Approximating Loops in a Shortest Homology Basis from Point Data

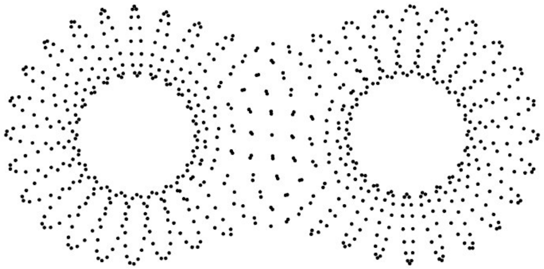
Tamal K. Dey

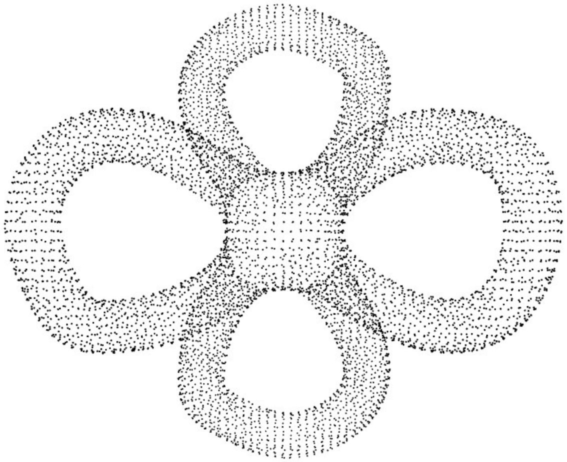
Department of Computer Science & Engineering

The Ohio State University

Joint work with Jian Sun and Yusu Wang

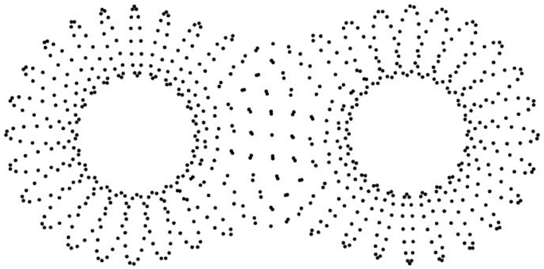
Our Goal



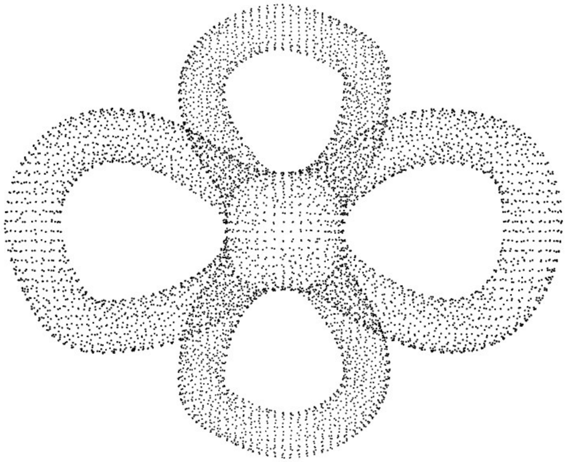


Point cloud

Our Goal



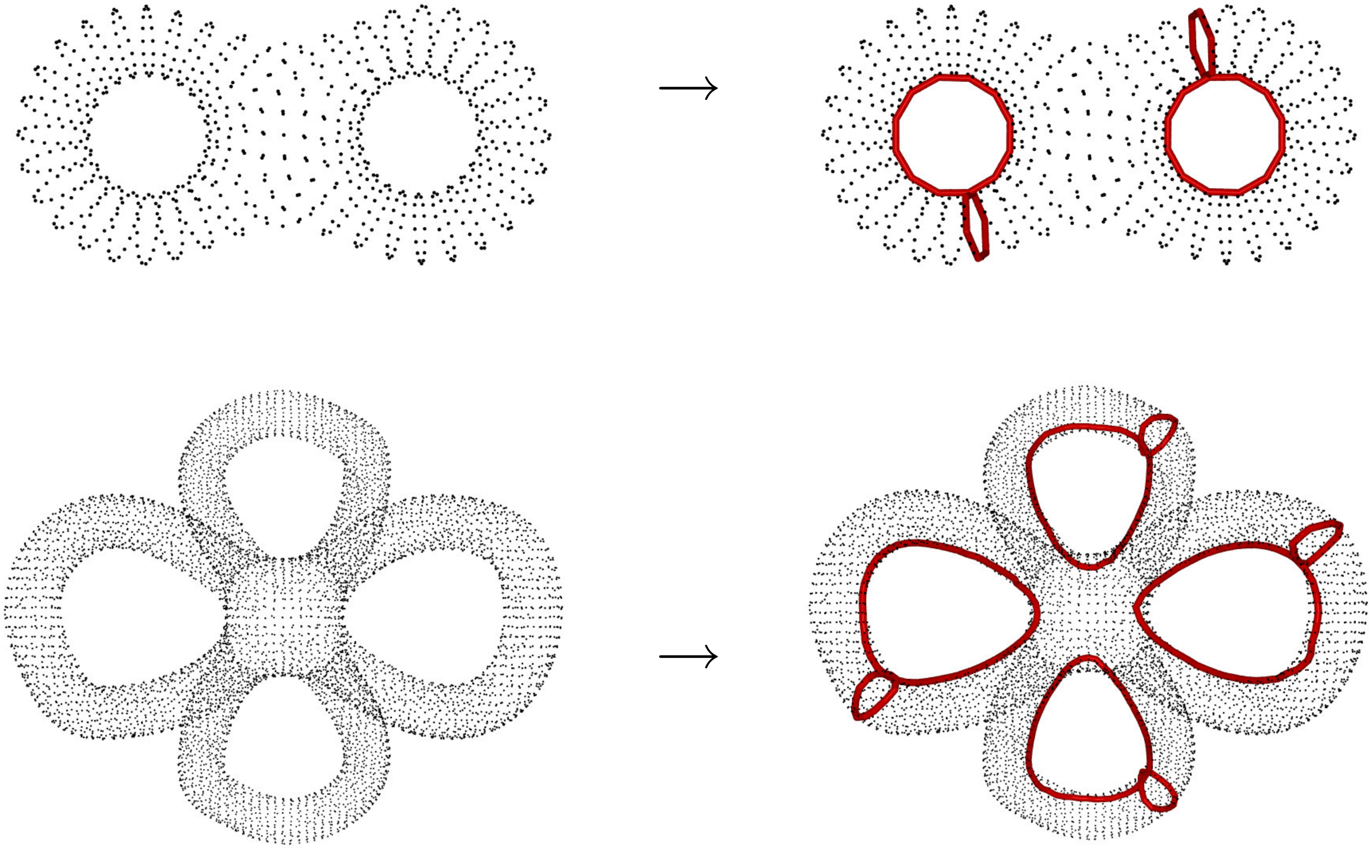
→



→

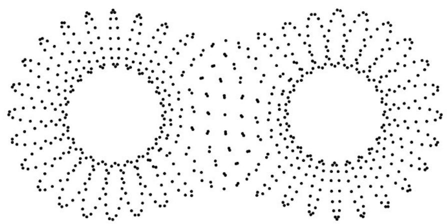
Point cloud

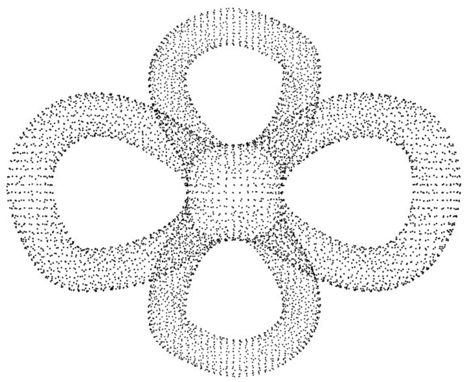
Our Goal



Point cloud Loops

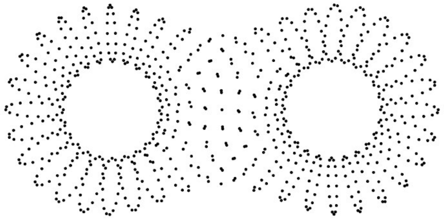
Our Method



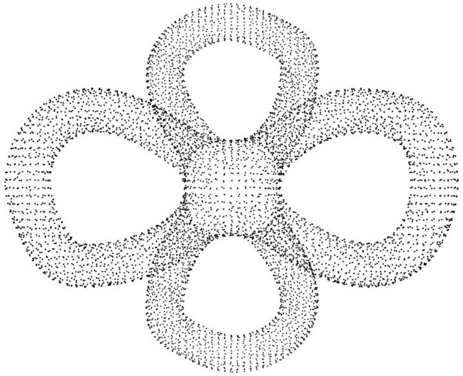


Point cloud

Our Method



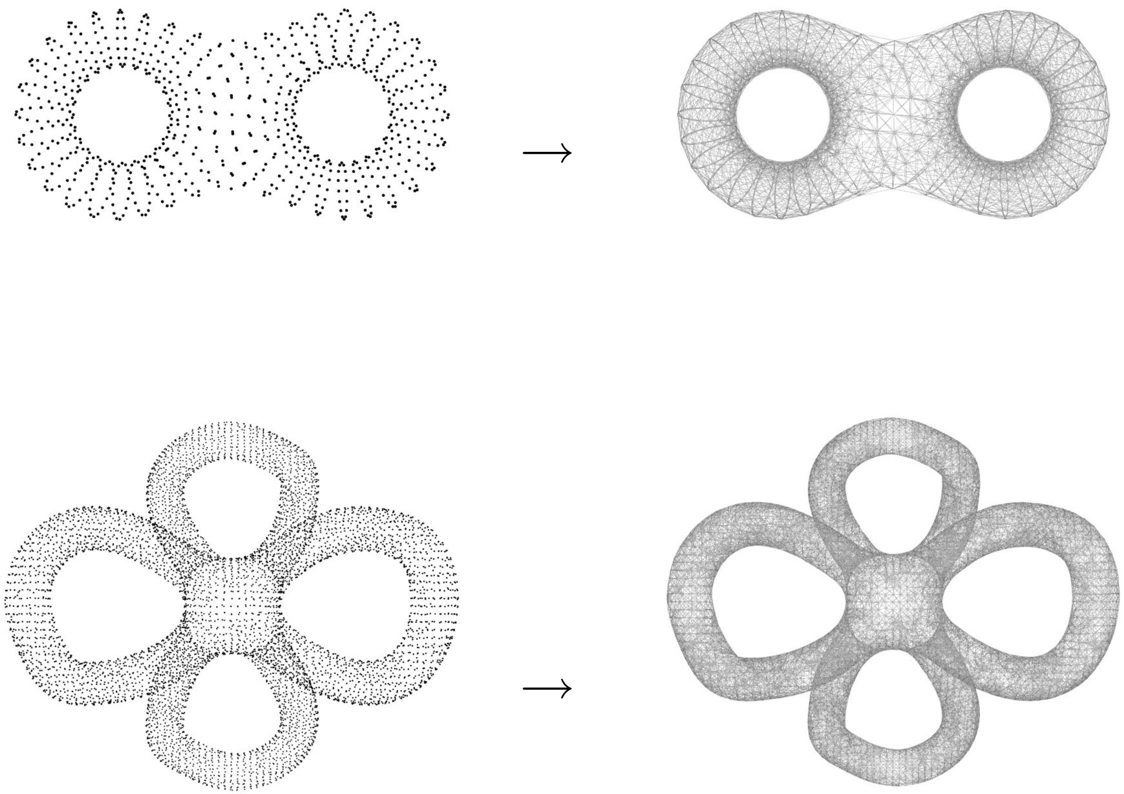
$\longrightarrow$



$\longrightarrow$

Point cloud

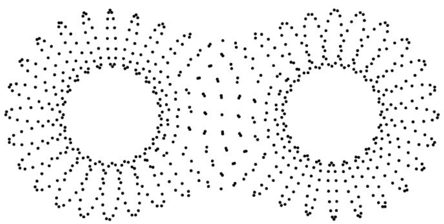
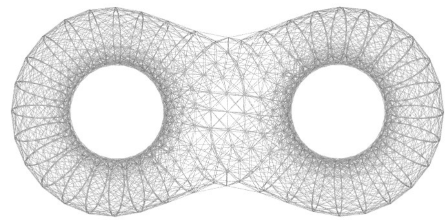
Our Method

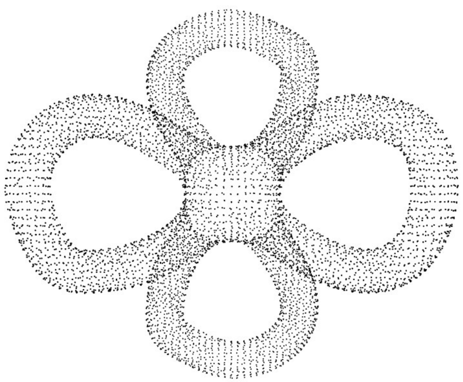
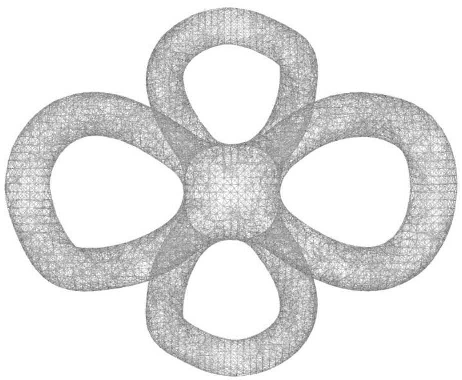


Point cloud

Rips complex

Our Method

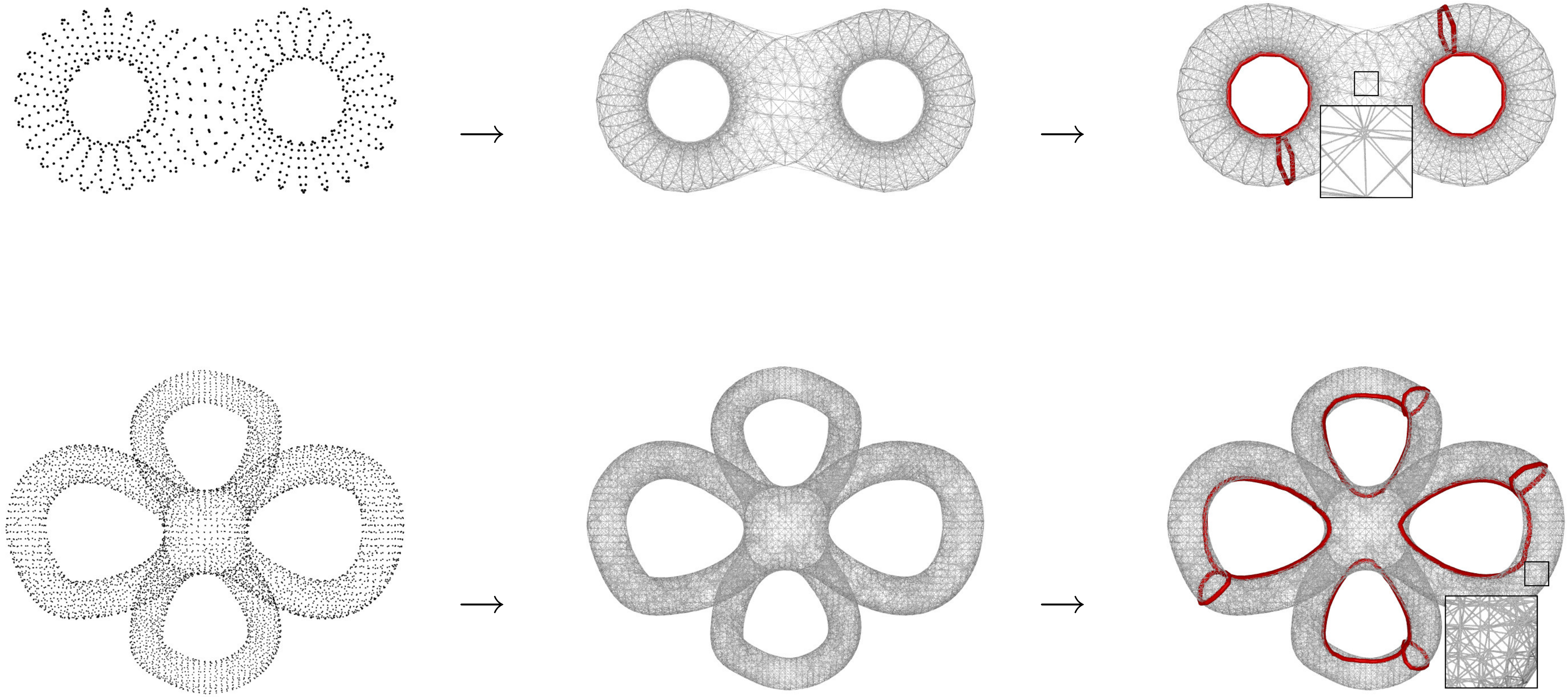
 →  →

 →  →

Point cloud

Rips complex

Our Method



Point cloud Rips complex Loops

Contributions

- Inference of topology and geometry of a hidden manifold from its point data is a fundamental problem.

Contributions

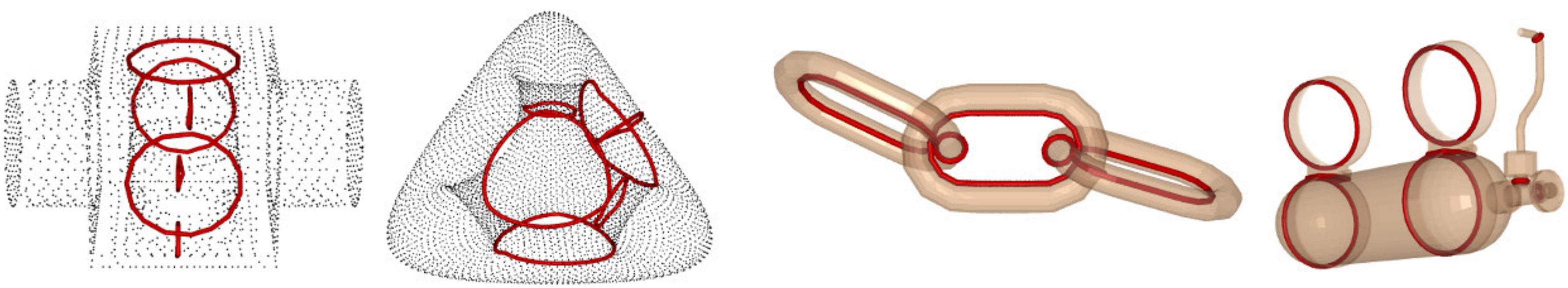
- Inference of topology and geometry of a hidden manifold from its point data is a fundamental problem.
- An algorithm to compute a set of loops from point data that approximate a **shortest** basis of the homology group $\mathsf{H}_1(M)$ of the sampled manifold M.

Contributions

- Inference of topology and geometry of a hidden manifold from its point data is a fundamental problem.
- An algorithm to compute a set of loops from point data that approximate a **shortest** basis of the homology group $\mathsf{H}_1(M)$ of the sampled manifold M.
- A polynomial time algorithm for computing a shortest basis of $\mathsf{H}_1(\mathcal{K})$ for any finite simplicial complex $\mathcal{K}$ embedded in an Euclidean space.

Contributions

- Inference of topology and geometry of a hidden manifold from its point data is a fundamental problem.
- An algorithm to compute a set of loops from point data that approximate a **shortest** basis of the homology group $\mathsf{H}_1(M)$ of the sampled manifold M.
- A polynomial time algorithm for computing a shortest basis of $\mathsf{H}_1(\mathcal{K})$ for any finite simplicial complex $\mathcal{K}$ embedded in an Euclidean space.



Background

- Algorithms for computing homology groups from the point data [CO08].

Background

- Algorithms for computing homology groups from the point data [CO08].
- Cycles from persistence algorithms[ELZ02] lack geometry.

Background

- Algorithms for computing homology groups from the point data [CO08].
- Cycles from persistence algorithms[ELZ02] lack geometry.
- Reconstruction of the sampled space (can be costly).

Background

- Algorithms for computing homology groups from the point data [CO08].
- Cycles from persistence algorithms[ELZ02] lack geometry.
- Reconstruction of the sampled space (can be costly).
- Rips, Čech or witness complexes are less constrained. We use Rips complex.

Background

- Algorithms for computing homology groups from the point data [CO08].
- Cycles from persistence algorithms[ELZ02] lack geometry.
- Reconstruction of the sampled space (can be costly).
- Rips, Čech or witness complexes are less constrained. We use Rips complex.
- 2-manifold triangulations [EW05].

Background

- Algorithms for computing homology groups from the point data [CO08].
- Cycles from persistence algorithms[ELZ02] lack geometry.
- Reconstruction of the sampled space (can be costly).
- Rips, Čech or witness complexes are less constrained. We use Rips complex.
- 2-manifold triangulations [EW05].
- NP-hard for higher dimenisonal homology groups [CF10].

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$: j**-dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$:j**-dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.
- The elements of $\mathsf{H}_1(\mathbb{T})$ are equivalent classes $[g]$ of 1-dimensional cycles g, also called **loops**.

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$: j**-dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.
- The elements of $\mathsf{H}_1(\mathbb{T})$ are equivalent classes $[g]$ of 1-dimensional cycles g, also called **loops**.
- A minimal set $\{[g_1], ..., [g_k]\}$ generating $\mathsf{H}_1(\mathbb{T})$ is called its **basis**.

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$:j**-dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.
- The elements of $\mathsf{H}_1(\mathbb{T})$ are equivalent classes $[g]$ of 1-dimensional cycles g, also called **loops**.
- A minimal set $\{[g_1], ..., [g_k]\}$ generating $\mathsf{H}_1(\mathbb{T})$ is called its **basis**.
- Here $k = \mathrm{rank}(\mathsf{H}_1(\mathbb{T}))$.

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$:j-**dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.
- The elements of $\mathsf{H}_1(\mathbb{T})$ are equivalent classes $[g]$ of 1-dimensional cycles g, also called **loops**.
- A minimal set $\{[g_1], ..., [g_k]\}$ generating $\mathsf{H}_1(\mathbb{T})$ is called its **basis**.
- Here $k = \mathrm{rank}(\mathsf{H}_1(\mathbb{T}))$.
- We associate a weight $w(g) \geq 0$ with each loop g in $\mathbb{T}$.

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$: j-**dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.
- The elements of $\mathsf{H}_1(\mathbb{T})$ are equivalent classes $[g]$ of 1-dimensional cycles g, also called **loops**.
- A minimal set $\{[g_1], ..., [g_k]\}$ generating $\mathsf{H}_1(\mathbb{T})$ is called its **basis**.
- Here $k = \mathrm{rank}(\mathsf{H}_1(\mathbb{T}))$.
- We associate a weight $w(g) \geq 0$ with each loop g in $\mathbb{T}$.
- The length of a set of loops $G = \{g_1, ..., g_k\}$ is given by $\mathrm{Len}(G) = \Sigma_{i=1}^{k} w(g_i)$.

Homology groups & generators

- $\mathsf{H}_j(\mathbb{T})$:j-**dimensional homology group** of $\mathbb{T}$ under $\mathbb{Z}_2$.
- The elements of $\mathsf{H}_1(\mathbb{T})$ are equivalent classes $[g]$ of 1-dimensional cycles g, also called **loops**.
- A minimal set $\{[g_1], ..., [g_k]\}$ generating $\mathsf{H}_1(\mathbb{T})$ is called its **basis**.
- Here $k = \mathrm{rank}(\mathsf{H}_1(\mathbb{T}))$.
- We associate a weight $w(g) \geq 0$ with each loop g in $\mathbb{T}$.
- The length of a set of loops $G = \{g_1, ..., g_k\}$ is given by $\mathrm{Len}(G) = \Sigma_{i=1}^{k} w(g_i)$.
- A **shortest basis** of $\mathsf{H}_1(\mathbb{T})$ is a set of k loops with minimal length generating $\mathsf{H}_1(\mathbb{T})$.

Theorem 1

- Let $\mathcal{K}$ be a finite simplicial complex with non-negative weights on edges.
- A shortest basis for $\mathrm{H}_1(\mathcal{K})$ can be computed in $O(n^4)$ time where $n = |\mathcal{K}|$.

Shortest basis

- $\mathsf{H}_1(\mathcal{K})$ is a vector space and supports matroid theory.

Shortest basis

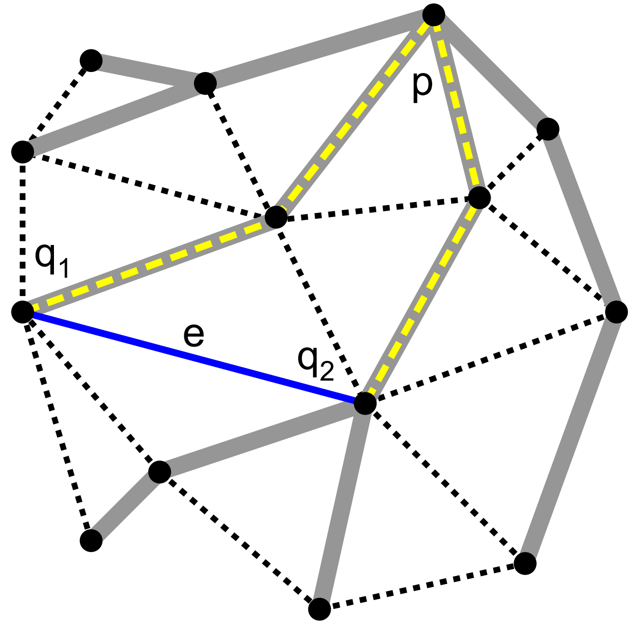
- $\mathrm{H}_1(\mathcal{K})$ is a vector space and supports matroid theory.
- If a set of loops $\mathcal{L}$ in $\mathcal{K}$ contains a shortest basis, then the *greedy* set G chosen from $\mathcal{L}$ is a shortest basis by matroid theory.

Shortest basis

- $\mathsf{H}_1(\mathcal{K})$ is a vector space and supports matroid theory.
- If a set of loops $\mathcal{L}$ in $\mathcal{K}$ contains a shortest basis, then the *greedy* set G chosen from $\mathcal{L}$ is a shortest basis by matroid theory.
- The **greedy set** G is an *ordered* set of loops $\{g_1, ..., g_k\}$ satisfying the following conditions:
  g_1 is the shortest loop in $\mathcal{L}$ nontrivial in $\mathsf{H}_1(\mathcal{K})$;
  g_{i+1} is the shortest loop in $\mathcal{L}$ independent of $g_1, ..., g_i$.

Canonical loop

- Let T be a shortest path tree in $\mathcal{K}$ rooted at p.
- For $q_1, q_2 \in P$, $\mathrm{sp}_T(q_1, q_2)$ denotes the unique path from q_1 to q_2 through p in T.
- Let E_T be the set of edges in T.
- The **canonical loop** for a non-tree edge e is defined as



$$T(e) = \mathrm{sp}_T(p, q_1) \circ e \circ \mathrm{sp}_T(q_2, p).$$

Candidate loops

- A simple cycle l is **tight** if it contains a shortest path between every pair of points in l.

Candidate loops

- A simple cycle l is **tight** if it contains a shortest path between every pair of points in l.
- **Proposition** Loops in a shortest basis of $\mathsf{H}_1(\mathcal{K})$ are tight.

Candidate loops

- A simple cycle l is **tight** if it contains a shortest path between every pair of points in l.
- **Proposition** Loops in a shortest basis of $\mathsf{H}_1(\mathcal{K})$ are tight.
- Let C_p be the set of all canonical loops with respect to p:

$$C_p = \{T(e) : e \in E \backslash E_T\}.$$

Candidate loops

- A simple cycle l is **tight** if it contains a shortest path between every pair of points in l.
- **Proposition** Loops in a shortest basis of $\mathsf{H}_1(\mathcal{K})$ are tight.
- Let C_p be the set of all canonical loops with respect to p:

$$C_p = \{T(e) : e \in E \backslash E_T\}.$$

- **Proposition** $\cup_{p \in P} C_p$ contains all tight loops and hence any shortest basis.

Candidate loops

- A simple cycle l is **tight** if it contains a shortest path between every pair of points in l.
- **Proposition** Loops in a shortest basis of $\mathsf{H}_1(\mathcal{K})$ are tight.
- Let C_p be the set of all canonical loops with respect to p:

$$C_p = \{T(e) : e \in E \backslash E_T\}.$$

- **Proposition** $\cup_{p \in P} C_p$ contains all tight loops and hence any shortest basis.
- Let G_p be the greedy set chosen from C_p.

Candidate loops

- A simple cycle l is **tight** if it contains a shortest path between every pair of points in l.
- **Proposition** Loops in a shortest basis of $\mathsf{H}_1(\mathcal{K})$ are tight.
- Let C_p be the set of all canonical loops with respect to p:

$$C_p = \{T(e) : e \in E \backslash E_T\}.$$

- **Proposition** $\cup_{p \in P} C_p$ contains all tight loops and hence any shortest basis.
- Let G_p be the greedy set chosen from C_p.
- **Proposition** The greedy set chosen from $\cup_{p \in P} G_p$ is a shortest basis of $\mathsf{H}_1(\mathcal{K})$.

Computing G_p

Proposition CANONGEN$(p, \mathcal{K})$ outputs G_p.

CANONGEN($p, \mathcal{K}$)

1: Construct a shortest path tree T in $\mathcal{K}$ with p as the root.

Computing G_p

Proposition CanonGen$(p, \mathcal{K})$ outputs G_p.

CanonGen$(p, \mathcal{K})$

1: Construct a shortest path tree T in $\mathcal{K}$ with p as the root.

2: For each non-tree edge $e = (q_1, q_2) \in E \backslash E_T$, let $T(e)$ be the canonical loop of e.

Computing G_p

Proposition CanonGen$(p, \mathcal{K})$ outputs G_p.

CanonGen($p, \mathcal{K}$)

1: Construct a shortest path tree T in $\mathcal{K}$ with p as the root.

2: For each non-tree edge $e = (q_1, q_2) \in E \backslash E_T$, let $T(e)$ be the canonical loop of e.

3: Run the persistence algorithm based on the following filtration of $\mathcal{K}$: vertices in $P = \mathrm{Vert}(\mathcal{K})$, tree edges in T, non-tree edges in the canonical order, triangles in $\mathcal{K}$. Return the set of canonical loops associated with $k = \mathrm{rank}(\mathsf{H}_1(\mathcal{K}))$ edges unpaired after the algorithm.

Computing shortest basis

SPGEN($\mathcal{K}$)

```
For each p ∈ P = Vert(K), set G_p := CANONGEN(p, K).
Sort all loops in ∪_p G_p by lengths in the increasing order.
Let g_1, ..., g_{k|P|} be this sorted list. Initialize G := {g_1}.
for i := 2 to k|P|, do
    if |G| = k, then
        Exit the for loop.
    else if g_i is independent of loops in G, then
        Add g_i to G.
    end if
end for
Return G.
```

Computing independence

- How to determine if g is independent of loops in G?

Computing independence

- How to determine if g is independent of loops in G?
- For each loop $g'_1, ..., g'_s$ in G we add triangles to fill it.

Computing independence

- How to determine if g is independent of loops in G?
- For each loop $g'_1, ..., g'_s$ in G we add triangles to fill it.
- We choose a dummy vertex, say v, and add triangles vv_iv_{i+1} for each edge v_iv_{i+1} of the loops to be filled.

Computing independence

- How to determine if g is independent of loops in G?
- For each loop $g'_1, ..., g'_s$ in G we add triangles to fill it.
- We choose a dummy vertex, say v, and add triangles vv_iv_{i+1} for each edge v_iv_{i+1} of the loops to be filled.
- These triangles destroy the generators $g'_1, ..., g'_s$. They destroy g as well if and only if g is dependent on $g'_1, ..., g'_s$ [Chen-Freedman].

Computing independence

- How to determine if g is independent of loops in G?
- For each loop $g'_1, ..., g'_s$ in G we add triangles to fill it.
- We choose a dummy vertex, say v, and add triangles vv_iv_{i+1} for each edge v_iv_{i+1} of the loops to be filled.
- These triangles destroy the generators $g'_1, ..., g'_s$. They destroy g as well if and only if g is dependent on $g'_1, ..., g'_s$ [Chen-Freedman].
- Whether g is rendeded trivial can be determined by augmenting the filtration of $\mathcal{K}$ with the simplices in $\mathcal{K}'\backslash\mathcal{K}$ and continuing the persistence algorithm.

Approximation from point cloud

- Let $P \subset \mathbb{R}^d$ be a point set sampled from a smooth closed manifold $M \subset \mathbb{R}^d$ embedded isometrically.

Approximation from point cloud

- Let $P \subset \mathbb{R}^d$ be a point set sampled from a smooth closed manifold $M \subset \mathbb{R}^d$ embedded isometrically.
- We want to approximate a shortest basis of $\mathsf{H}_1(M)$ from P.

Approximation from point cloud

- Let $P \subset \mathbb{R}^d$ be a point set sampled from a smooth closed manifold $M \subset \mathbb{R}^d$ embedded isometrically.
- We want to approximate a shortest basis of $\mathsf{H}_1(M)$ from P.
- Compute a complex $\mathcal{K}$ from P. Compute a shortest basis of $\mathsf{H}_1(\mathcal{K})$. Argue that if P is *d*ense, a subset of computed loops approximate a shortest basis of $\mathsf{H}_1(M)$ within constant factors.

Complexes

- Let $P \subset \mathbb{R}^d$ be a point set. $B(p, r)$ denotes an open d-ball centered at p with radius r.

Complexes

- Let $P \subset \mathbb{R}^d$ be a point set. $B(p, r)$ denotes an open d-ball centered at p with radius r.
- The **Čech complex** $\mathcal{C}^r(P)$ is a simplicial complex where a simplex $\sigma \in \mathcal{C}^r(P)$ iff $\mathrm{Vert}(\sigma)$, the vertices of σ, are in P and $\cap_{p \in \mathrm{Vert}(\sigma)} B(p, r/2) \neq 0$.

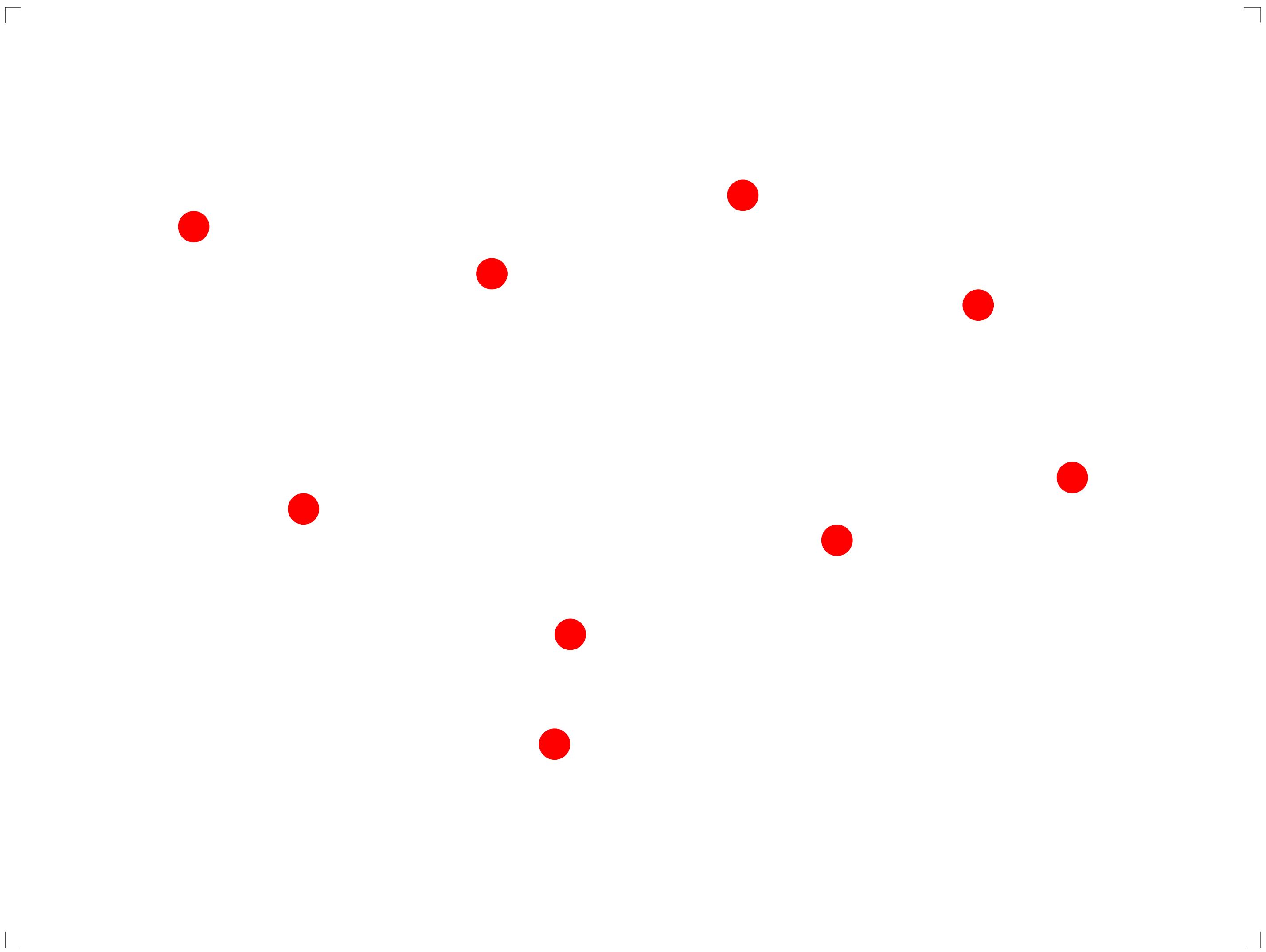
Complexes

- Let $P \subset \mathbb{R}^d$ be a point set. $B(p, r)$ denotes an open d-ball centered at p with radius r.
- The **Čech complex** $\mathcal{C}^r(P)$ is a simplicial complex where a simplex $\sigma \in \mathcal{C}^r(P)$ iff $\mathrm{Vert}(\sigma)$, the vertices of σ, are in P and $\cap_{p \in \mathrm{Vert}(\sigma)} B(p, r/2) \neq 0$.
- The **Rips complex** $\mathcal{R}^r(P)$ is a simplicial complex where a simplex $\sigma \in \mathcal{R}^r(P)$ if and only if $\mathrm{Vert}(\sigma)$ are within pairwise Euclidean distance of r.

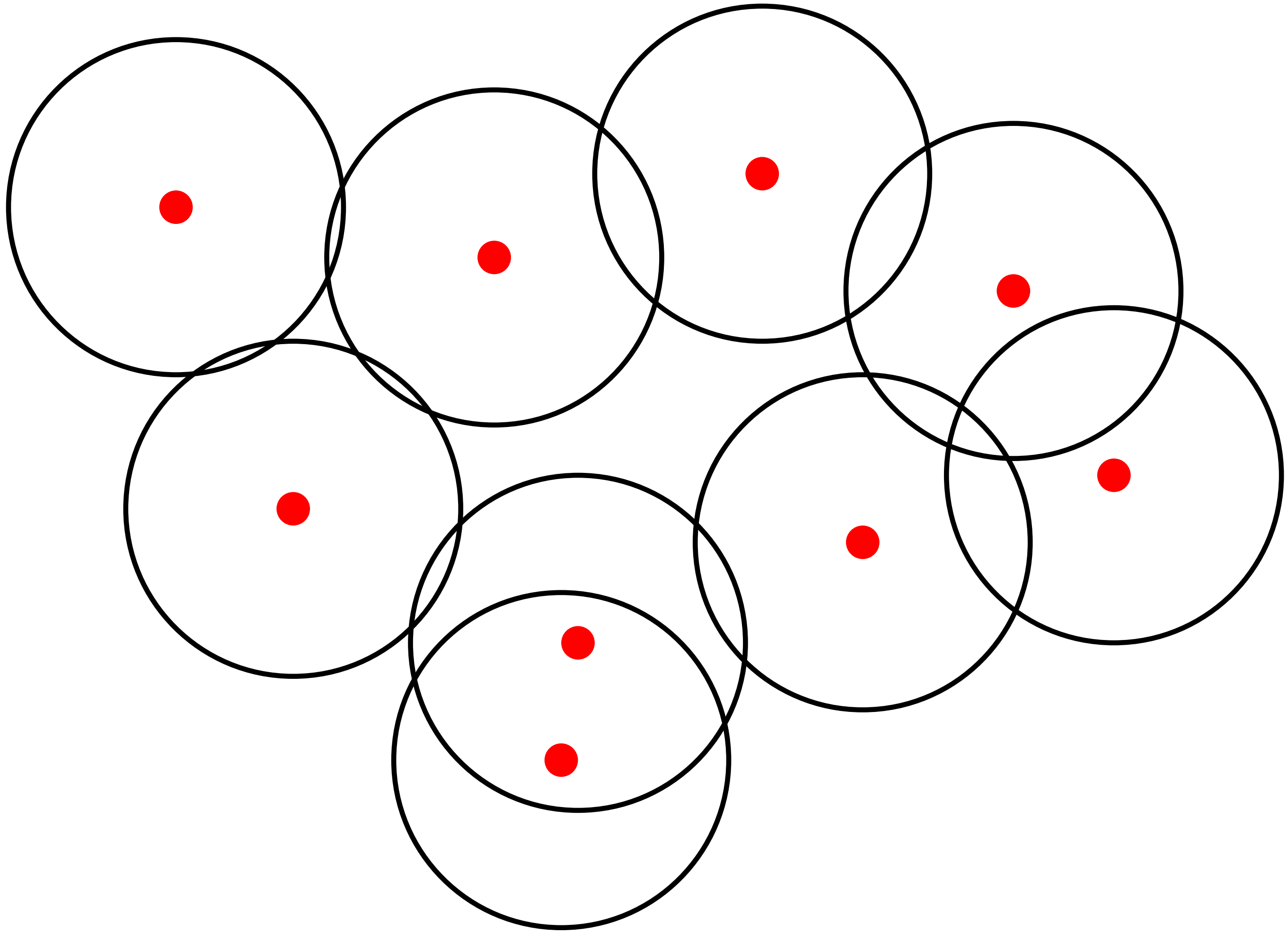
Complexes

- Let $P \subset \mathbb{R}^d$ be a point set. $B(p, r)$ denotes an open d-ball centered at p with radius r.
- The **Čech complex** $\mathcal{C}^r(P)$ is a simplicial complex where a simplex $\sigma \in \mathcal{C}^r(P)$ iff $\mathrm{Vert}(\sigma)$, the vertices of σ, are in P and $\cap_{p \in \mathrm{Vert}(\sigma)} B(p, r/2) \neq 0$.
- The **Rips complex** $\mathcal{R}^r(P)$ is a simplicial complex where a simplex $\sigma \in \mathcal{R}^r(P)$ if and only if $\mathrm{Vert}(\sigma)$ are within pairwise Euclidean distance of r.
- **Proposition** For any finite set $P \subset \mathbb{R}^d$ and any $r \geq 0$, one has $\mathcal{C}^r(P) \subseteq \mathcal{R}^r(P) \subseteq \mathcal{C}^{2r}(P)$.

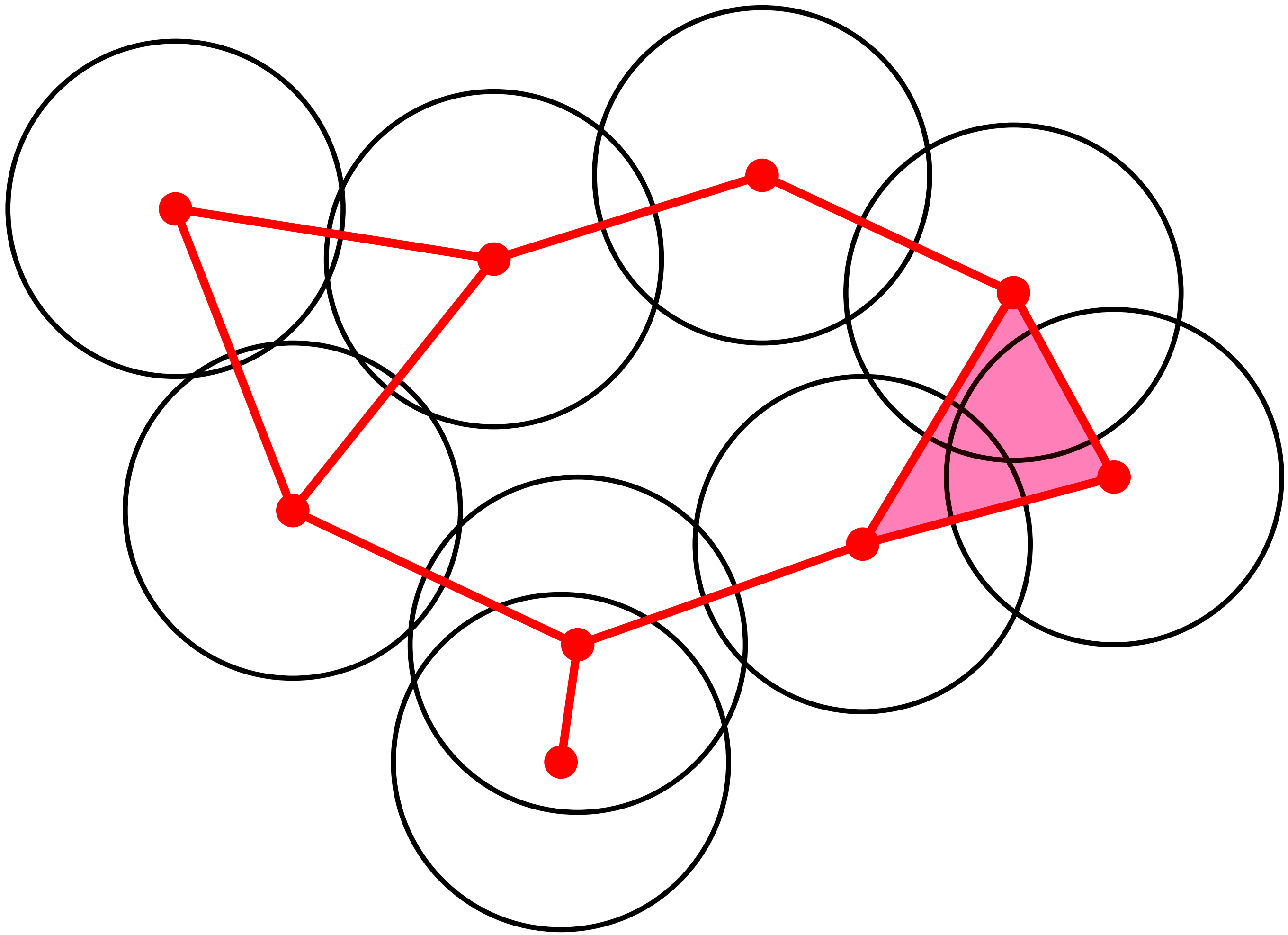
Point set P



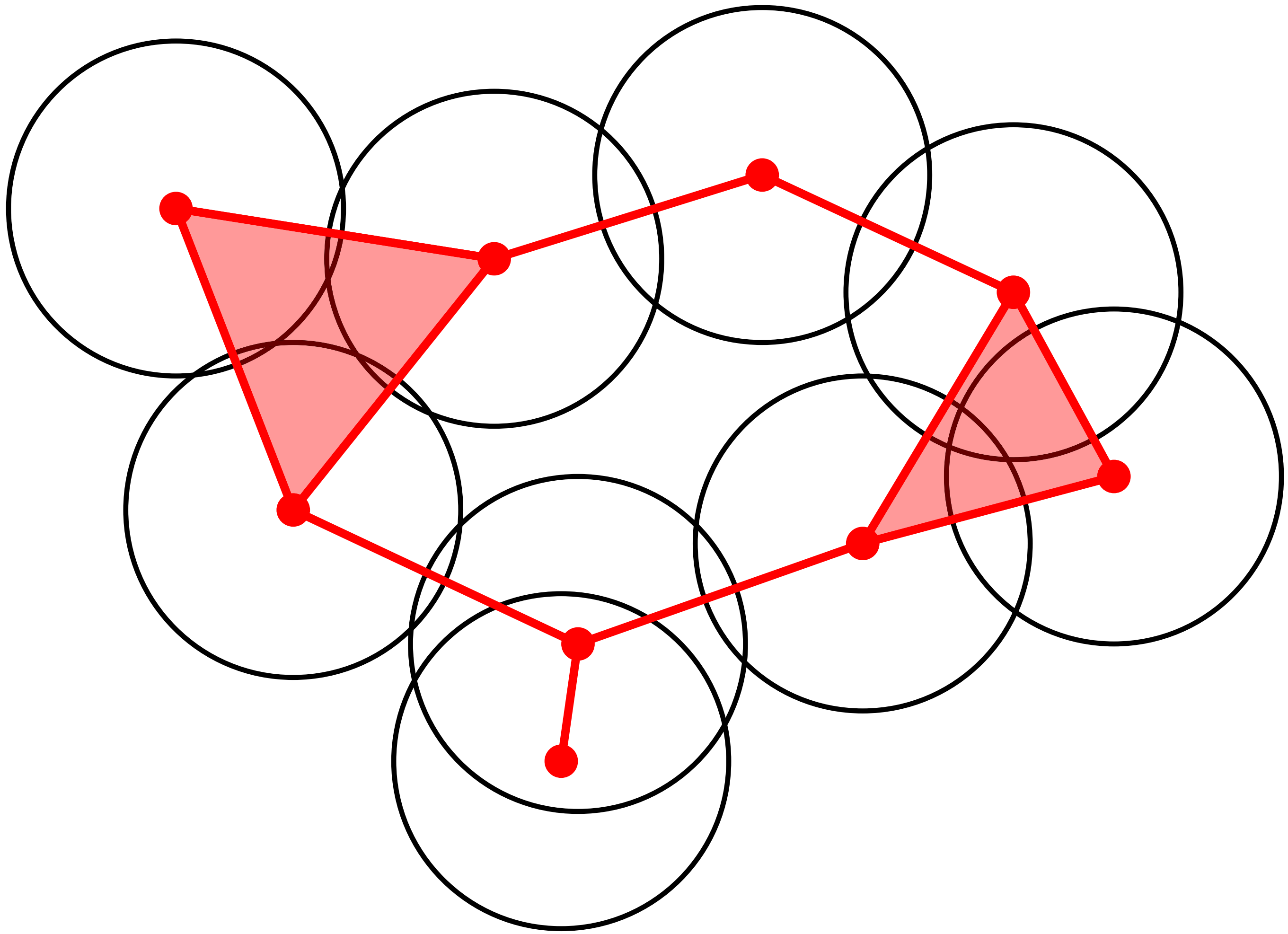
Balls $B(p, r/2)$ for $p \in P$



Čech complex $\mathcal{C}^r(P)$



Rips complex $\mathcal{R}^r(P)$



Manifold constants and sampling

- **Geodesic ball**: $B_M(p, r) = \{q \mid d_M(p, q) < r\}$.

Manifold constants and sampling

- **Geodesic ball**: $B_M(p,r) = \{q \mid d_M(p,q) < r\}$.
- There is a $r_p > 0$ for each $p \in M$ where r_p is the supremum r so that $B_M(p,r)$ is convex (the minimizing geodesics between any two points in $B_M(p,r)$ lie in $B_M(p,r)$).

Manifold constants and sampling

- **Geodesic ball**: $B_M(p,r) = \{q \mid d_M(p,q) < r\}$.
- There is a $r_p > 0$ for each $p \in M$ where r_p is the supremum r so that $B_M(p,r)$ is convex (the minimizing geodesics between any two points in $B_M(p,r)$ lie in $B_M(p,r)$).
- **Convexity radius** of M: $\rho_c(M) = \inf_{p \in M} r_p$.

Manifold constants and sampling

- **Geodesic ball**: $B_M(p, r) = \{q \mid d_M(p, q) < r\}$.
- There is a $r_p > 0$ for each $p \in M$ where r_p is the supremum r so that $B_M(p, r)$ is convex (the minimizing geodesics between any two points in $B_M(p, r)$ lie in $B_M(p, r)$).
- **Convexity radius** of M: $\rho_c(M) = \inf_{p \in M} r_p$.
- $\rho(M)$ is the **reach** defined as the minimum distance between M and its medial axis.

Manifold constants and sampling

- **Geodesic ball**: $B_M(p, r) = \{q \mid d_M(p, q) < r\}$.
- There is a $r_p > 0$ for each $p \in M$ where r_p is the supremum r so that $B_M(p, r)$ is convex (the minimizing geodesics between any two points in $B_M(p, r)$ lie in $B_M(p, r)$).
- **Convexity radius** of M: $\rho_c(M) = \inf_{p \in M} r_p$.
- $\rho(M)$ is the **reach** defined as the minimum distance between M and its medial axis.
- P is an ε-sample of M if $B(x, \varepsilon) \cap P \neq \emptyset$ for each $x \in M$.

Approximation Theorem

- Let $M \subset \mathbb{R}^d$ be a smooth, closed manifold with l as the length of a shortest basis of $\mathsf{H}_1(M)$ and $k = rank\ \mathsf{H}_1(M)$.

Approximation Theorem

- Let $M \subset \mathbb{R}^d$ be a smooth, closed manifold with l as the length of a shortest basis of $\mathsf{H}_1(M)$ and $k = rank\ \mathsf{H}_1(M)$.
- Given a set $P \subset M$ of n points which is an ε-sample of M and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\sqrt{\frac{3}{5}}\rho(M), \rho_c(M)\}$, one can compute a set of loops G in $O(nn_e^2 n_t)$ time where

$$\frac{1}{1+\frac{4r^2}{3\rho^2(M)}} l \leq \mathrm{Len}(G) \leq (1+\frac{4\varepsilon}{r})l.$$

Here n_e, n_t are the number of edges and triangles in $\mathcal{R}^{2r}(P)$.

Using Rips complexes

- Let ι^*: $\mathsf{H}_1(\mathcal{R}^r(P)) \rightarrow \mathsf{H}_1(\mathcal{R}^{2r}(P))$ and $\mathsf{H}_1^{r,2r}(P)$=im ι^*.

Using Rips complexes

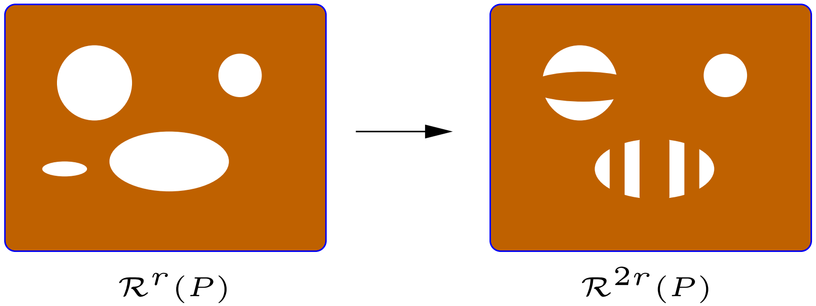
- Let ι^*: $\mathsf{H}_1(\mathcal{R}^r(P)) \rightarrow \mathsf{H}_1(\mathcal{R}^{2r}(P))$ and $\mathsf{H}_1^{r,2r}(P)$=im ι^*.
- A shortest basis of $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$ coincides with a shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$.

Using Rips complexes

- Let ι^*: $\mathsf{H}_1(\mathcal{R}^r(P)) \rightarrow \mathsf{H}_1(\mathcal{R}^{2r}(P))$ and $\mathsf{H}_1^{r,2r}(P)$=im ι^*.
- A shortest basis of $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$ coincides with a shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$.
- A shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$ has length within a small factor of the length of a shortest basis of $\mathsf{H}_1(M)$.

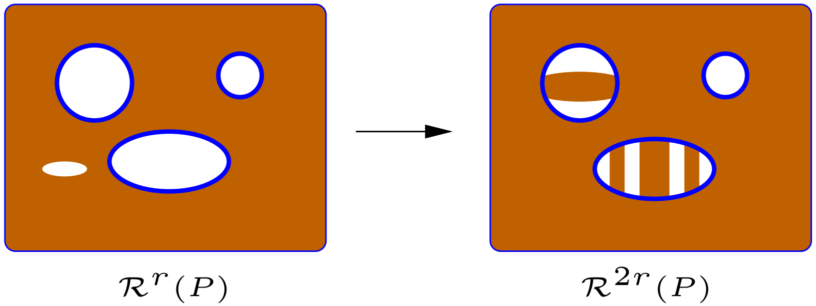
Using Rips complexes

- Let ι^*: $\mathsf{H}_1(\mathcal{R}^r(P)) \to \mathsf{H}_1(\mathcal{R}^{2r}(P))$ and $\mathsf{H}_1^{r,2r}(P)$=im ι^*.
- A shortest basis of $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$ coincides with a shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$.
- A shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$ has length within a small factor of the length of a shortest basis of $\mathsf{H}_1(M)$.
- We weight edges in $\mathcal{R}^{2r}(P)$, creating a complex $\mathcal{K}$. Each edge $e \in \mathcal{R}^{2r}(P) \setminus \mathcal{R}^r(P)$ has a large weight W. Other edges have lengths as their weights.



Using Rips complexes

- Let ι^*: $\mathsf{H}_1(\mathcal{R}^r(P)) \to \mathsf{H}_1(\mathcal{R}^{2r}(P))$ and $\mathsf{H}_1^{r,2r}(P)$=im ι^*.
- A shortest basis of $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$ coincides with a shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$.
- A shortest basis of $\mathsf{H}_1(\mathcal{C}^r(P))$ has length within a small factor of the length of a shortest basis of $\mathsf{H}_1(M)$.
- We weight edges in $\mathcal{R}^{2r}(P)$, creating a complex $\mathcal{K}$. Each edge $e \in \mathcal{R}^{2r}(P) \setminus \mathcal{R}^r(P)$ has a large weight W. Other edges have lengths as their weights.



Approximation Algorithm

Theorem SHORTLOOP(P, r) computes a shortest basis for the persistent homology group $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$.

SHORTLOOP(P, r)

1: Compute Rips complex $\mathcal{R}^{2r}(P)$.

Approximation Algorithm

Theorem ShortLoop(P, r) computes a shortest basis for the persistent homology group $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$.

ShortLoop(P, r)

1: Compute Rips complex $\mathcal{R}^{2r}(P)$.

2: Let $\mathcal{K}$ be $\mathcal{R}^{2r}(P)$ where edges of $\mathcal{R}^{2r}(P) \setminus \mathcal{R}^{r}(P)$ are weighted with large weight W.

Approximation Algorithm

Theorem ShortLoop(P, r) computes a shortest basis for the persistent homology group $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$.

ShortLoop(P, r)

1: Compute Rips complex $\mathcal{R}^{2r}(P)$.

2: Let $\mathcal{K}$ be $\mathcal{R}^{2r}(P)$ where edges of $\mathcal{R}^{2r}(P) \backslash \mathcal{R}^{r}(P)$ are weighted with large weight W.

3: Compute the shortest basis for $\mathsf{H}_1(\mathcal{K})$.

Approximation Algorithm

Theorem SHORTLOOP(P, r) computes a shortest basis for the persistent homology group $\mathsf{H}_1^{r,2r}(\mathcal{R}(P))$.

SHORTLOOP(P, r)

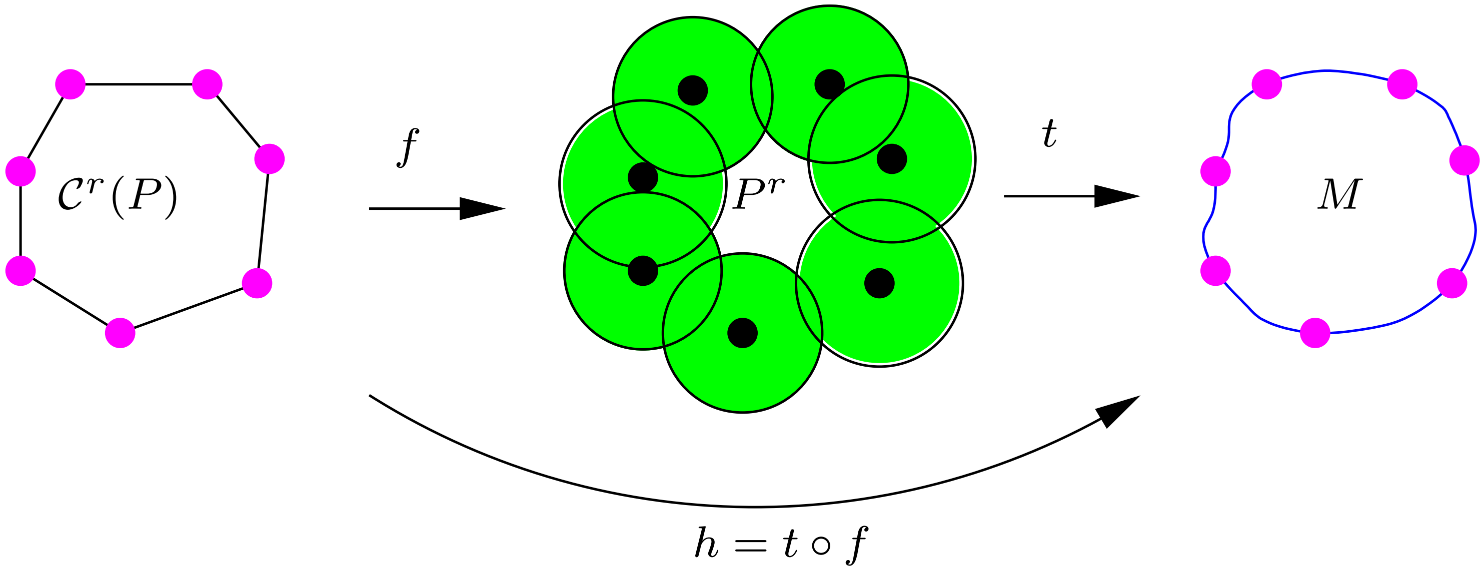
1: Compute Rips complex $\mathcal{R}^{2r}(P)$.

2: Let $\mathcal{K}$ be $\mathcal{R}^{2r}(P)$ where edges of $\mathcal{R}^{2r}(P) \setminus \mathcal{R}^{r}(P)$ are weighted with large weight W.

3: Compute the shortest basis for $\mathsf{H}_1(\mathcal{K})$.

4: Return first k loops from the computed basis where k is the rank of the $\mathsf{H}_1(\mathcal{R}^r(P)) \to \mathsf{H}_1(\mathcal{R}^{2r}(P))$.

Connecting $\mathcal{C}^r(P)$ and M



Bounding Lengths

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.

Bounding Lengths

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.
- Let g be a geodesic loop in M. There is a loop $\hat{g}$ in $\mathcal{C}^r(P)$ so that $[h(\hat{g})] = [g]$ where h is a homotopy equivalence and $Len(\hat{g}) \leq (1 + \frac{4\varepsilon}{r})Len(g)$.

Upper bound

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.

Upper bound

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.
- If $G = \{g_1, ..., g_k\}$ and $G' = \{g'_1, ..., g'_k\}$ are the generators of a shortest basis of $\mathsf{H}_1(M)$ and $\mathsf{H}_1(\mathcal{K})$ respectively, then we have $\mathsf{Len}(G') \leq (1 + \frac{4\varepsilon}{r})\mathsf{Len}(G)$.

Lower bound

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.

Lower bound

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.
- Let G and G' be defined as before.

Lower bound

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\rho(M), \rho_c(M)\}$.
- Let G and G' be defined as before.
- We have Len$(G) \leq (1 + \frac{4r^2}{3\rho^2(M)})Len(G')$.

Length Approximation Theorem

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\sqrt{\frac{3}{5}}\rho(M), \rho_c(M)\}$.

Length Approximation Theorem

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\sqrt{\frac{3}{5}}\rho(M), \rho_c(M)\}$.
- Let G and G' be a shortest basis of $\mathsf{H}_1(M)$ and $\mathsf{H}_1(\mathcal{K})$ respectively.

Length Approximation Theorem

- Let $P \subset M$ be an ε-sample and $4\varepsilon \leq r \leq \min\{\frac{1}{2}\sqrt{\frac{3}{5}}\rho(M), \rho_c(M)\}$.
- Let G and G' be a shortest basis of $\mathsf{H}_1(M)$ and $\mathsf{H}_1(\mathcal{K})$ respectively.
- We have $\frac{1}{1+\frac{4r^2}{3\rho^2(M)}}\mathsf{Len}(G) \leq \mathsf{Len}(G') \leq (1+\frac{4\varepsilon}{r})\mathsf{Len}(G)$.

Conclusions

- Algorithms for shortest basis of the first homology groups of a simplicial complex and point sampled manifolds.

Conclusions

- Algorithms for shortest basis of the first homology groups of a simplicial complex and point sampled manifolds.
- What about higher dimensional homology groups? Recent results of [CF10] indicate that this problem is NP-hard.

Conclusions

- Algorithms for shortest basis of the first homology groups of a simplicial complex and point sampled manifolds.
- What about higher dimensional homology groups? Recent results of [CF10] indicate that this problem is NP-hard.
- What about shortest homologous loop/cycle? Under Z_2 the problem is NP-hard [CF10]. Under Z, surprisingly this optimization is polynomial time solvable for a large class [DHK10].

Conclusions

- Algorithms for shortest basis of the first homology groups of a simplicial complex and point sampled manifolds.
- What about higher dimensional homology groups? Recent results of [CF10] indicate that this problem is NP-hard.
- What about shortest homologous loop/cycle? Under Z_2 the problem is NP-hard [CF10]. Under Z, surprisingly this optimization is polynomial time solvable for a large class [DHK10].
- Software **ShortLoop** is available from authors' web-pages.

Thank you!