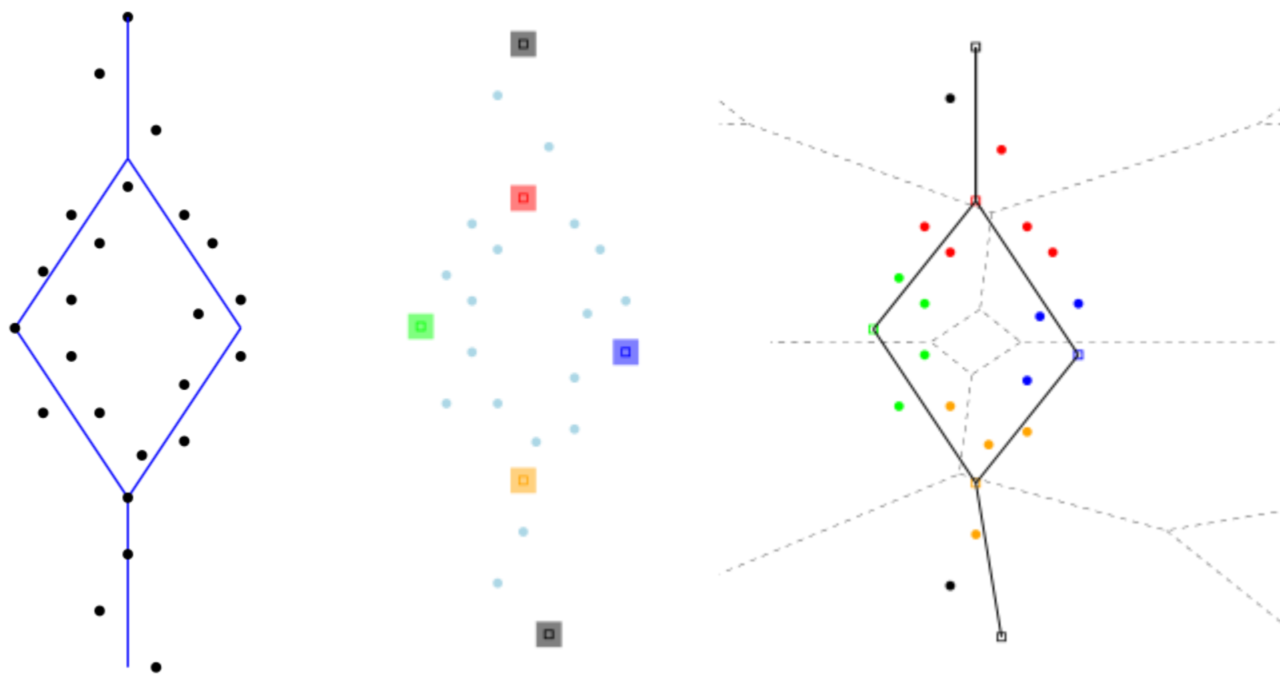
- However:
  - Huge sizes
  - Computation also costly

Witness Complexes

- A simplex $\sigma = \{q_0, \ldots, q_k\}$ is *weakly witnessed* by a point x if $d(q_i, x) \leq d(q, x)$ for any $i \in [0,\ k]$ and $q \in Q \setminus \{q_0, \ldots, q_k\}$.
  - is *strongly witnessed* if in addition $d(q_i, x) = d(q_j, x), \forall i, j \in [0, k]$
- Given a set of points $P = \{p_1, p_2, \ldots, p_n\} \subset R^d$ 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

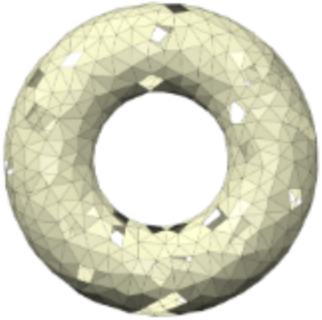
- L: landmarks from P, a way to subsample.



P $\qquad$ $L \subseteq P$ $\qquad$ $W(L, P)$

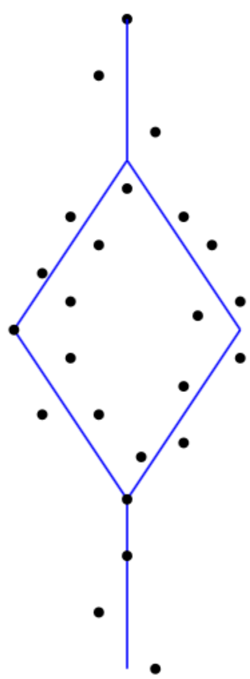
Witness Complexes

- Greatly reduce size of complex
  - Similar to Delaunay triangulation, remove redundancy
- Relation to Delaunay complex
  - $W(Q,P) \subseteq Del\ Q$ if $Q \subseteq P \subset R^d$
  - $W(Q,R^d) = Del\ Q$
  - $W(Q,M) = Del|_M\ Q$ if $M \subseteq R^d$ is a smooth 1- c
    - *[Attali et al, 2007]*

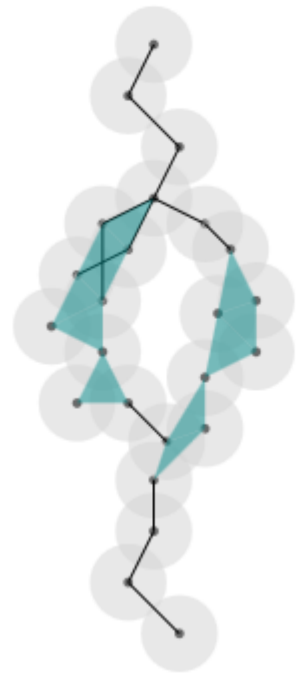


- However,
  - Does not capture full topology easily for high-dimensional manifolds

Subsampling

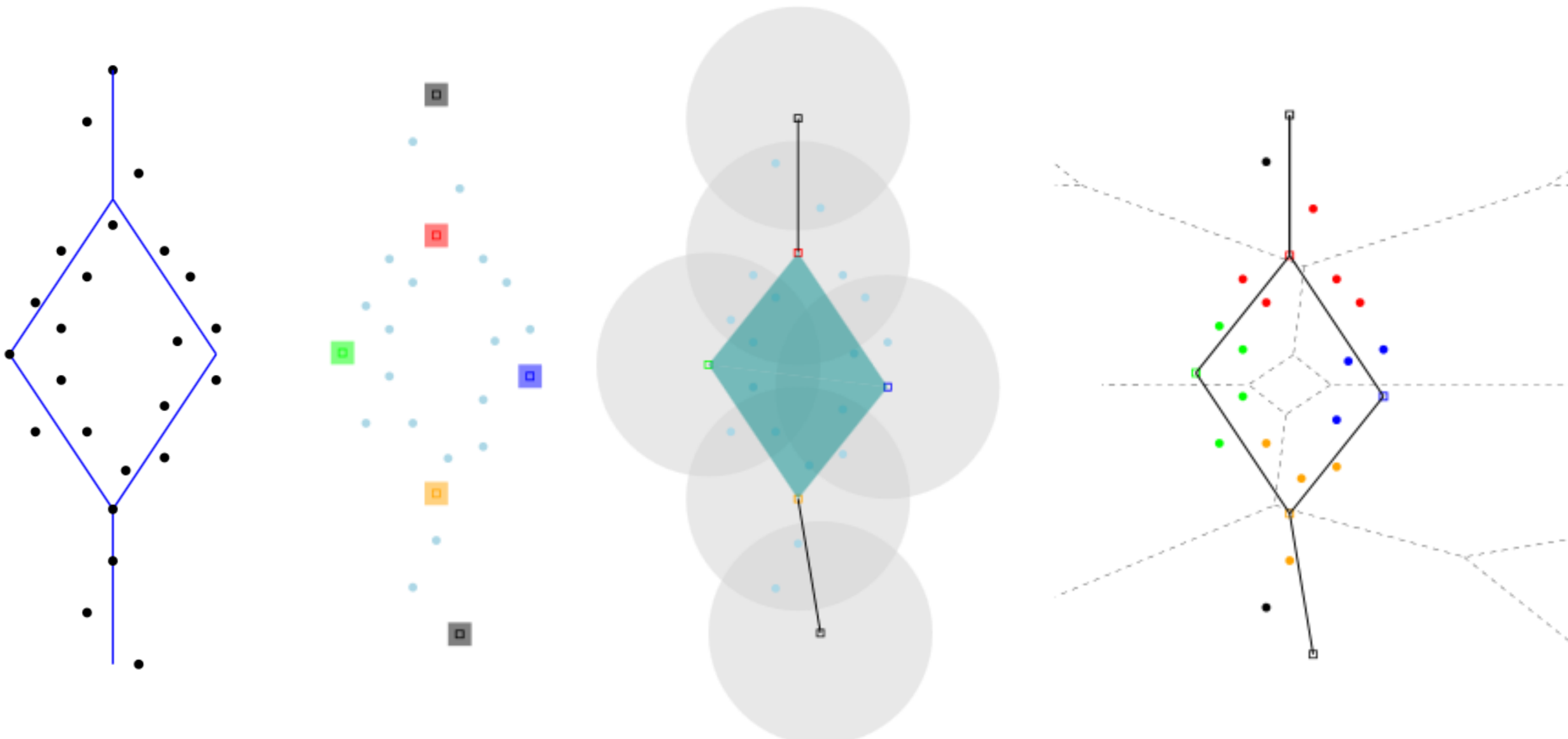


P



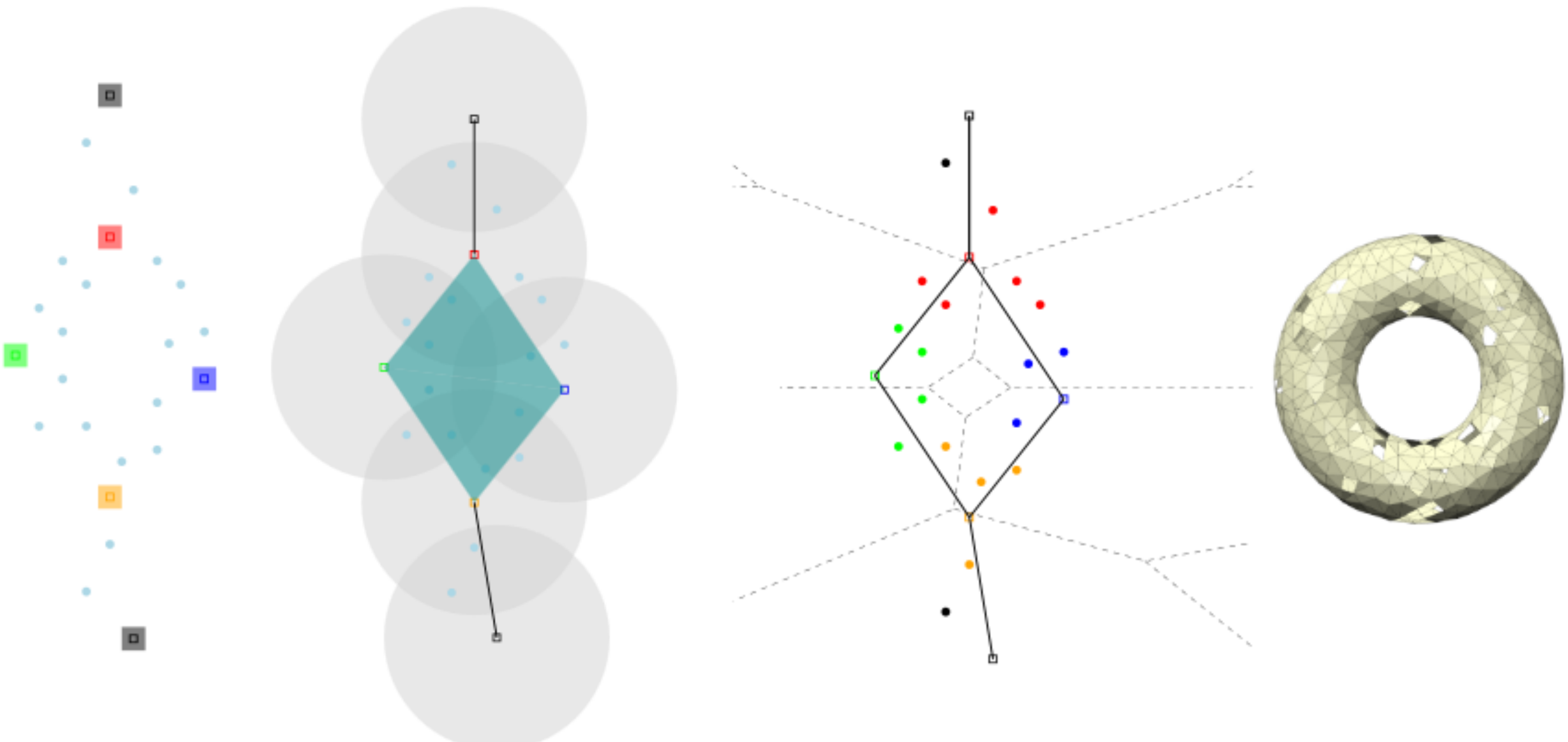
$R^{\epsilon}(P)$

Subsampling -cont



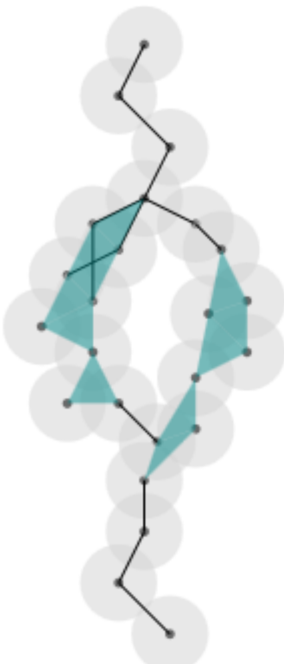
P $\quad Q \subseteq P$ $\quad R^r(Q)$ $\quad W(Q,P)$

Subsampling -cont

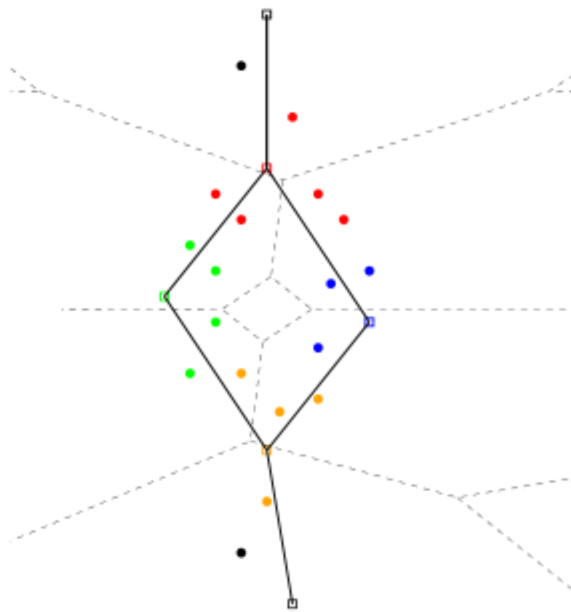


$Q \subseteq P$ $\quad$ $R^r(Q)$ $\quad$ $W(Q,P)$

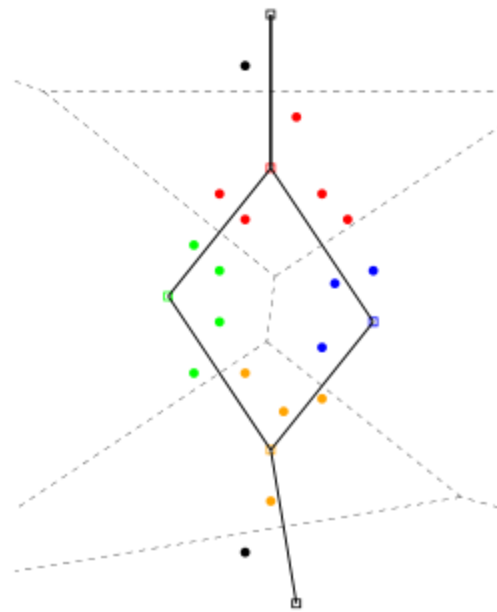
Subsampling - cont



$R^{\epsilon}(P)$



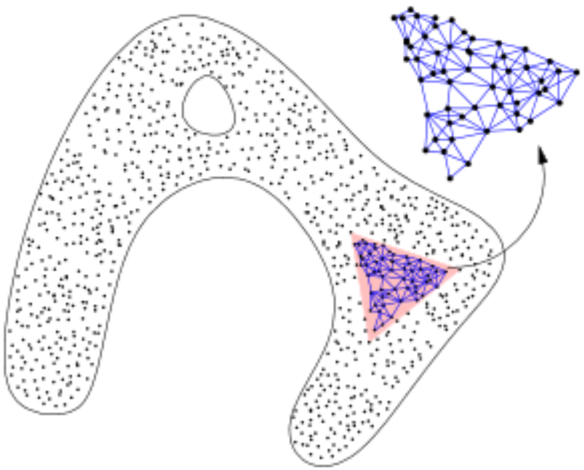
$W(Q,P)$



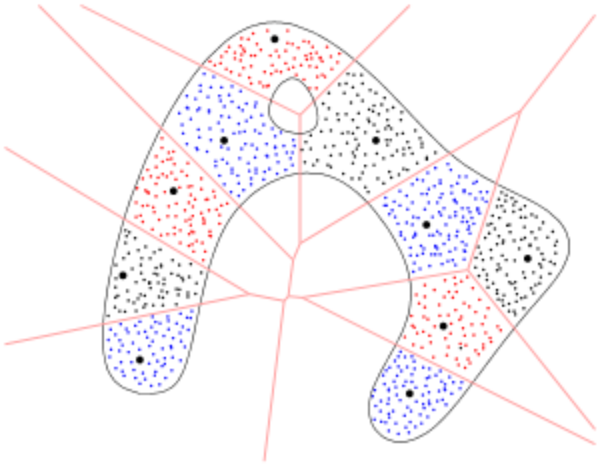
$\mathcal{G}^{r}(Q,P)$

Graph Induced Complex

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

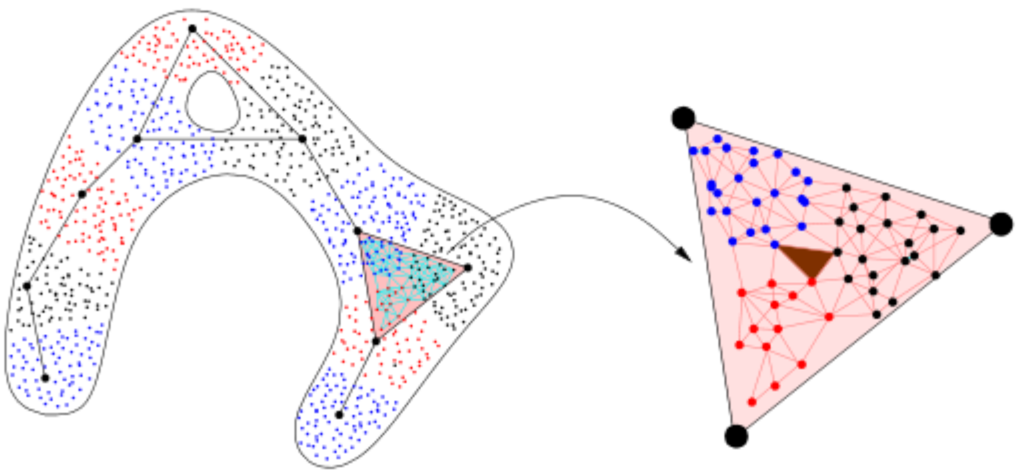


- $Q \subset P$: a subset
- $\pi(p)$: the closest point of $p \in P$ in Q



Graph Induced Complex

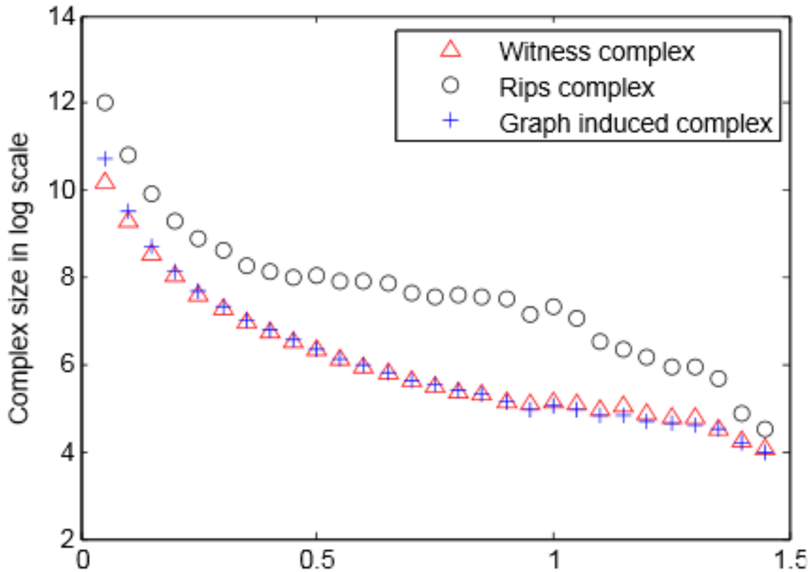
- *Graph induced complex* $\mathcal{G}(P, Q, d)$: $\{q_0, \ldots, q_k\} \subseteq Q$
  - if and only if there is a (k+1)-clique in $G(P)$ with vertices $p_0, \ldots, p_k$ such that $\pi(p_i) = q_i$, for any $i \in [0, k]$.

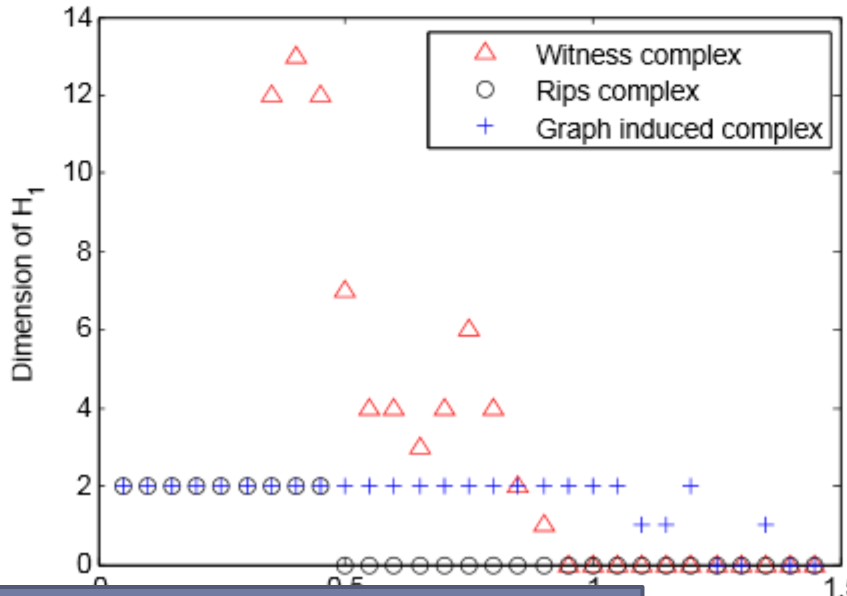


- Graph induced complex depends on the metric d:
  - Euclidean metric
  - Graph based distance d_G

Graph Induced Complex

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

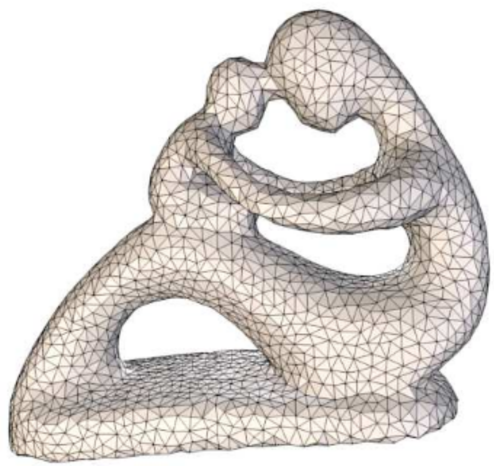


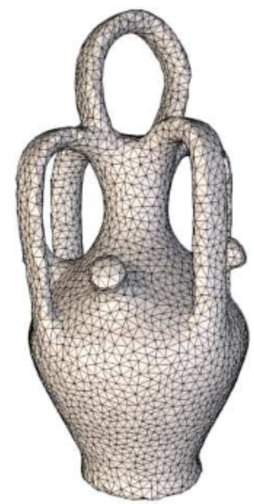


GIC can also be used as a way to sparsify graphs while maintaining global structure.

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