CTDA Reading Group Chapter 2.1 – 2.3

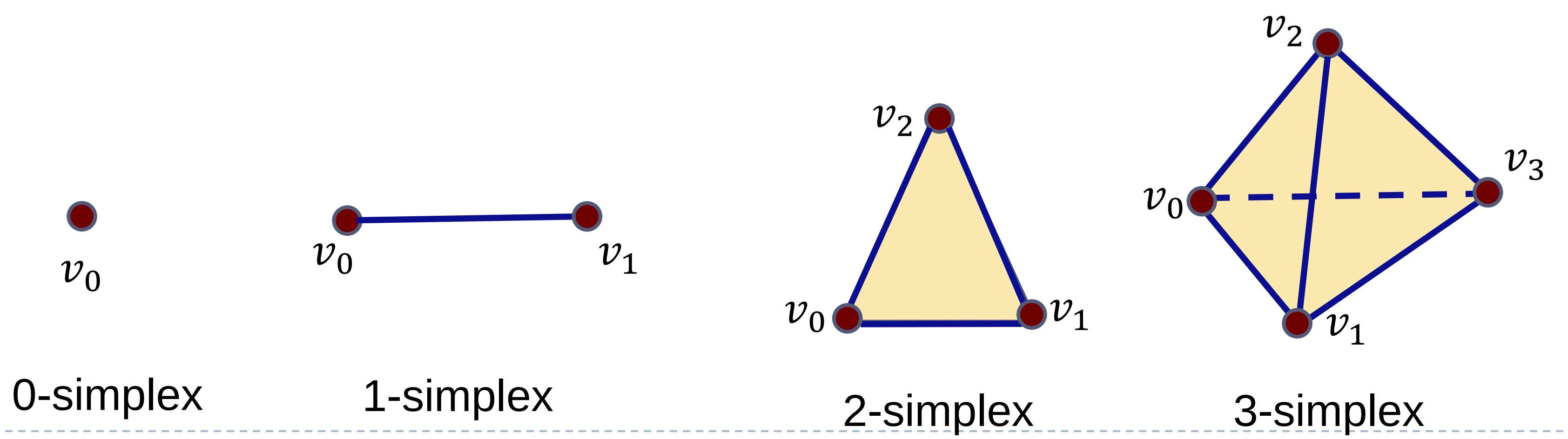
Anping Zhang

Lecture Notes From Course DSD 291 Taught by Prof. Yusu Wang

2.1 Simplicial Complex

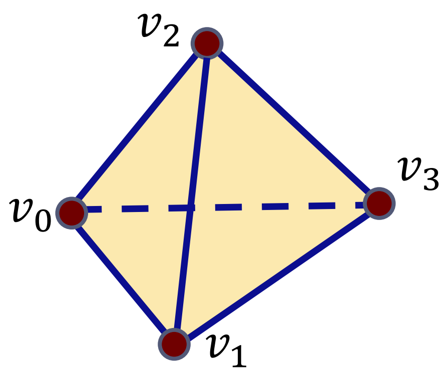
A (Geometric) Simplex

- Points $\{p_0, p_1, \dots, p_d\} \subset R^N$ are (affinely) 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 $p + 1$ *affinely-independent* points in R^N
    - $\sigma = \{ \sum_{i=0}^{p} a_i v_i \mid a_i \geq 0, \sum a_i = 1 \}$
- Examples



A (Geometric) Simplex

- Points $\{p_0, p_1, \dots, p_d\} \subset R^N$ are (affinely) 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 $p + 1$ *affinely-independent* points in R^N
    - $\sigma = \{ \sum_{i=0}^{p} a_i v_i \mid a_i \geq 0, \sum a_i = 1 \}$

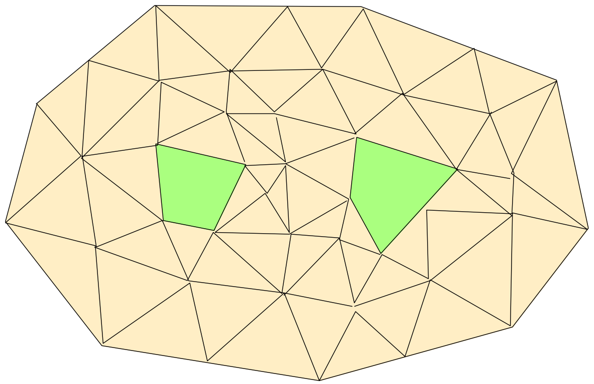


3-simplex

- Simplex τ formed by a subset of $\{ v_0, v_1, \dots, v_p\}$ is called a face of σ, denoted by $\boldsymbol{\tau} \subseteq \sigma$
  - A ***proper face*** $of\ \sigma\ is\ a\ simplex\ that\ is\ the\ convex\ hull\ of\ a\ proper\ subset\ of\ P;$
  - $(i.e. any\ face\ except\ \sigma)$
  - $The\ (k - 1) - faces\ of\ \sigma\ are\ called$ ***facets*** $of\ \sigma$ (σ has k + 1 facets)

Simplicial complex

- A geometric simplicial complex K
  - A collection of simplices such that
    - If $\sigma \in K$, then any face $\tau \subseteq \sigma$ 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
- The ***p*-skeleton** of K consists of all simplices in K of dimension at most p

- Underlying space $|K|$ of K
  - is the pointwise union of all points in all simplices of K,
  - i.e, $|K| = \cup_{\sigma \in K} \{x \mid x \in \sigma\}$

Abstract simplicial complex

- An (abstract) p-simplex $\sigma = \{ v_0, v_1, \ldots, v_p\}$
  - 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

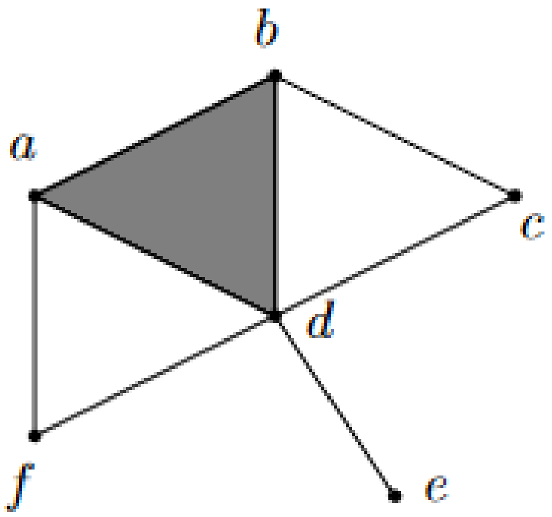
- Geometric realization of S in the standard simplex $\Delta \subset R^N$ with $N = |\mathrm{Vert}(S)|$

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

- 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: $Lk(\tau) = \{ \sigma \in clSt(\tau) \mid \sigma \cap \tau = \emptyset \}$



- $\mathrm{St}(a) = \{\{a\}, \{a,b\}, \{a,d\}, \{a,f\}, \{a,b,d\}\}$, $\overline{\mathrm{St}}(a) = \mathrm{St}(a) \cup \{\{b\}, \{d\}, \{f\}, \{b,d\}\}$
- $\mathrm{St}(f) = \{\{f\}, \{a,f\}, \{d,f\}\}$, $\overline{\mathrm{St}}(f) = \mathrm{St}(f) \cup \{\{a\}, \{d\}\}$
- $\mathrm{St}(\{a,b\}) = \{\{a,b\}, \{a,b,d\}\}$, $\overline{\mathrm{St}}(\{a,b\}) = \mathrm{St}(\{a,b\}) \cup \{\{a\}, \{b\}, \{d\}, \{a,d\}, \{b,d\}\}$
- $\mathrm{Lk}(a) = \{\{b\}, \{d\}, \{f\}, \{b,d\}\}$, $\mathrm{Lk}(f) = \{\{a\}, \{d\}\}$, $\mathrm{Lk}(\{a,b\}) = \{\{d\}\}$.

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
- 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 $x = \sum_{i \in [0,d]} a_i p_i \in \sigma = \{p_0, \dots, p_d\}$

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

A topological invariant – Euler Characteristics

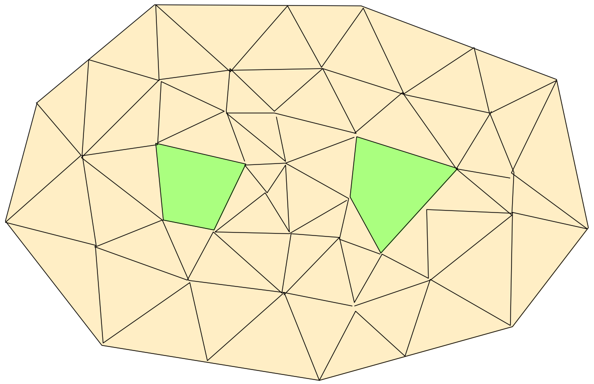
- Given a d-dim simplicial complex K with n_i number of i-simplices
- the *Euler characteristics* of K is defined as:
  - $\chi(K) := \sum_{i=0}(-1)^i n_i$

- Euler characteristics is a topological invariant, meaning that it does not change under homeomorphism.

- Fact:
  - Any two simplicial complexes K and L with homeomorphic underlying spaces $|K| \cong |L|$ have identical Euler characteristics.

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



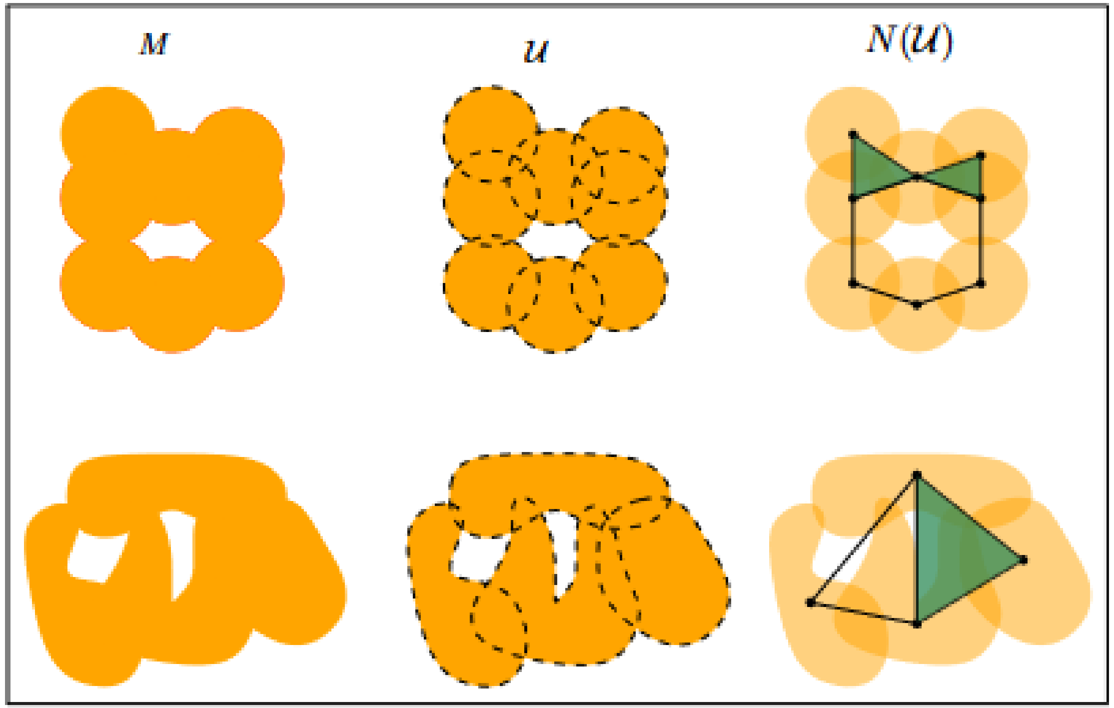
- If K is a triangulation of d-manifold M
  - then the dimension of K is also d
  - for any vertex $v \in Vert(K), St(v) \cong B_d^o \cong R^d$

2.2 Nerves, Cech and Rips complex

Nerves

Definition 2.8 (Nerve). Given a finite collection of sets $\mathfrak{U} = \{U_\alpha\}_{\alpha \in A}$, we define the *nerve* of the set $\mathfrak{U}$ to be the simplicial complex $N(\mathfrak{U})$ whose vertex set is the index set A, and where a subset $\{\alpha_0, \alpha_1, \ldots, \alpha_k\} \subseteq A$ spans a k-simplex in $N(\mathfrak{U})$ if and only if $U_{\alpha_0} \cap U_{\alpha_1} \cap \ldots \cap U_{\alpha_k} \neq \emptyset$.

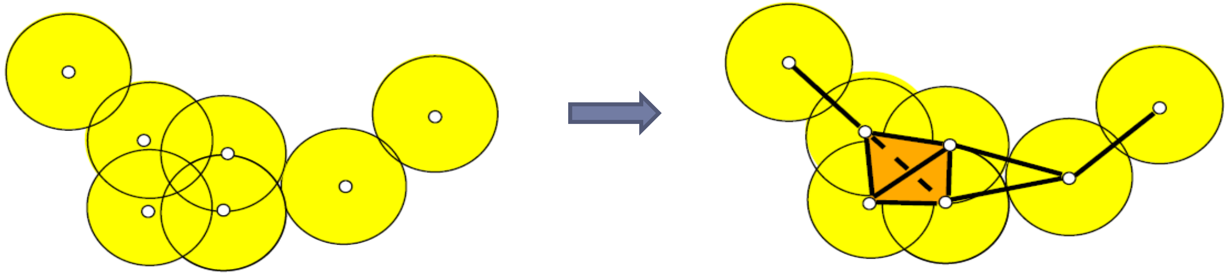
- Hence Čech complex $C^r(P)$
  - is the nerve of $F = \{ B(p,r) \mid p \in P \}$
  - i.e, $C^r(P) = Nrv(F)$



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

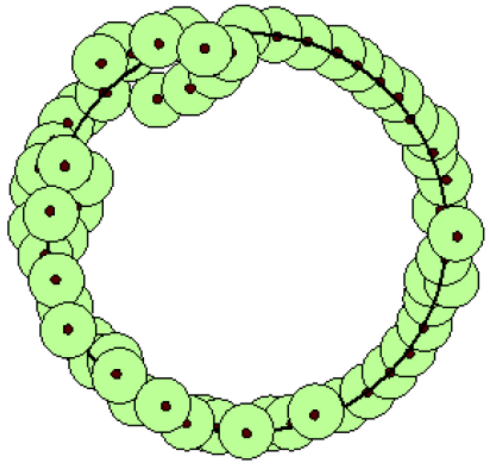


Nerve Lemma

- 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 $|F|$.

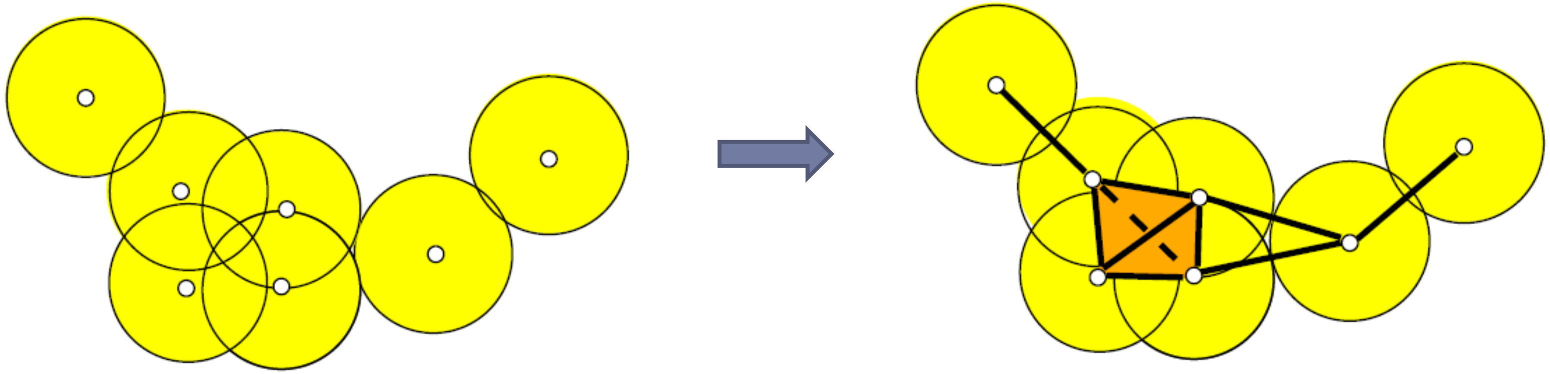
- Corollary:
  - $C^r(P) \simeq \cup_{p \in P} B(p, r)$, i.e, $C^r(P)$ is homotopy equivalent to the union of r-balls around points in P

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



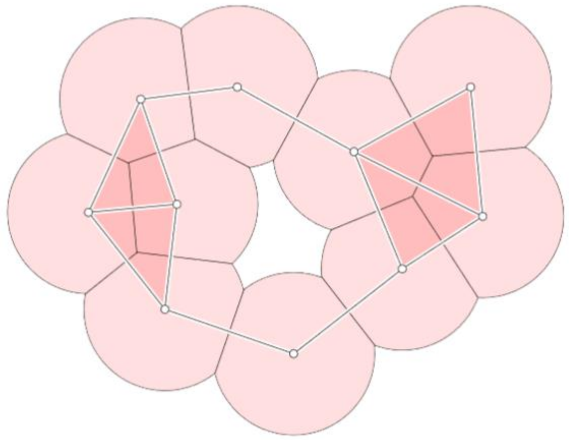
Čech Complex

- 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 $\{B(p_i, r)\}_{i \in [1,n]}$
  - i.e, $\sigma = \{p_{i_0}, \dots, p_{i_s}\} \in C^r(P)$ 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.



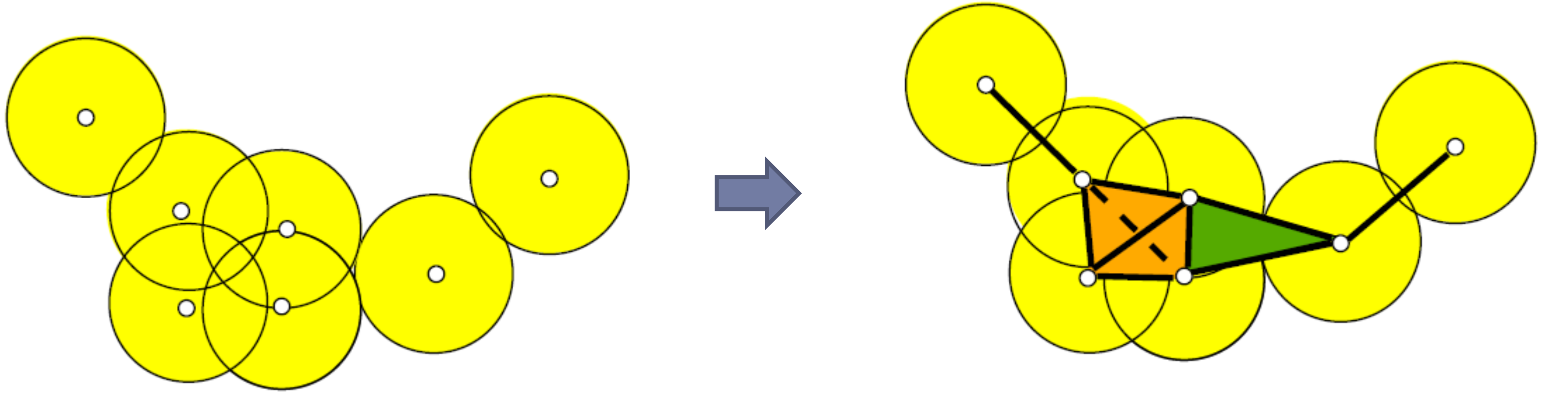
More on Čech

- 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



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* $Rips^r(P)$ is:
  - $\{ (p_{i_0}, p_{i_1}, \dots, p_{i_k}) \mid B_r(p_{i_l}) \cap B_r(p_{i_j}) \neq \emptyset, \forall\, l, j \in [0, k\,]\}$.
- Equivalently, purely metric view:
  - $\mathrm{Rips}^{\mathrm{r}}(\mathrm{P}) = \{ (p_{i_0}, p_{i_1}, \dots, p_{i_k}) \mid d(p_{i_l}, p_{i_j}) \leq 2r, \forall\, l, j \in [0, k\,]\}$.



- 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
  - $C^r(P) \subseteq Rips^r(P) \subseteq C^{2r}(P)$
  - Bounds better in Euclidean space
- Simple to compute
- Able to capture geometry and topology
  - One of the most popular choices for topology inference from PCD in recent years
- However:
  - Huge sizes
  - Computation also costly
  - Much work on sparsified Rips complex

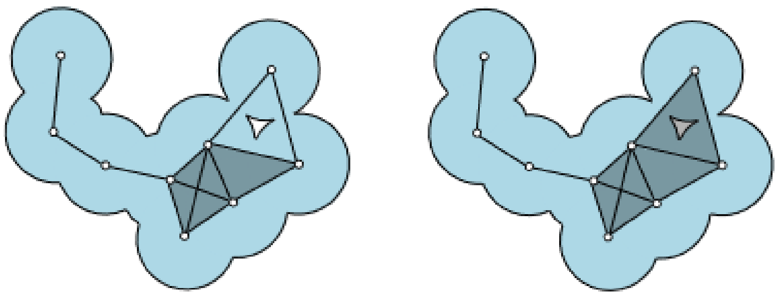
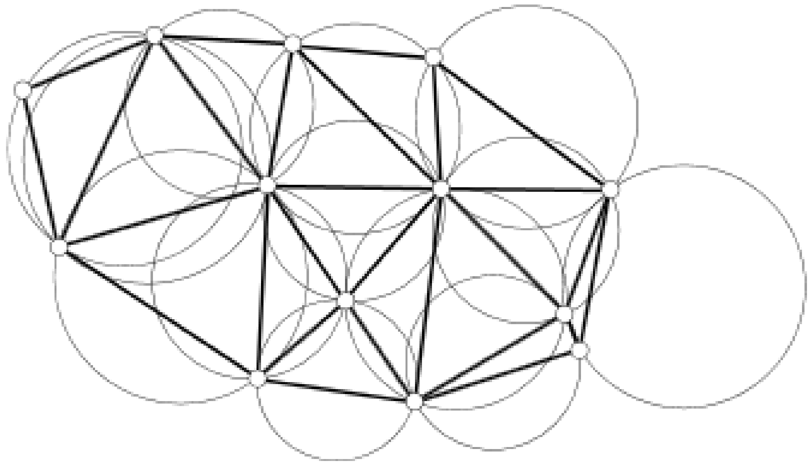


Figure 2.3: (left) Čech complex $\mathbb{C}^r(P)$, (right) Rips complex $\mathbb{VR}^r(P)$.

2.3 Sparse complexes

Delaunay Complex

- Given a set of points $P \subset R^d$
- 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.

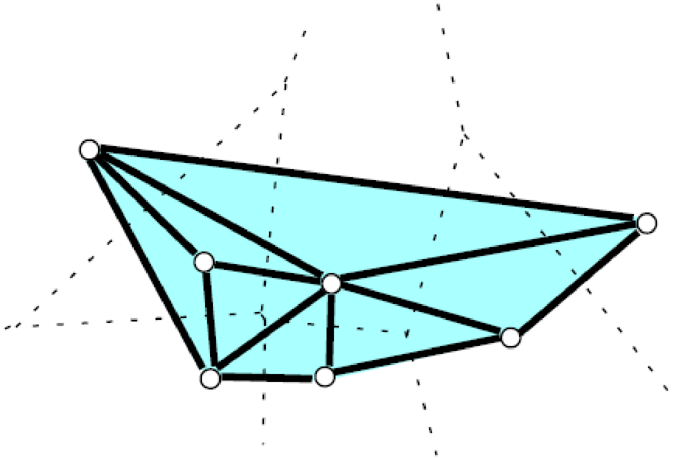


Delaunay Complex

- Many beautiful properties
  - Connection to Voronoi diagram: given $p \in P$
    - Voronoi cell of p is $Vor(p) \coloneqq \{x \in R^d \mid d(x,p) = d(x,P)\}$
  - If points from R^d are in generic positions, then a geometric simplicial complex in R^d

Fact 2.4. *For $P \subset \mathbb{R}^d$, $\mathrm{Del}(P)$ is the nerve of the set of Voronoi cells $\{V_p\}_{p \in P}$ which is a closed cover of $\mathbb{R}^d$.*

- However,
  - Computationally very expensive in high dimensions



Čech and Delaunay

- 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

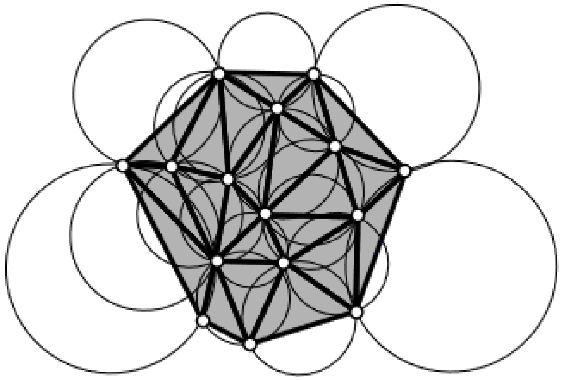


Figure 2.4: Every triangle in a Delaunay complex has an empty open circumdisk.

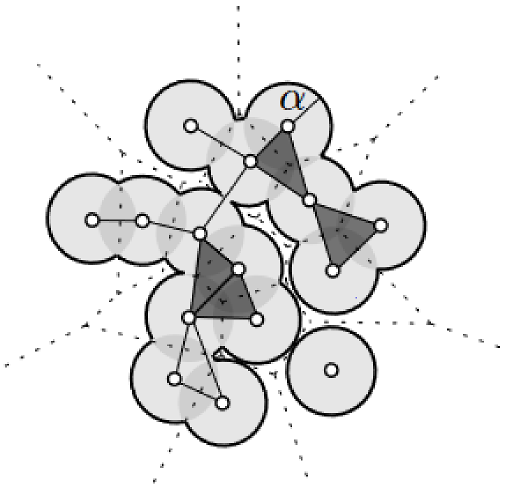
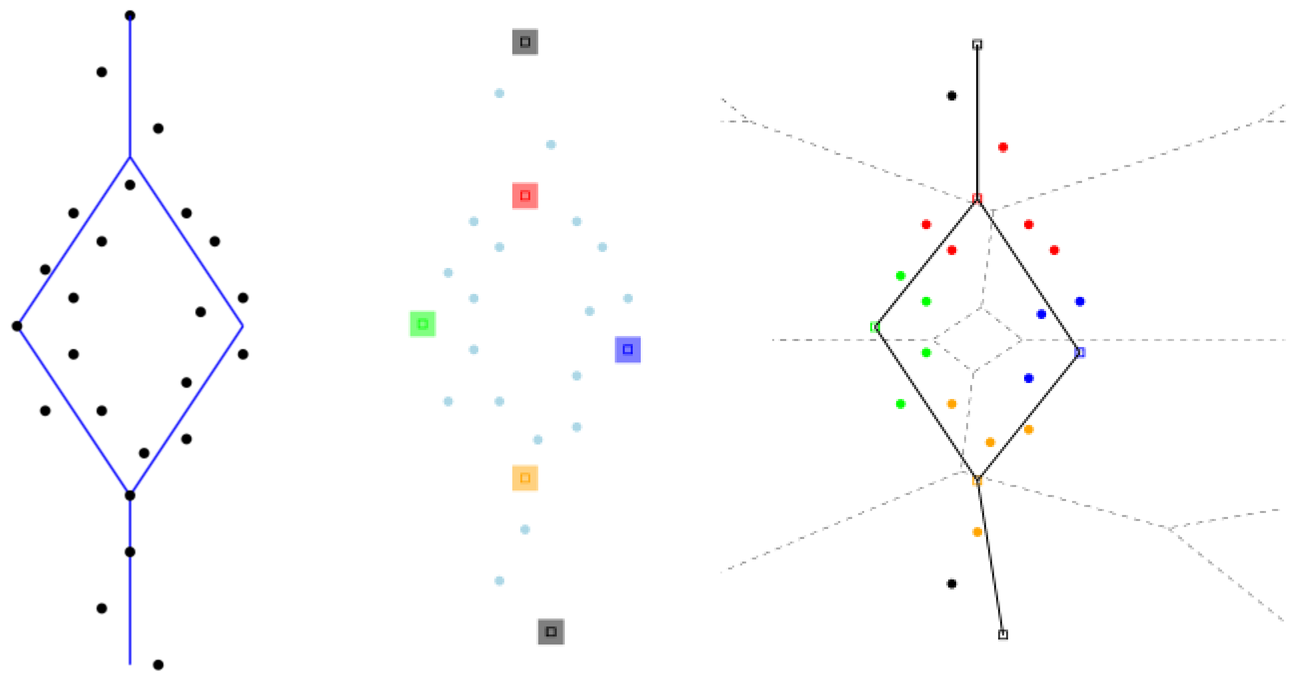


Figure 2.5: Alpha complex of the point set in Figure 2.4 for an α indicated in the figure. The Voronoi diagram of the point set is shown with dotted edges. The triangles and edges in the complex are shown with solid edges which are subset of the Delaunay complex.

Witness complex Intuition

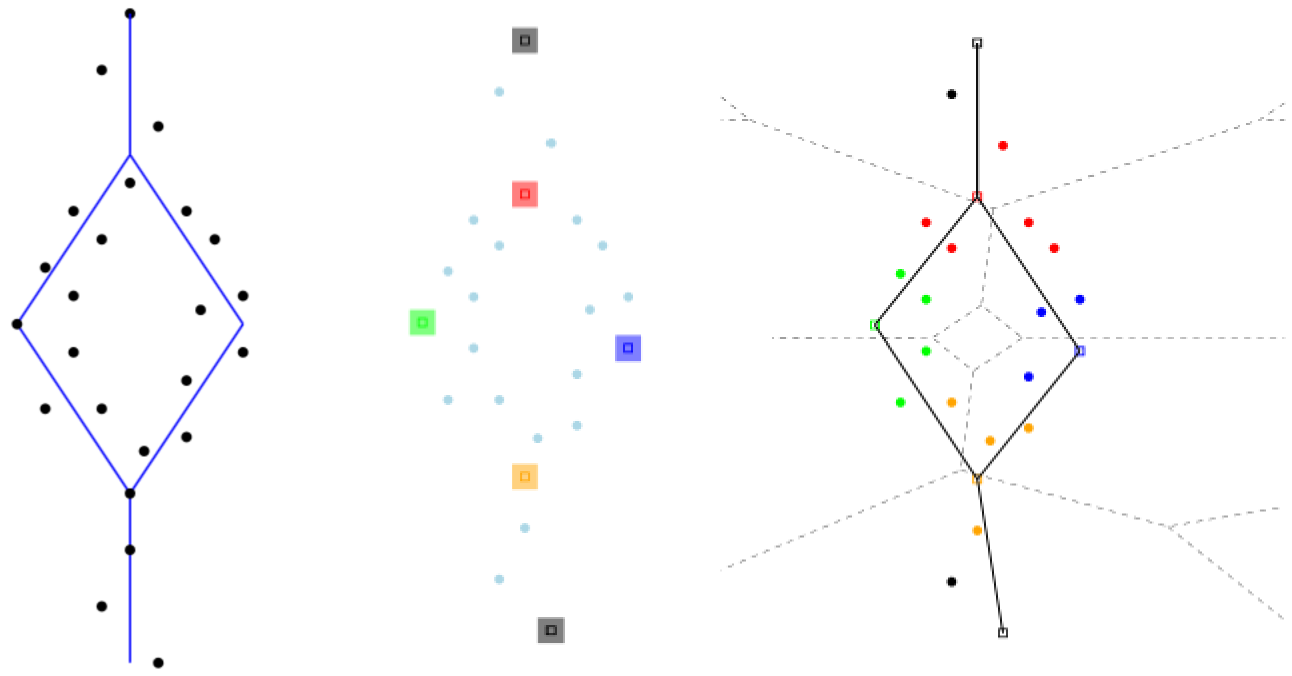
- L: landmarks from P, a way to subsample.



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

Witness complex

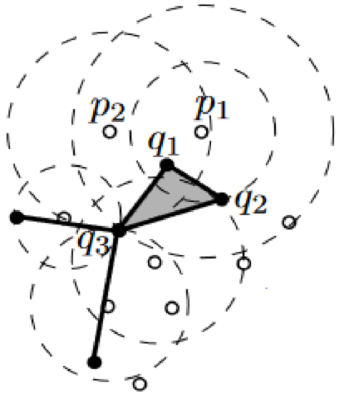
- Using landmarks, but leveraging full points to build complex
- L: landmarks from P, a way to subsample.



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

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 $q \in Q \setminus \{q_0, \dots, q_k\}$.



$q_1 q_2 q_3$ is **weakly witnessed** by p_1
$q_1 q_3$ is **weakly witnessed** by p_2

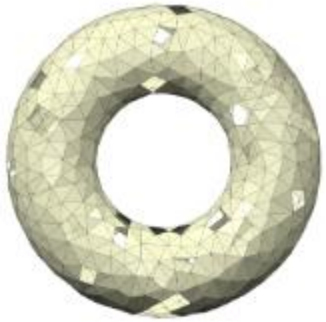
- Given a set of points $P = \{ p_1, p_2, \dots, p_n\} \subset R^d$ and a subset $Q \subseteq P$
- The *witness complex* $W(Q, P)$ is the collection of simplices with vertices from Q whose all subsimplices are weakly witnessed by a point in P.
  - *[de Silva and Carlsson, 2004] [de Silva 2003]*
  - Can be defined for a general metric space
  - P does not have to be a finite subset of points

Witness Complexes

- Greatly reduce size of complex
  - Similar to Delaunay triangulation, remove redundancy

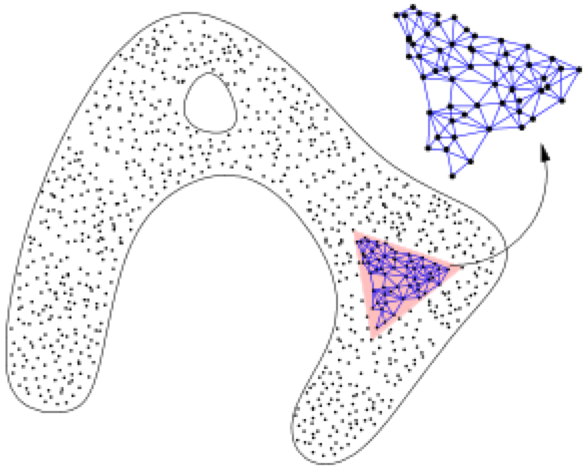
- Relation to Delaunay complex
  - $W(Q,P) \subseteq Del\ Q$ if $Q \subseteq P \subset R^d$
  - $W(Q,R^d) = Del\ Q$

- However,
  - Does not capture full topology easily for high-dimensional manifolds
  - Also expensive to compute

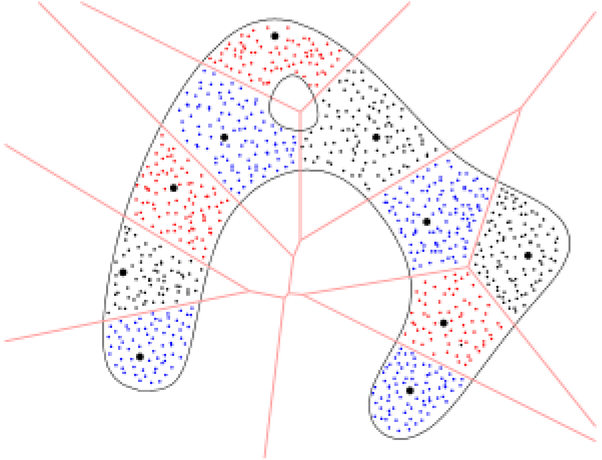


Graph Induced Complex

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



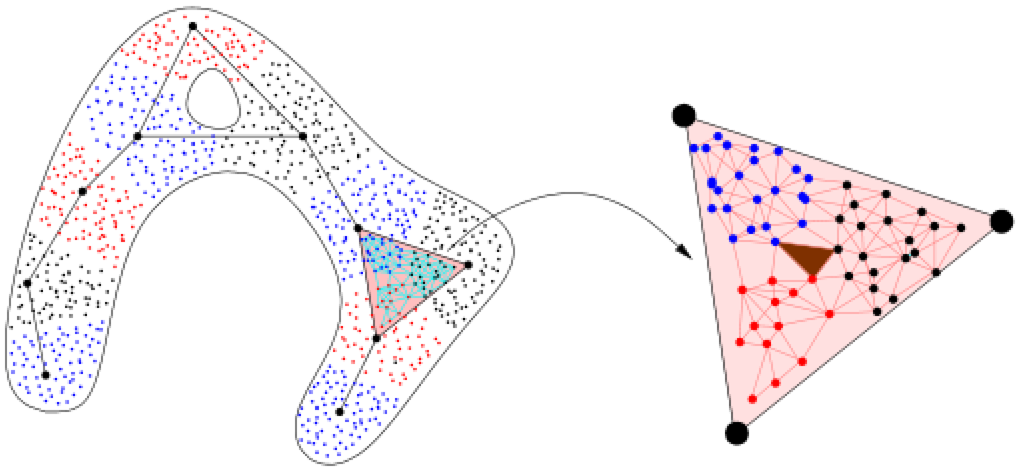
- $Q \subset P$: a subset
- $\pi(p)$: the closest point of $p \in P$ in Q



Dey, Tamal Krishna, Fengtao Fan, and Yusu Wang. "Graph induced complex on point data." *Proceedings of the twenty-ninth annual symposium on Computational geometry*. 2013.

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 $i \in [0,k]$.



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

Dey, Tamal Krishna, Fengtao Fan, and Yusu Wang. "Graph induced complex on point data." *Proceedings of the twenty-ninth annual symposium on Computational geometry*. 2013.

An example pipeline for high-D PCDs (point cloud data)

- Given a PCD $P \subset R^d$
  - First, use a small radius $\delta > 0$, construct the 1-skeleton of $Rips^{\delta}(P)$, denoted by $G_{\delta}(P)$; let d_G be the shortest path distance metric on this graph.
  - Compute a sparse subset (subsample) $Q \subset P$
    - e.g, via the furthest point sampling, which guarantees to be a good subsample (more precisely later in class when we talk about analysis of PCDs)
  - Compute the graph induced complex (GIC) $\mathcal{G}(P, Q, d_G)$

Dey, Tamal Krishna, Fengtao Fan, and Yusu Wang. "Graph induced complex on point data." *Proceedings of the twenty-ninth annual symposium on Computational geometry*. 2013.

Constructing a GIC

Definition 2.17. A subset $Q \subseteq P$ is called a δ-*sample* of a metric space (P, d), if the following condition holds:

- $\forall p \in P$, there exists a $q \in Q$, so that $\mathsf{d}(p, q) \leq \delta$.

Q is called δ-*sparse* if the following condition holds:

- $\forall (q, r) \in Q \times Q$ with $q \neq r$, $\mathsf{d}(q, r) \geq \delta$.

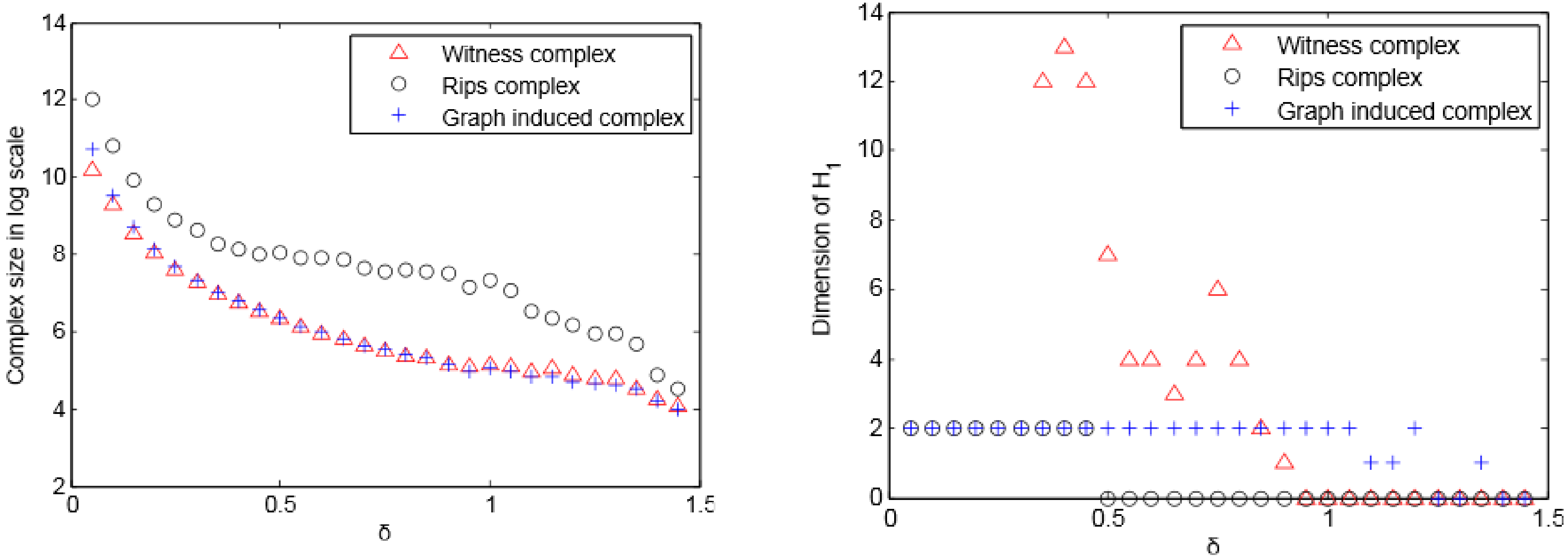
The first condition ensures Q to be a good sample of P with respect to the parameter δ and the second condition enforces that the points in Q cannot be too close relative to the distance δ.

1. Cover the manifold
2. Not too dense

Dey, Tamal Krishna, Fengtao Fan, and Yusu Wang. "Graph induced complex on point data." *Proceedings of the twenty-ninth annual symposium on Computational geometry*. 2013.

Graph Induced Complex

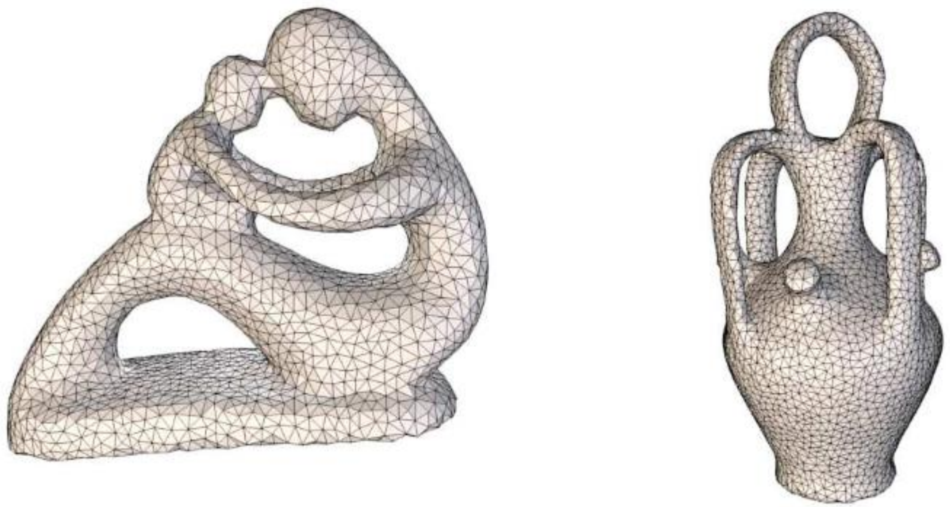
- Small size, but with homology inference guarantees
- In particular:
  - H_1 inference from a lean sample (40,000 sample points from a Klein bottle in R^4)



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

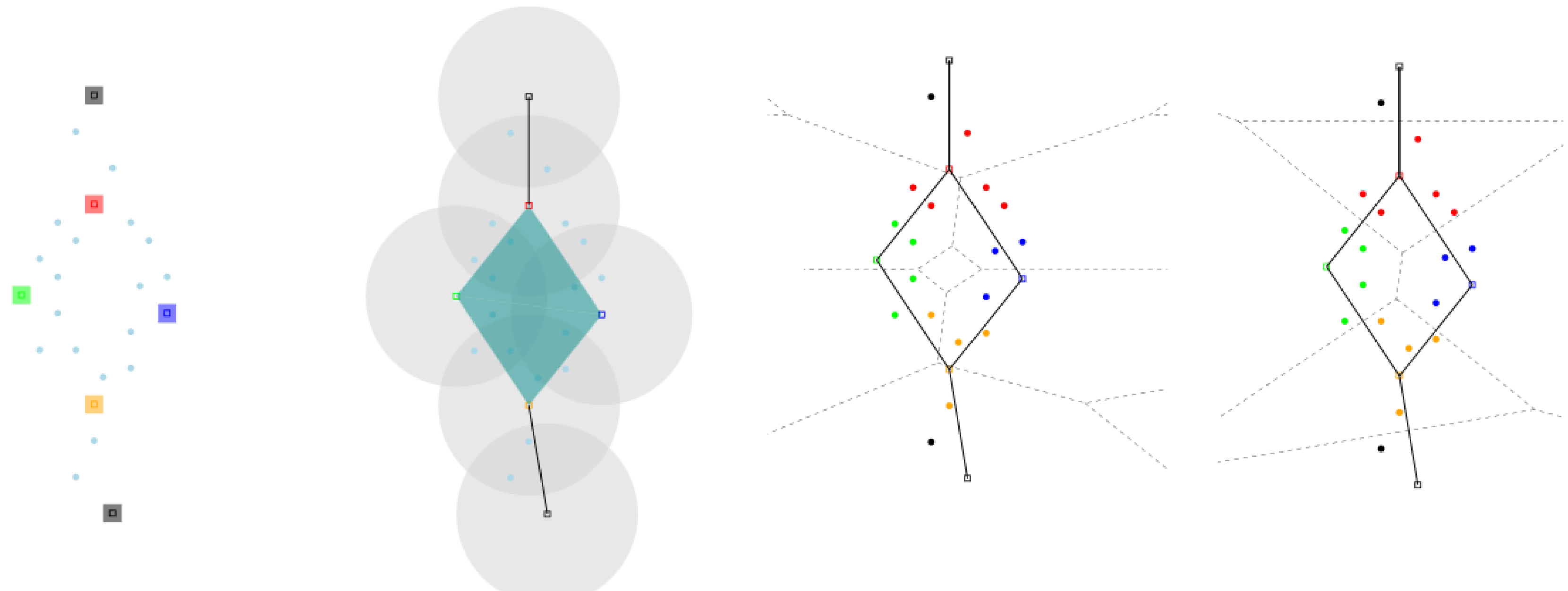


  - Topological inference for compact sets in R^d using persistence

Dey, Tamal Krishna, Fengtao Fan, and Yusu Wang. "Graph induced complex on point data." *Proceedings of the twenty-ninth annual symposium on Computational geometry*. 2013.

Comparisons

- Given a set of dense points P and a sparse subsample $Q \subset P$
  - Rips complex: completely ignores information in P
  - Witness complex: uses info. in P, but hard to compute and also weaker topo guarantee
  - GIC: uses info. in P, easy to compute, and with topo inference guarantees



$Q \subseteq P$ $\quad$ $Rips^r(Q)$ $\quad$ $W(Q,P)$ $\quad$ GIC $\mathcal{G}^r(Q,P)$