Persistent homology

From theory to computation

Ulrich Bauer

TUM

Oct 24, 2019

GAMM workshop
Computational and Mathematical Methods in Data Science
Zuse Instutite Berlin

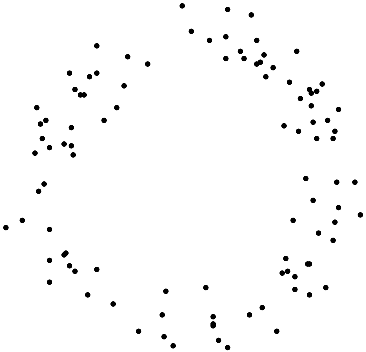
$\cong$

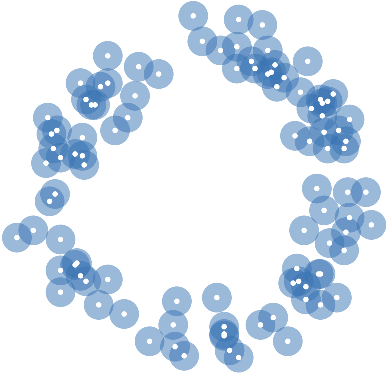
?

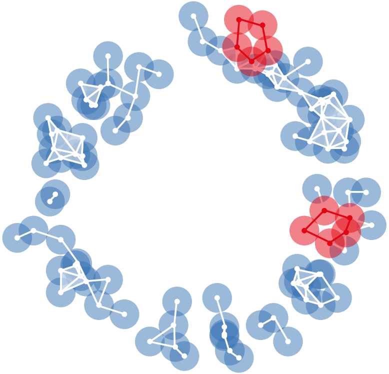
$\cong$

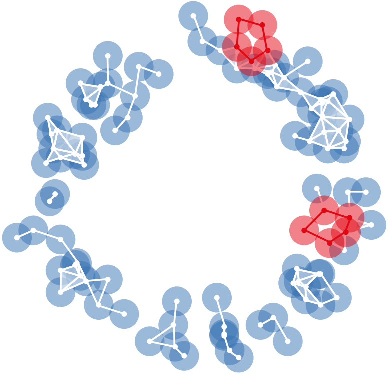
?

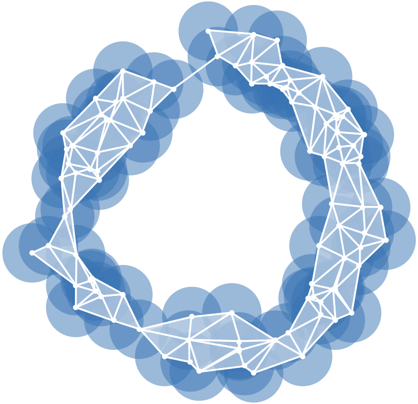
Persistent homology

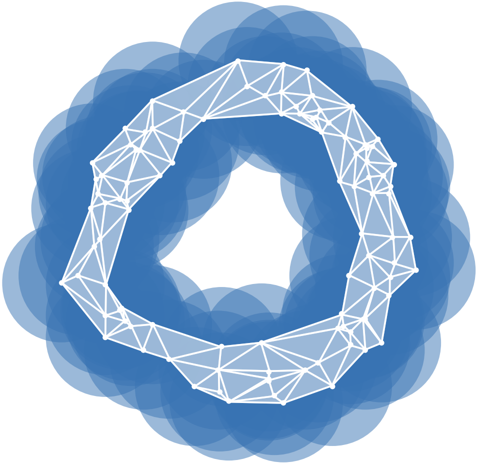


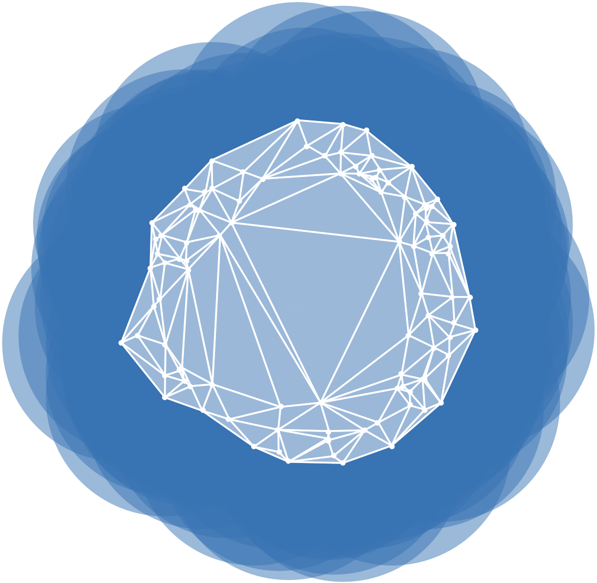




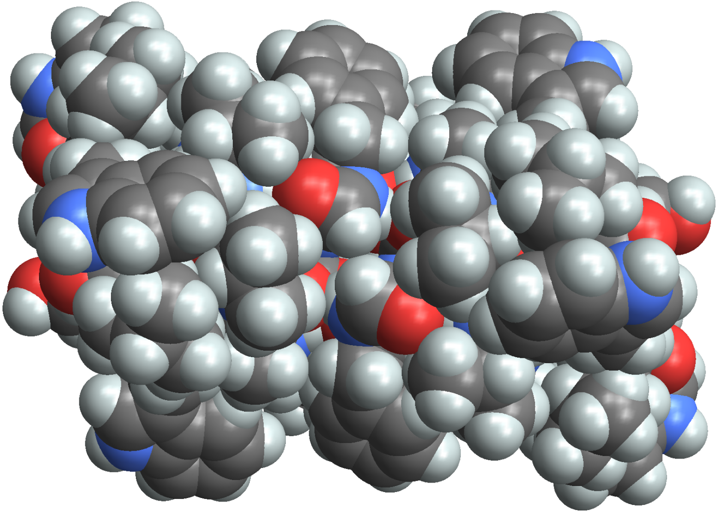




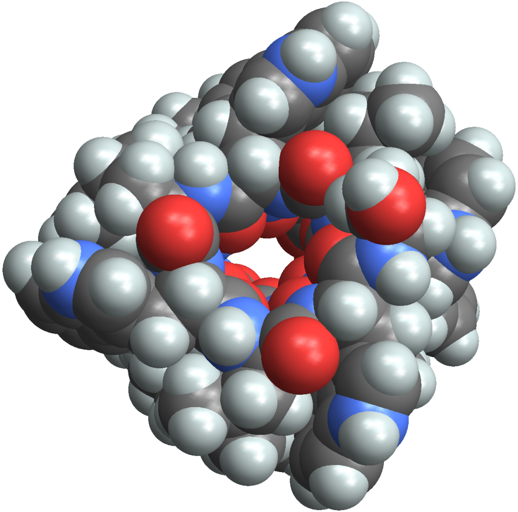




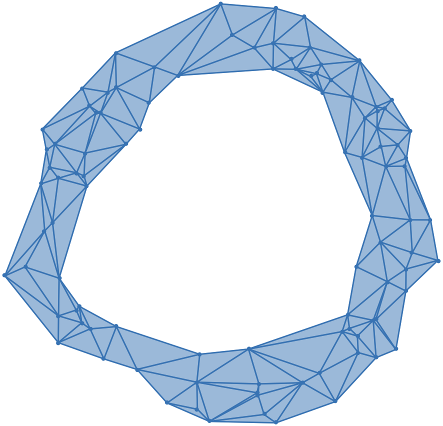
Gramicidin: an antibiotic functioning as ion channel



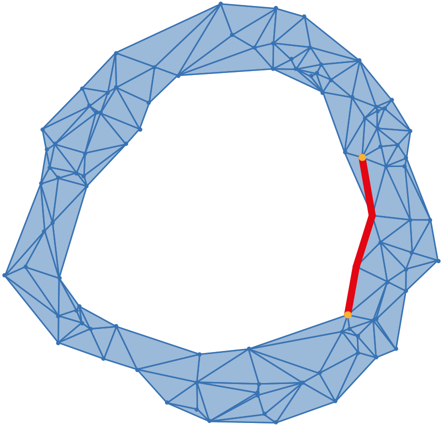
Gramicidin: an antibiotic functioning as ion channel



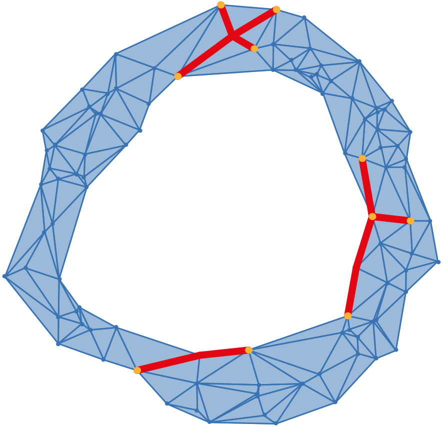
Homology



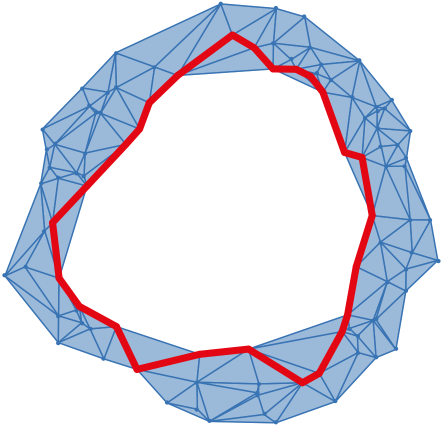
Homology



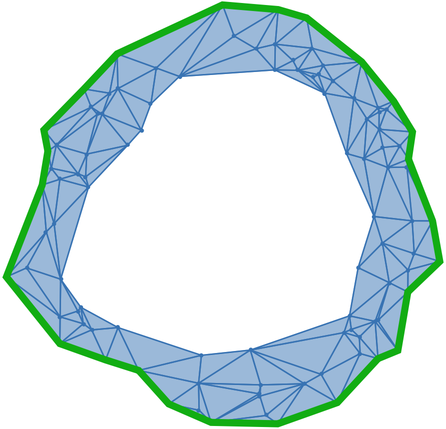
Homology



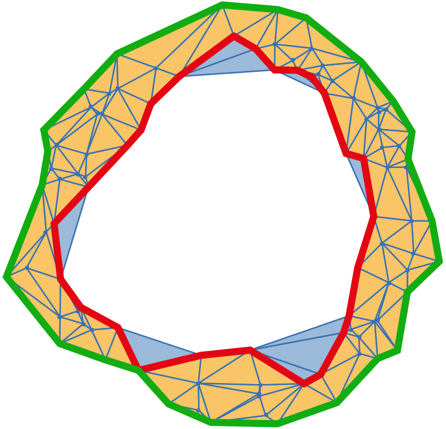
Homology



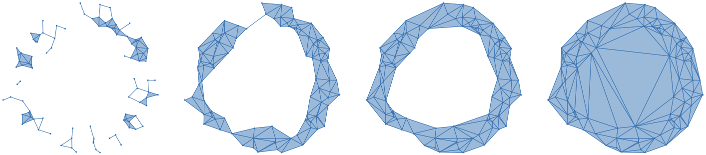
Homology



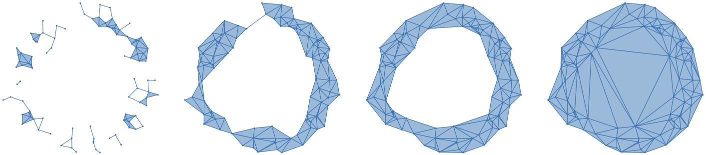
Homology

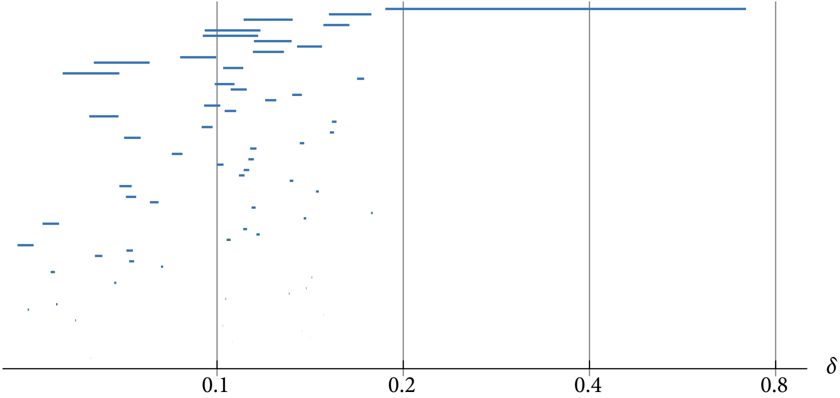


What is persistent homology?

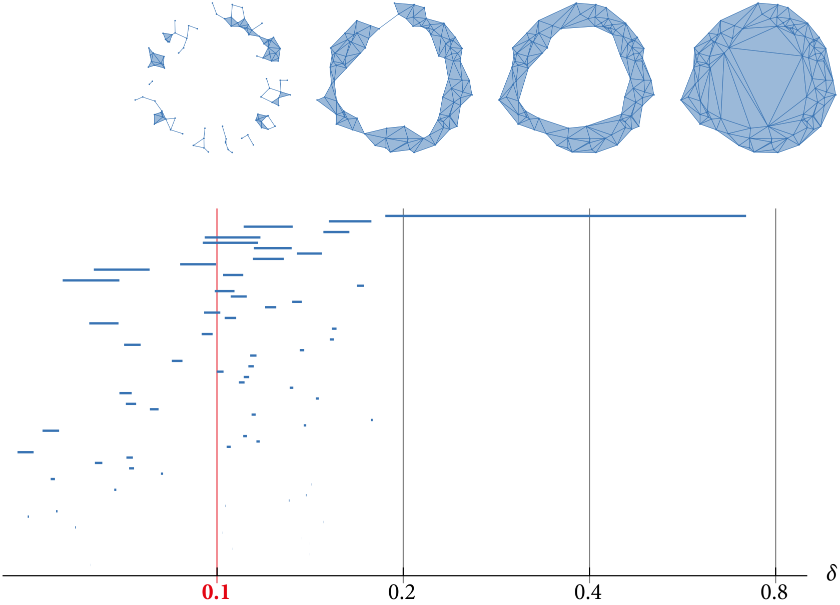


What is persistent homology?

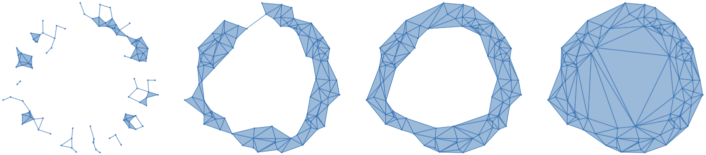


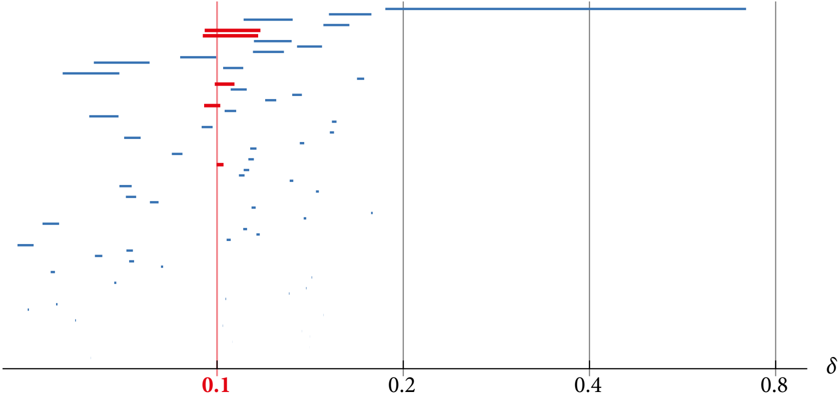


What is persistent homology?

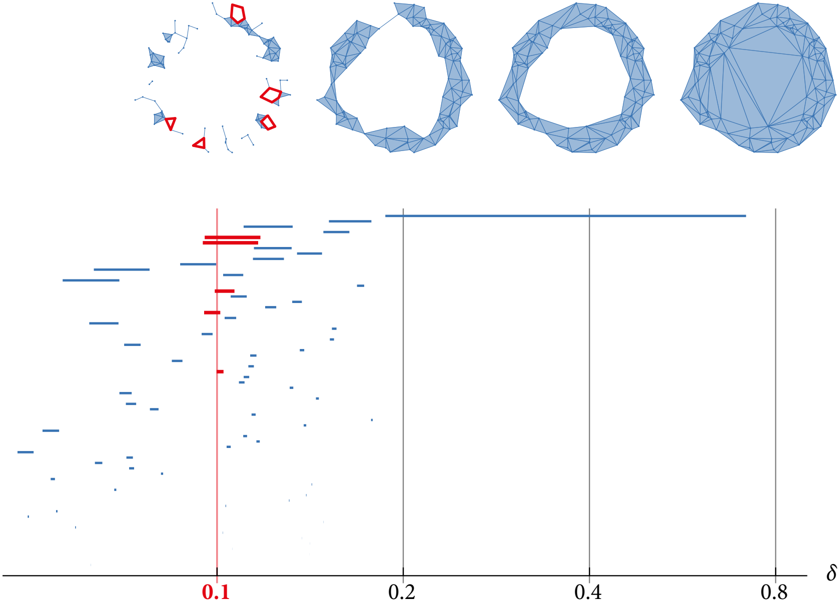


What is persistent homology?

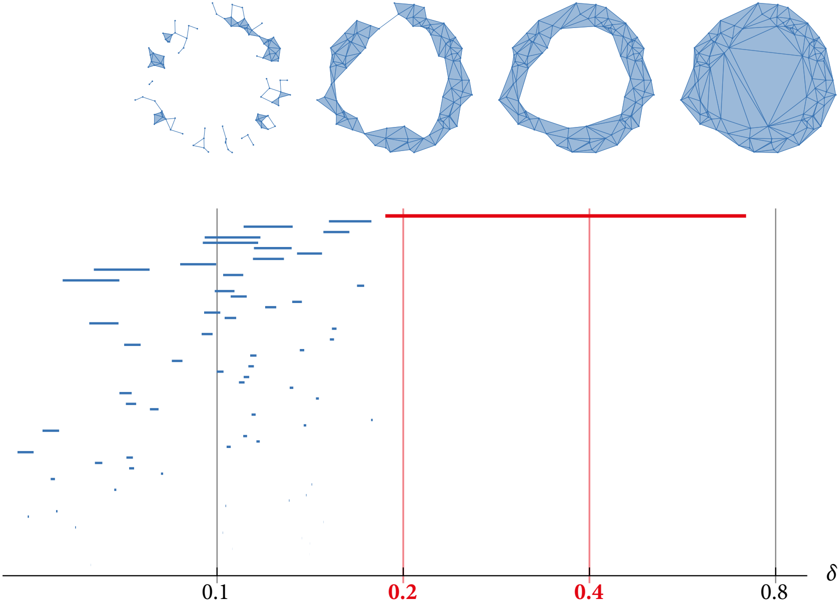




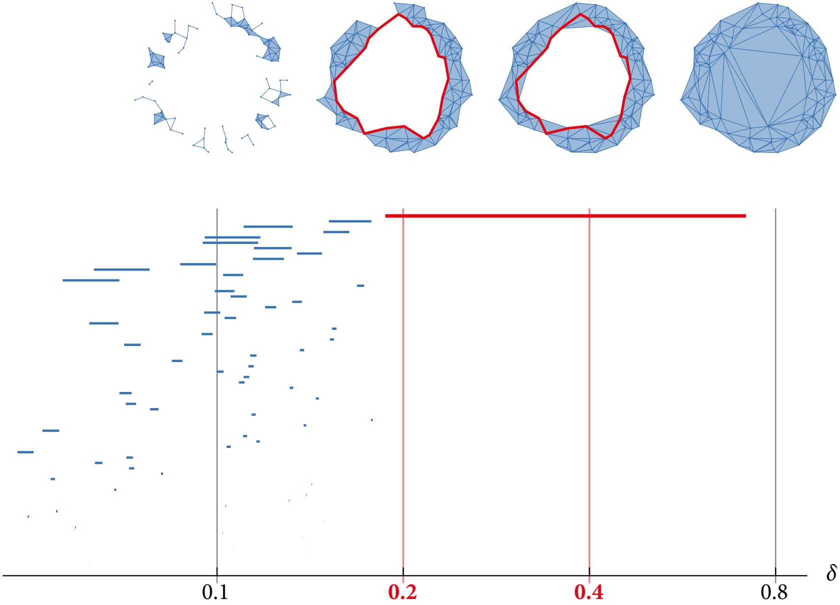
What is persistent homology?



What is persistent homology?



What is persistent homology?



What is persistent homology?

What is persistent homology?

Persistent homology is the homology of a filtration.

What is persistent homology?

Persistent homology is the homology of a filtration.

- A filtration is a certain commutative diagram $K : \mathbf{R} \to \mathbf{Top}$:
  - $\mathbf{R}$ is the partially ordered set $(\mathbb{R}, \leq)$
  - A topological space K_t for each $t \in \mathbb{R}$
  - An inclusion map $K_s \hookrightarrow K_t$ for each $s \leq t \in \mathbb{R}$

What is persistent homology?

Persistent homology is the homology of a filtration.

- A filtration is a certain commutative diagram $K : \mathbf{R} \to \mathbf{Top}$:
  - $\mathbf{R}$ is the partially ordered set $(\mathbb{R}, \leq)$
  - A topological space K_t for each $t \in \mathbb{R}$
  - An inclusion map $K_s \hookrightarrow K_t$ for each $s \leq t \in \mathbb{R}$
- Apply homology $H_* : \mathbf{Top} \to \mathbf{Vect}$

What is persistent homology?

Persistent homology is the homology of a filtration.

- A filtration is a certain commutative diagram $K : \mathbf{R} \to \mathbf{Top}$:
  - $\mathbf{R}$ is the partially ordered set $(\mathbb{R}, \leq)$
  - A topological space K_t for each $t \in \mathbb{R}$
  - An inclusion map $K_s \hookrightarrow K_t$ for each $s \leq t \in \mathbb{R}$
- Apply homology $H_* : \mathbf{Top} \to \mathbf{Vect}$
- Persistent homology is a commutative diagram $M : \mathbf{R} \to \mathbf{Vect}$ *(persistence module)*:
  - A vector space M_t for each $t \in \mathbb{R}$
  - A linear map $M_s \to M_t$ for each $s \leq t \in \mathbb{R}$

What is persistent homology?

Persistent homology is the homology of a filtration.

- A filtration is a certain commutative diagram $K : \mathbf{R} \to \mathbf{Top}$:
  - $\mathbf{R}$ is the partially ordered set $(\mathbb{R}, \leq)$
  - A topological space K_t for each $t \in \mathbb{R}$
  - An inclusion map $K_s \hookrightarrow K_t$ for each $s \leq t \in \mathbb{R}$
- Apply homology $H_* : \mathbf{Top} \to \mathbf{Vect}$
- Persistent homology is a commutative diagram $M : \mathbf{R} \to \mathbf{Vect}$ *(persistence module)*:
  - A vector space M_t for each $t \in \mathbb{R}$
  - A linear map $M_s \to M_t$ for each $s \leq t \in \mathbb{R}$

In this talk, all vector spaces will be finite dimensional.

Barcodes: the structure of persistence modules

Theorem (Crawley-Boevey 2015)

Any persistence module $M : \mathbf{R} \to \mathbf{vect}$ *(of finite dim. vector spaces over some field* $\mathbb{F}$*) decomposes as a direct sum of* interval modules

$$0 \to \cdots \to 0 \to \mathbb{F} \to \cdots \to \mathbb{F} \to 0 \to \cdots \to 0$$

(in an essentially unique way).

Barcodes: the structure of persistence modules

Theorem (Crawley-Boevey 2015)

Any persistence module $M : \mathbf{R} \to \mathbf{vect}$ *(of finite dim. vector spaces over some field* $\mathbb{F}$*) decomposes as a direct sum of* interval modules

$$0 \to \cdots \to 0 \to \color{red}{\mathbb{F} \to \cdots \to \mathbb{F}} \to 0 \to \cdots \to 0$$

(in an essentially unique way).

- The supporting intervals form the *persistence barcode* $B(M)$.

Barcodes: the structure of persistence modules

Theorem (Crawley-Boevey 2015)

Any persistence module $M : \mathbf{R} \to \mathbf{vect}$ *(of finite dim. vector spaces over some field* $\mathbb{F}$*) decomposes as a direct sum of* interval modules

$$0 \to \cdots \to 0 \to \color{red}{\mathbb{F} \to \cdots \to \mathbb{F}} \to 0 \to \cdots \to 0$$

(in an essentially unique way).

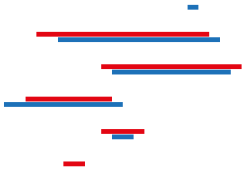
- The supporting intervals form the *persistence barcode* $B(M)$.
- The barcode completely describes the persistence module (up to isomorphism).

Stability

Stability of persistence barcodes for functions

Theorem (Cohen-Steiner, Edelsbrunner, Harer 2005)

Let $f, g : X \to \mathbb{R}$ *with* $\|f - g\|_\infty = \delta$ *(and some regularity assumptions).*
Consider the persistence barcodes of (sublevel set filtrations of) f *and* g*.*
Then there exists a matching between their intervals such that

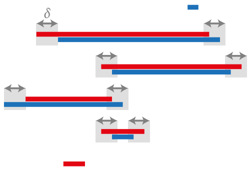


Stability of persistence barcodes for functions

Theorem (Cohen-Steiner, Edelsbrunner, Harer 2005)

Let $f, g : X \to \mathbb{R}$ with $\|f - g\|_\infty = \delta$ (and some regularity assumptions).
Consider the persistence barcodes of (sublevel set filtrations of) f and g.
Then there exists a matching between their intervals such that

- *matched intervals have endpoints within distance $\leq \delta$, and*

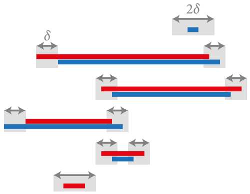


Stability of persistence barcodes for functions

Theorem (Cohen-Steiner, Edelsbrunner, Harer 2005)

Let $f, g : X \to \mathbb{R}$ with $\|f - g\|_\infty = \delta$ (and some regularity assumptions).
Consider the persistence barcodes of (sublevel set filtrations of) f and g.
Then there exists a matching between their intervals such that

- *matched intervals have endpoints within distance $\leq \delta$, and*
- *unmatched intervals have length $\leq 2\delta$.*



Persistence and stability: the big picture

Data	point cloud	$P \subset \mathbb{R}^d$

Persistence and stability: the big picture

$$
\begin{tikzcd}
\text{Data} \arrow[d] & \text{point cloud} \arrow[d, "\text{distance}"] & P \subset \mathbb{R}^d \\
\text{Geometry} & \text{function} & f : \mathbb{R}^d \to \mathbb{R}
\end{tikzcd}
$$

Persistence and stability: the big picture

$$
\begin{array}{ccc}
\text{Data} & \text{point cloud} & P \subset \mathbb{R}^d \\
\downarrow & \big\downarrow \text{distance} & \\
\text{Geometry} & \text{function} & f : \mathbb{R}^d \to \mathbb{R} \\
\downarrow & \big\downarrow \text{sublevel sets} & \\
\text{Topology} & \text{topological spaces (filtration)} & K : \mathbf{R} \to \mathbf{Top}
\end{array}
$$

Persistence and stability: the big picture

$$
\begin{tikzcd}
\text{Data} \arrow[d] & \text{point cloud} \arrow[d, "\text{distance}"] & P \subset \mathbb{R}^d \\
\text{Geometry} \arrow[d] & \text{function} \arrow[d, "\text{sublevel sets}"] & f : \mathbb{R}^d \to \mathbb{R} \\
\text{Topology} \arrow[d] & \text{topological spaces (filtration)} \arrow[d, "\text{homology}"] & K : \mathbf{R} \to \mathbf{Top} \\
\text{Algebra} & \text{vector spaces (persistence module)} & M : \mathbf{R} \to \mathbf{Vect}
\end{tikzcd}
$$

Persistence and stability: the big picture

$$
\begin{tikzcd}
\text{Data} \arrow[d] & \text{point cloud} \arrow[d, "\text{distance}"] & P \subset \mathbb{R}^d \\
\text{Geometry} \arrow[d] & \text{function} \arrow[d, "\text{sublevel sets}"] & f : \mathbb{R}^d \to \mathbb{R} \\
\text{Topology} \arrow[d] & \text{topological spaces (filtration)} \arrow[d, "\text{homology}"] & K : \mathbf{R} \to \mathbf{Top} \\
\text{Algebra} \arrow[d] & \text{vector spaces (persistence module)} \arrow[d, "\text{structure theorem}"] & M : \mathbf{R} \to \mathbf{Vect} \\
\text{Combinatorics} & \text{intervals (persistence barcode)} & B : \mathbf{R} \to \mathbf{Mch}
\end{tikzcd}
$$

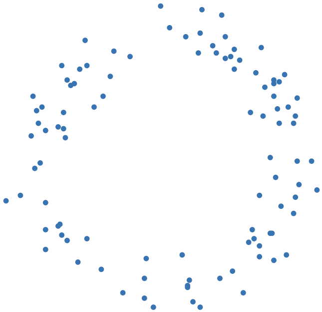
Persistence and stability: the big picture

$$
\begin{tikzcd}
\text{Data} \arrow[d] & \text{point cloud} \arrow[d, "\text{distance}"] & P \subset \mathbb{R}^d \\
\text{Geometry} \arrow[d] & \text{function} \arrow[d, "\text{sublevel sets}"] \arrow[dddd, bend right, red] & f : \mathbb{R}^d \to \mathbb{R} \\
\text{Topology} \arrow[d] & \text{topological spaces (filtration)} \arrow[d, "\text{homology}"] & K : \mathbf{R} \to \mathbf{Top} \\
\text{Algebra} \arrow[d] & \text{vector spaces (persistence module)} \arrow[d, "\text{structure theorem}"] & M : \mathbf{R} \to \mathbf{Vect} \\
\text{Combinatorics} & \text{intervals (persistence barcode)} & B : \mathbf{R} \to \mathbf{Mch}
\end{tikzcd}
$$

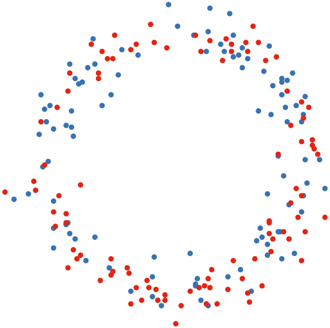
Persistence and stability: the big picture

$$
\begin{tikzcd}
\text{Data} \arrow[d] & \text{point cloud} \arrow[d, "\text{distance}"] & P \subset \mathbb{R}^d \\
\text{Geometry} \arrow[d] & \text{function} \arrow[d, "\text{sublevel sets}"] & f : \mathbb{R}^d \to \mathbb{R} \\
\text{Topology} \arrow[d] & \text{topological spaces (filtration)} \arrow[d, "\text{homology}"] & K : \mathbf{R} \to \mathbf{Top} \\
\text{Algebra} \arrow[d] & \text{vector spaces (persistence module)} \arrow[d, "\text{structure theorem}"] & M : \mathbf{R} \to \mathbf{Vect} \\
\text{Combinatorics} & \text{intervals (persistence barcode)} & B : \mathbf{R} \to \mathbf{Mch}
\end{tikzcd}
$$

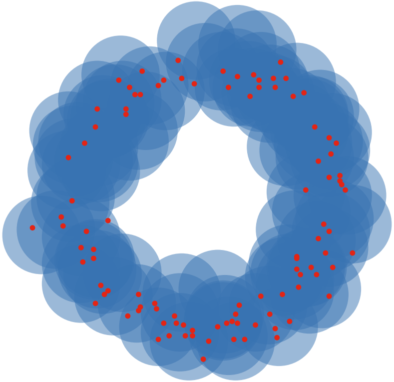
Geometric interleavings



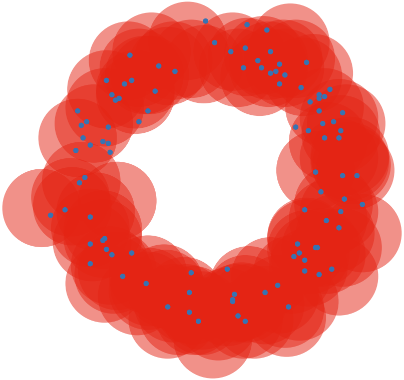
Geometric interleavings



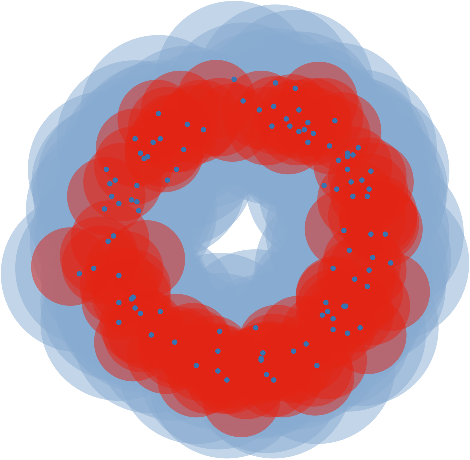
Geometric interleavings



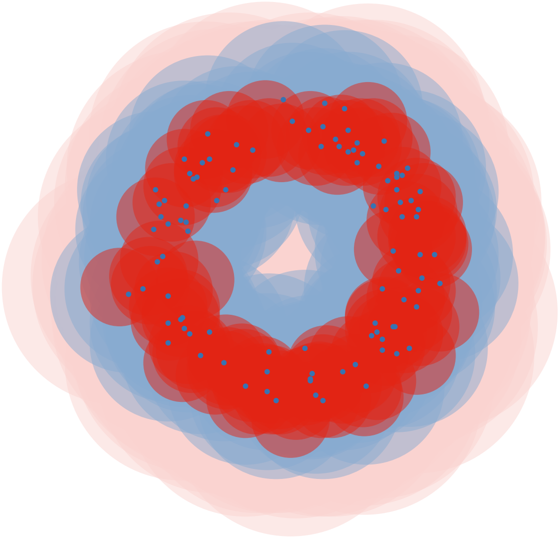
Geometric interleavings



Geometric interleavings



Geometric interleavings



Interleavings

Let $\delta = \|f - g\|_\infty$. Write $F_t = f^{-1}(-\infty, t]$ for the t-sublevel set of f.

Interleavings

Let $\delta = \|f - g\|_\infty$. Write $F_t = f^{-1}(-\infty, t]$ for the t-sublevel set of f.

Then the sublevel set filtrations $F, G : \mathbf{R} \to \mathbf{Top}$ are δ-*interleaved:*

$$\begin{tikzcd}
\cdots \arrow[r, dotted] & F_t \arrow[r, hook] \arrow[dr, hook] & F_{t+\delta} \arrow[r, hook] \arrow[dr, hook] & F_{t+2\delta} \arrow[r, dotted] & \cdots \\
\cdots \arrow[r, dotted] & G_t \arrow[r, hook] \arrow[ur, hook] & G_{t+\delta} \arrow[r, hook] \arrow[ur, hook] & G_{t+2\delta} \arrow[r, dotted] & \cdots
\end{tikzcd} \qquad \forall t \in \mathbb{R}.$$

Interleavings

Let $\delta = \|f - g\|_\infty$. Write $F_t = f^{-1}(-\infty, t]$ for the t-sublevel set of f.

Then the sublevel set filtrations $F, G : \mathbf{R} \to \mathbf{Top}$ are δ-*interleaved:*

$$\begin{tikzcd}
\cdots \arrow[r, dotted] & F_t \arrow[r, hook] \arrow[dr, hook] & F_{t+\delta} \arrow[r, hook] \arrow[dr, hook] & F_{t+2\delta} \arrow[r, dotted] & \cdots \\
\cdots \arrow[r, dotted] & G_t \arrow[r, hook] \arrow[ur, hook] & G_{t+\delta} \arrow[r, hook] \arrow[ur, hook] & G_{t+2\delta} \arrow[r, dotted] & \cdots
\end{tikzcd} \qquad \forall t \in \mathbb{R}.$$

Applying homology (a functor) preserves commutativity

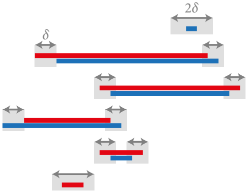
- persistent homology of f, g yields δ-interleaved persistence modules $\mathbf{R} \to \mathbf{Vect}$

Algebraic stability of persistence barcodes

Theorem (Chazal et al. 2009, 2012; B, Lesnick 2015)

If two persistence modules are δ-interleaved, then there exists a δ-matching of their barcodes:

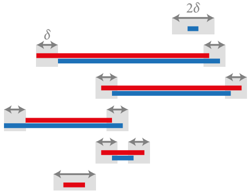
- matched intervals have endpoints within distance $\leq \delta$,
- unmatched intervals have length $\leq 2\delta$.



Algebraic stability of persistence barcodes

Theorem (Chazal et al. 2009, 2012; B, Lesnick 2015)

If two persistence modules are δ-interleaved,
then there exists a δ-matching of their barcodes:



- matched intervals have endpoints within distance $\leq \delta$,
- unmatched intervals have length $\leq 2\delta$.

Equivalently: there exists a δ-interleaving of their barcodes (as diagrams $\mathbf{R} \to \mathbf{Mch}$).

Structure of persistence submodules and quotients

Proposition

Let $M \twoheadrightarrow N$ be an epimorphism.

Then there is an injection of barcodes $B(N) \hookrightarrow B(M)$ such that if J is mapped to I, then

- *I and J are aligned below, and*
- *I bounds J above.*

This construction is functorial (compatible with composition and identity).

Structure of persistence submodules and quotients

Proposition

Let $M \twoheadrightarrow N$ be an epimorphism.

Then there is an injection of barcodes $B(N) \hookrightarrow B(M)$ such that if J is mapped to I, then

- *I and J are aligned below, and*
- *I bounds J above.*

This construction is functorial (compatible with composition and identity).

Dually, there is an injection $B(N) \hookrightarrow B(O)$ for monomorphisms $N \hookrightarrow O$.

Induced matchings

For $f : M \to N$ a morphism of pfd persistence modules, the epi-mono factorization

$$M \twoheadrightarrow \operatorname{im} f \hookrightarrow N$$

gives an *induced matching* $\chi(f)$ between their barcodes:

Induced matchings

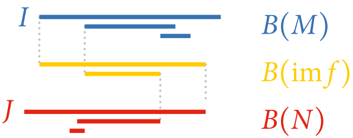
For $f : M \to N$ a morphism of pfd persistence modules, the epi-mono factorization

$$M \twoheadrightarrow \operatorname{im} f \hookrightarrow N$$

gives an *induced matching* $\chi(f)$ between their barcodes:

- compose the injections $B(M) \hookleftarrow B(\operatorname{im} f) \hookrightarrow B(N)$ from before to a matching

$$\chi(f) : B(M) \nrightarrow B(N).$$



Induced matchings

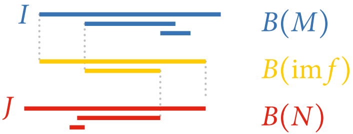
For $f : M \to N$ a morphism of pfd persistence modules, the epi-mono factorization

$$M \twoheadrightarrow \operatorname{im} f \hookrightarrow N$$

gives an *induced matching* $\chi(f)$ between their barcodes:

- compose the injections $B(M) \hookleftarrow B(\operatorname{im} f) \hookrightarrow B(N)$ from before to a matching

$$\chi(f) : B(M) \nrightarrow B(N).$$



When applied to a δ-interleaving, this yields a δ-matching of barcodes.

Homology inference

Homology inference

Given: finite sample $P \subset X$ of unknown shape $X \subset \mathbb{R}^d$

Problem (Homology inference)

Determine the homology $H_*(X)$.

Homology inference

Given: finite sample $P \subset X$ of unknown shape $X \subset \mathbb{R}^d$

Problem (Homology inference)

Determine the homology $H_*(X)$.

Problem (Homological reconstruction)

Construct a shape R *with* $H_*(R) \cong H_*(X)$.

Homology inference

Given: finite sample $P \subset X$ of unknown shape $X \subset \mathbb{R}^d$

Problem (Homology inference)

Determine the homology $H_*(X)$.

Problem (Homological reconstruction)

Construct a shape R *with* $H_*(R) \cong H_*(X)$.

- Approach: approximate the shape by a thickening $P_\delta = \bigcup_{p \in P} B_\delta(p)$ covering X

Homology inference

Given: finite sample $P \subset X$ of unknown shape $X \subset \mathbb{R}^d$

Problem (Homology inference)

Determine the homology $H_*(X)$.

Problem (Homological reconstruction)

Construct a shape R with $H_*(R) \cong H_*(X)$.

- Approach: approximate the shape by a thickening $P_\delta = \bigcup_{p \in P} B_\delta(p)$ covering X
- Requires strong assumptions:

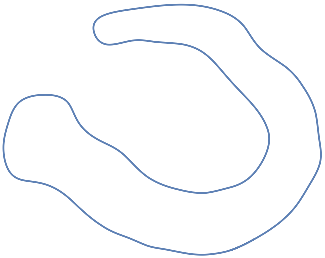
Homology reconstruction by thickening

Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.



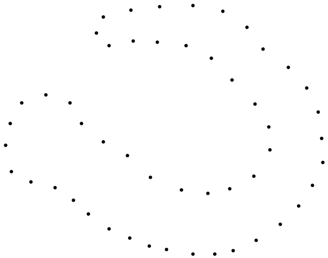
Homology reconstruction by thickening

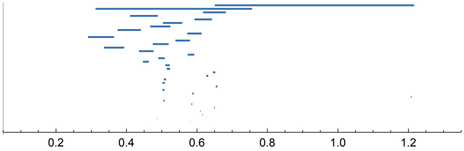
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





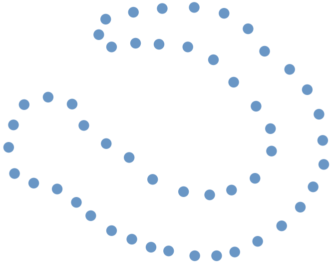
Homology reconstruction by thickening

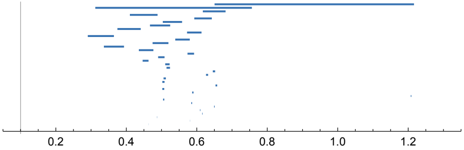
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





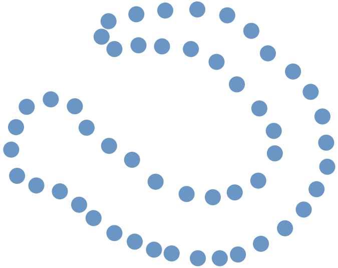
Homology reconstruction by thickening

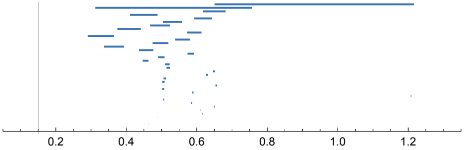
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





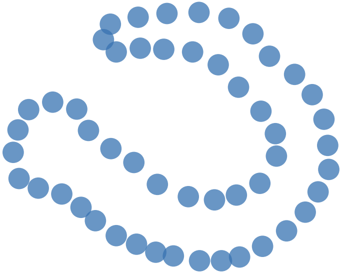
Homology reconstruction by thickening

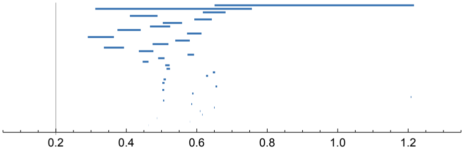
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





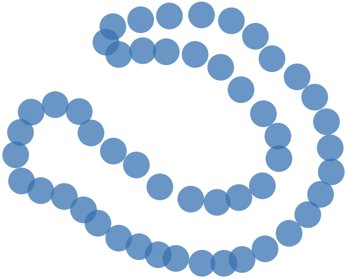
Homology reconstruction by thickening

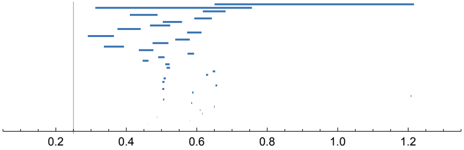
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- *$\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.*

Then $H_*(X) \cong H_*(P_{2\delta})$.





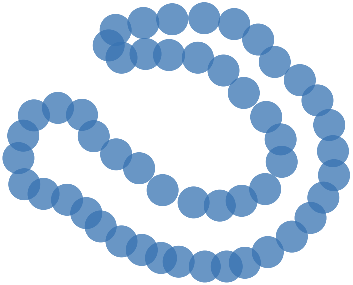
Homology reconstruction by thickening

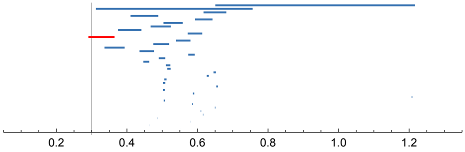
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- *$\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.*

Then $H_*(X) \cong H_*(P_{2\delta})$.





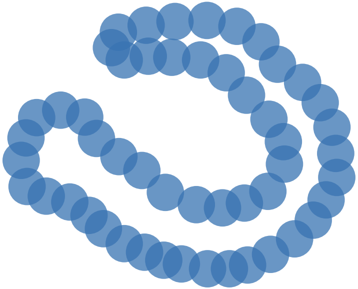
Homology reconstruction by thickening

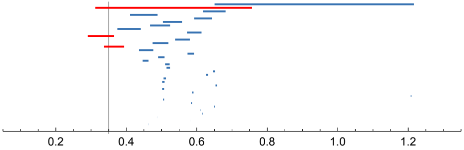
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





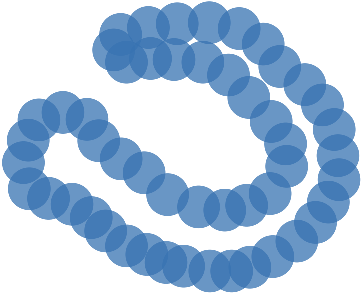
Homology reconstruction by thickening

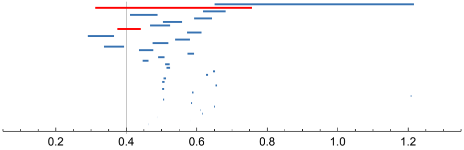
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





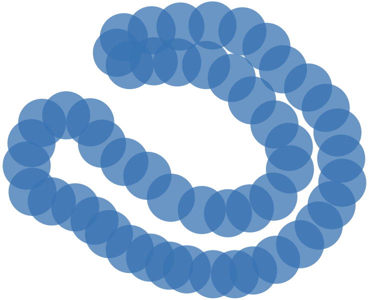
Homology reconstruction by thickening

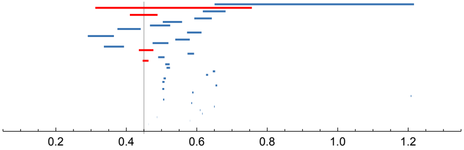
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- *$\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.*

Then $H_*(X) \cong H_*(P_{2\delta})$.





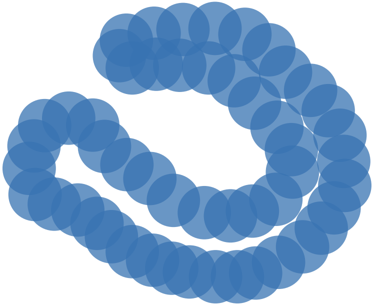
Homology reconstruction by thickening

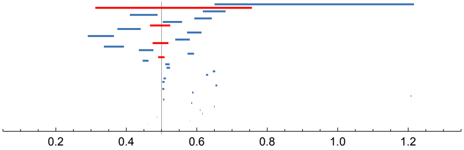
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





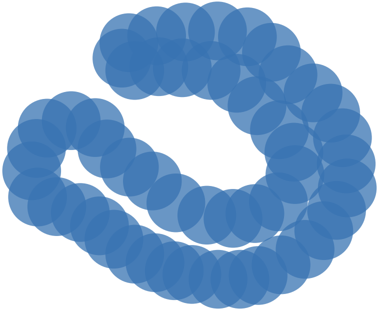
Homology reconstruction by thickening

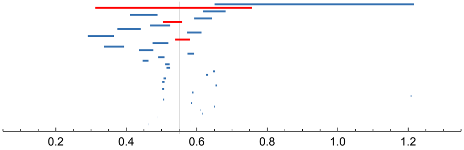
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





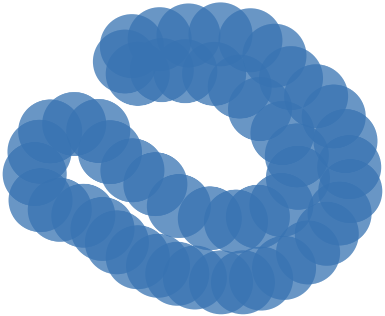
Homology reconstruction by thickening

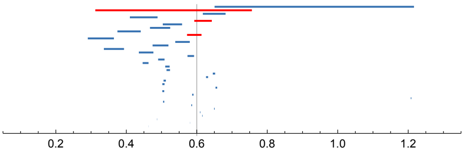
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





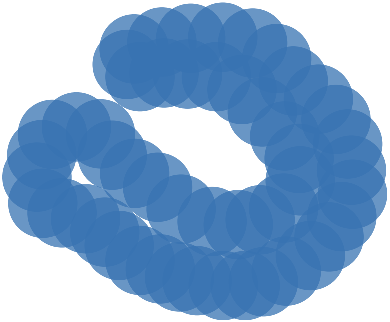
Homology reconstruction by thickening

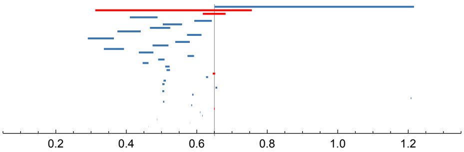
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





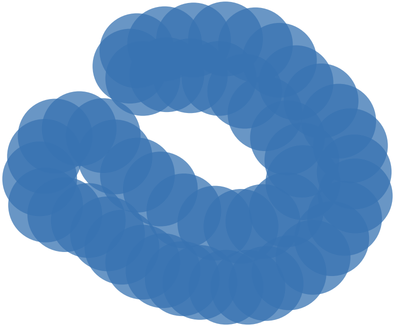
Homology reconstruction by thickening

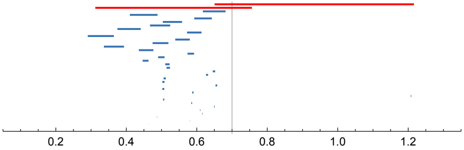
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\operatorname{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





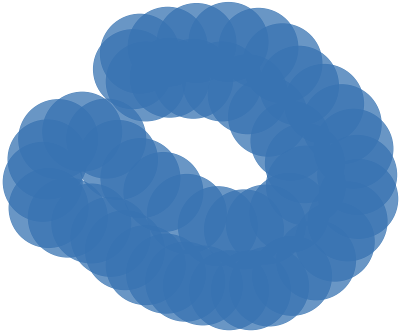
Homology reconstruction by thickening

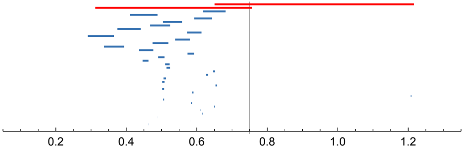
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





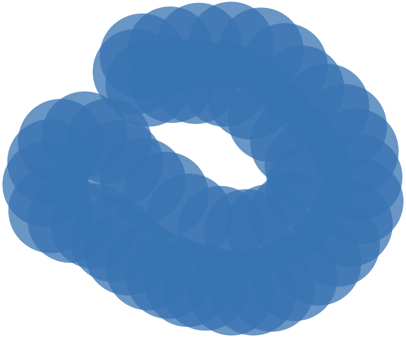
Homology reconstruction by thickening

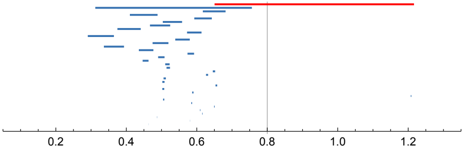
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- *$\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.*

Then $H_*(X) \cong H_*(P_{2\delta})$.





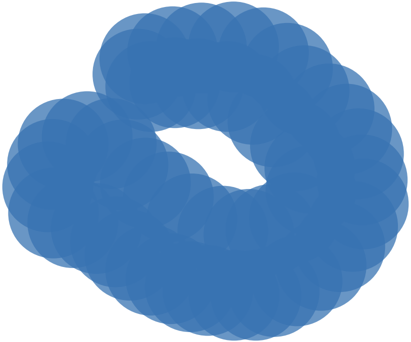
Homology reconstruction by thickening

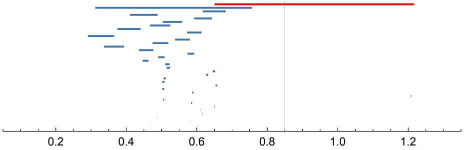
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





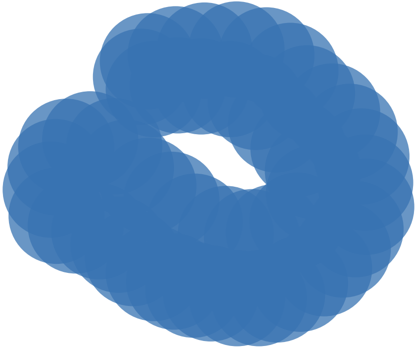
Homology reconstruction by thickening

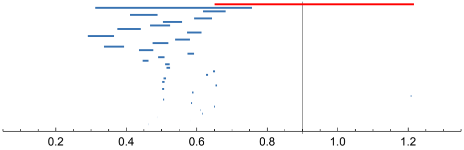
Theorem (Niyogi, Smale, Weinberger 2006)

Let X be a submanifold of $\mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- $\delta < \sqrt{3/20}\,\mathrm{reach}(X)$.

Then $H_*(X) \cong H_*(P_{2\delta})$.





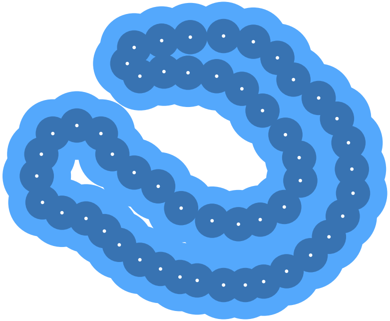
Homology inference using persistence

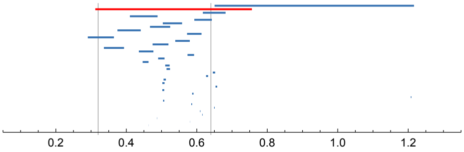
Theorem (Cohen-Steiner, Edelsbrunner, Harer 2005)

Let $X \subset \mathbb{R}^d$. Let $P \subset X$ and $\delta > 0$ be such that

- *P_δ covers X, and*
- *the induced maps $H_*(X \hookrightarrow X_\delta)$ and $H_*(X_\delta \hookrightarrow X_{2\delta})$ are isomorphisms.*

Then $H_*(X) \cong \operatorname{im} H_*(P_\delta \hookrightarrow P_{2\delta})$.





Homological realization

This motivates the *homological realization problem*:

Problem

Given a pair $L \subseteq K$ of simplicial complexes, find a third complex X with $L \subseteq X \subseteq K$ such that $H_*(X) \cong \operatorname{im} H_*(L \hookrightarrow K)$.

Theorem (Attali, B, Devillers, Glisse, Lieutier 2013)

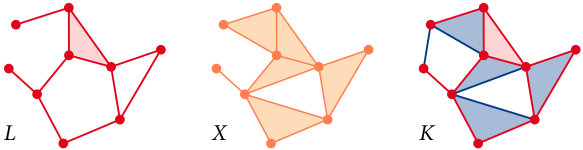
The homological realization problem is NP-hard, even in $\mathbb{R}^3$.

Homological realization

This motivates the *homological realization problem*:

Problem

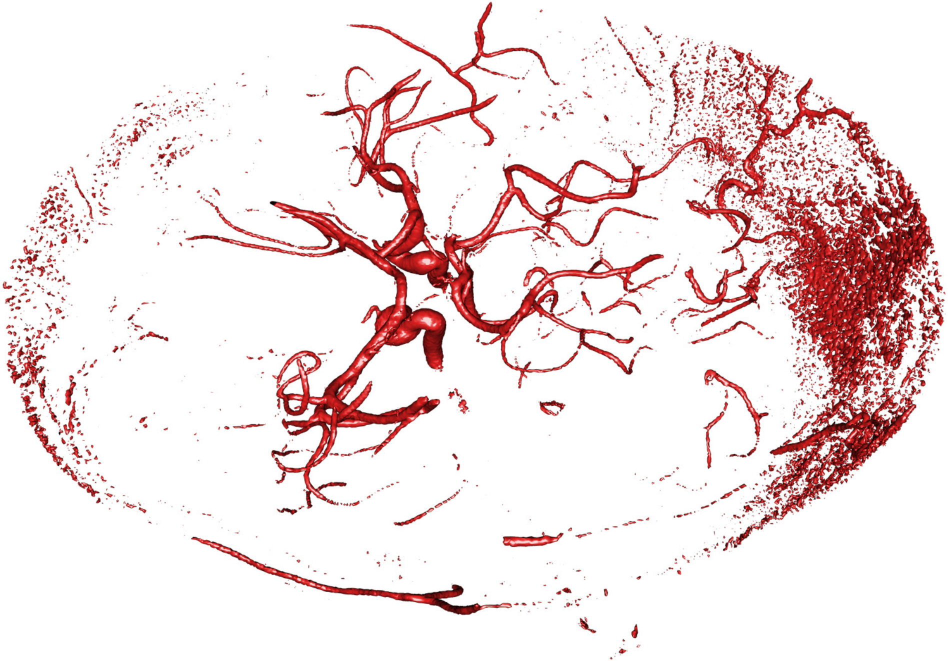
Given a pair $L \subseteq K$ of simplicial complexes, find a third complex X with $L \subseteq X \subseteq K$ such that $H_*(X) \cong \operatorname{im} H_*(L \hookrightarrow K)$.

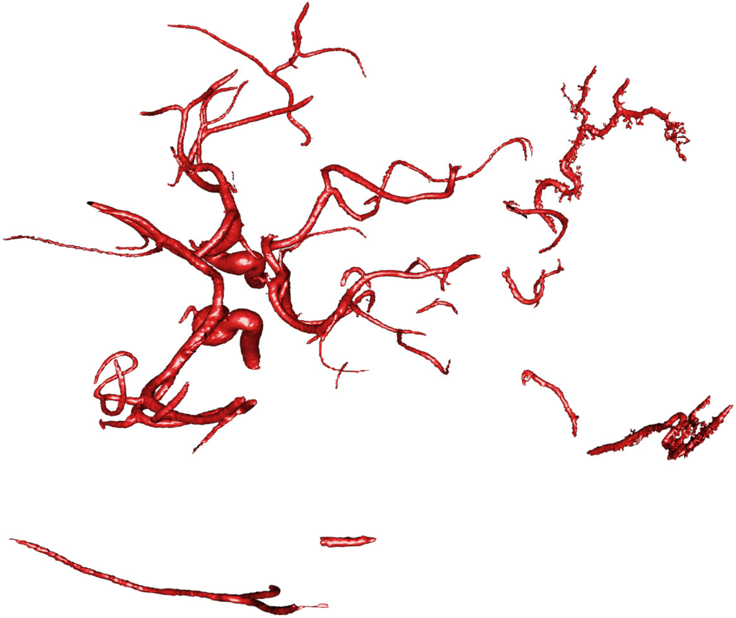


Theorem (Attali, B, Devillers, Glisse, Lieutier 2013)

The homological realization problem is NP-hard, even in $\mathbb{R}^3$.

Simplification





Sublevel set simplification

Let $F_t = f^{-1}(-\infty, t]$ denote the t-sublevel set of f.

Problem (Sublevel set simplification)

Given a function $f : X \subset \mathbb{R}^3 \to \mathbb{R}$ and $t \in \mathbb{R}$, $\delta > 0$,
find a function g with $\|g - f\|_\infty \leq \delta$ minimizing $\dim H_*(G_t)$.

Sublevel set simplification

Let $F_t = f^{-1}(-\infty, t]$ denote the t-sublevel set of f.

Problem (Sublevel set simplification)

Given a function $f : X \subset \mathbb{R}^3 \to \mathbb{R}$ *and* $t \in \mathbb{R}$, $\delta > 0$,
find a function g *with* $\|g - f\|_\infty \leq \delta$ *minimizing* $\dim H_*(G_t)$.

- Since $F_{t-\delta} \subseteq G_t \subseteq F_{t+\delta}$, the map $H_*(F_{t-\delta} \hookrightarrow F_{t+\delta})$ factors through $H_*(G_t)$.
- This gives a lower bound

$$\dim H_*(G_t) \geq \operatorname{rank} H_*(F_{t-\delta} \hookrightarrow F_{t+\delta}).$$

Sublevel set simplification

Let $F_t = f^{-1}(-\infty, t]$ denote the t-sublevel set of f.

Problem (Sublevel set simplification)

Given a function $f : X \subset \mathbb{R}^3 \to \mathbb{R}$ *and* $t \in \mathbb{R}$, $\delta > 0$,
find a function g *with* $\|g - f\|_\infty \leq \delta$ *minimizing* $\dim H_*(G_t)$.

- Since $F_{t-\delta} \subseteq G_t \subseteq F_{t+\delta}$, the map $H_*(F_{t-\delta} \hookrightarrow F_{t+\delta})$ factors through $H_*(G_t)$.
- This gives a lower bound

$$\dim H_*(G_t) \geq \operatorname{rank} H_*(F_{t-\delta} \hookrightarrow F_{t+\delta}).$$

- But it is NP-hard to decide whether this bound can be achieved.

Sublevel set simplification

Let $F_t = f^{-1}(-\infty, t]$ denote the t-sublevel set of f.

Problem (Sublevel set simplification)

Given a function $f : X \subset \mathbb{R}^3 \to \mathbb{R}$ and $t \in \mathbb{R}$, $\delta > 0$,
find a function g with $\|g - f\|_\infty \leq \delta$ minimizing $\dim H_*(G_t)$.

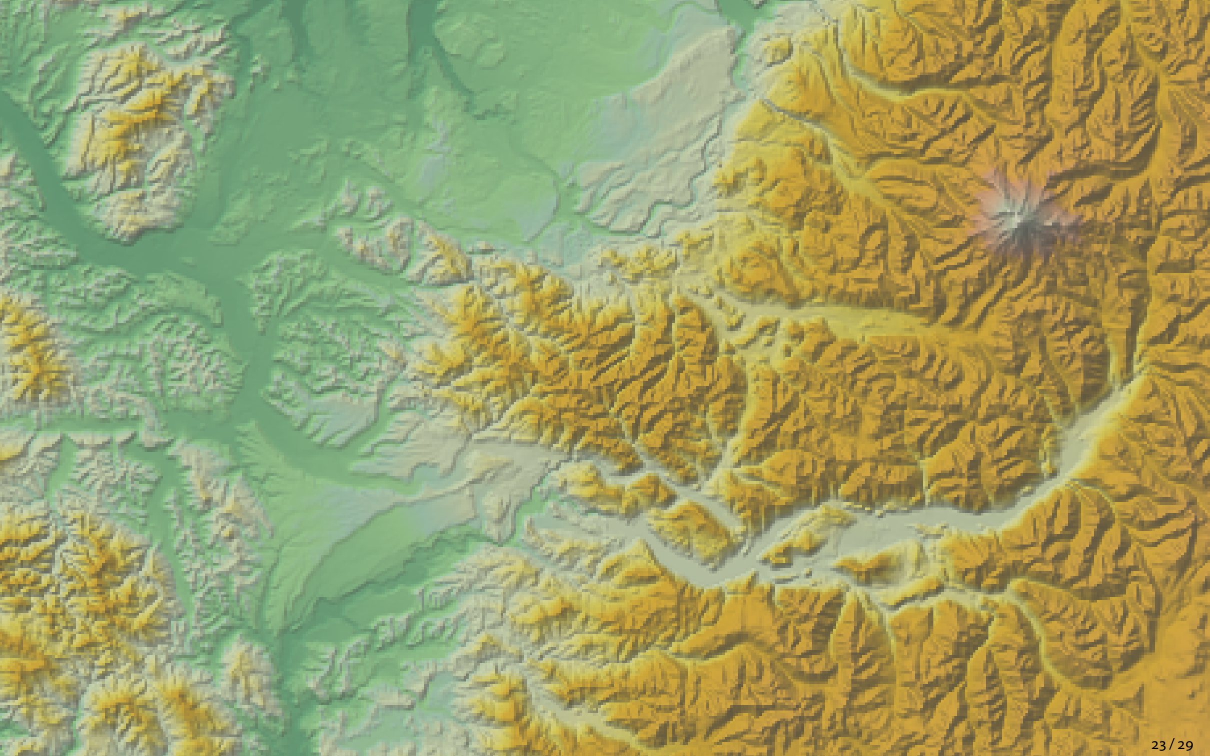
- Since $F_{t-\delta} \subseteq G_t \subseteq F_{t+\delta}$, the map $H_*(F_{t-\delta} \hookrightarrow F_{t+\delta})$ factors through $H_*(G_t)$.
- This gives a lower bound

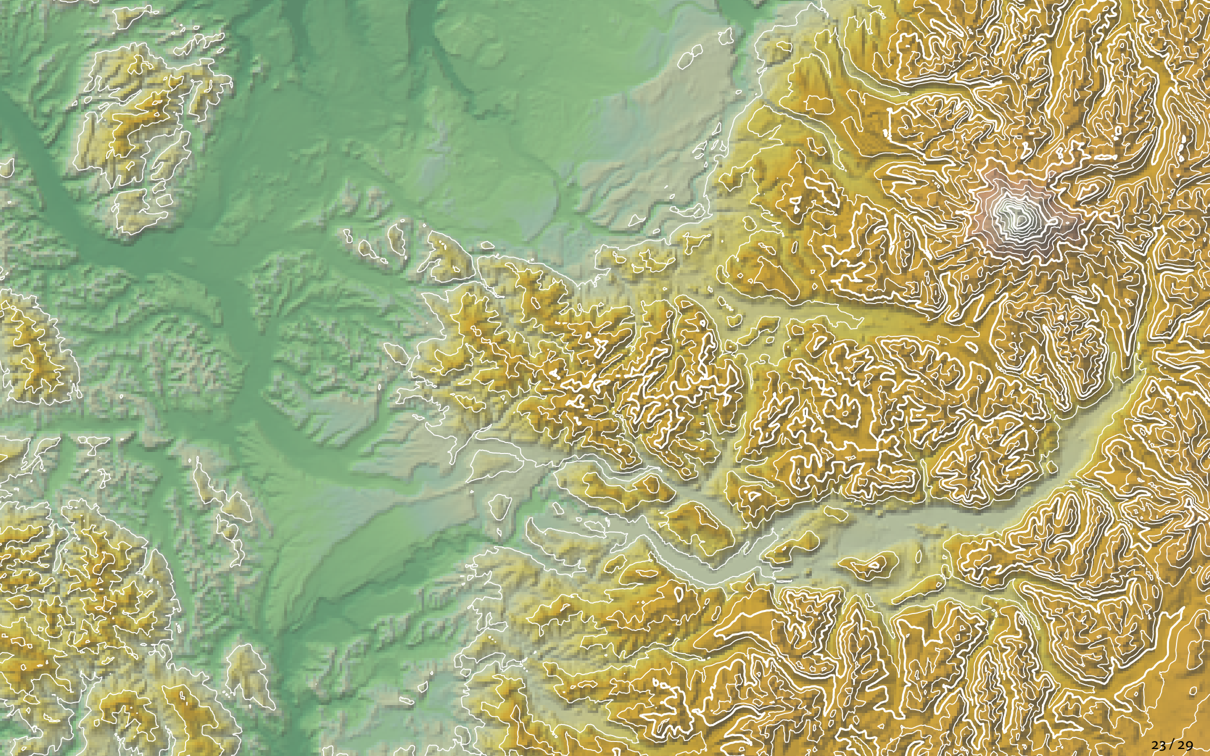
$$\dim H_*(G_t) \geq \operatorname{rank} H_*(F_{t-\delta} \hookrightarrow F_{t+\delta}).$$

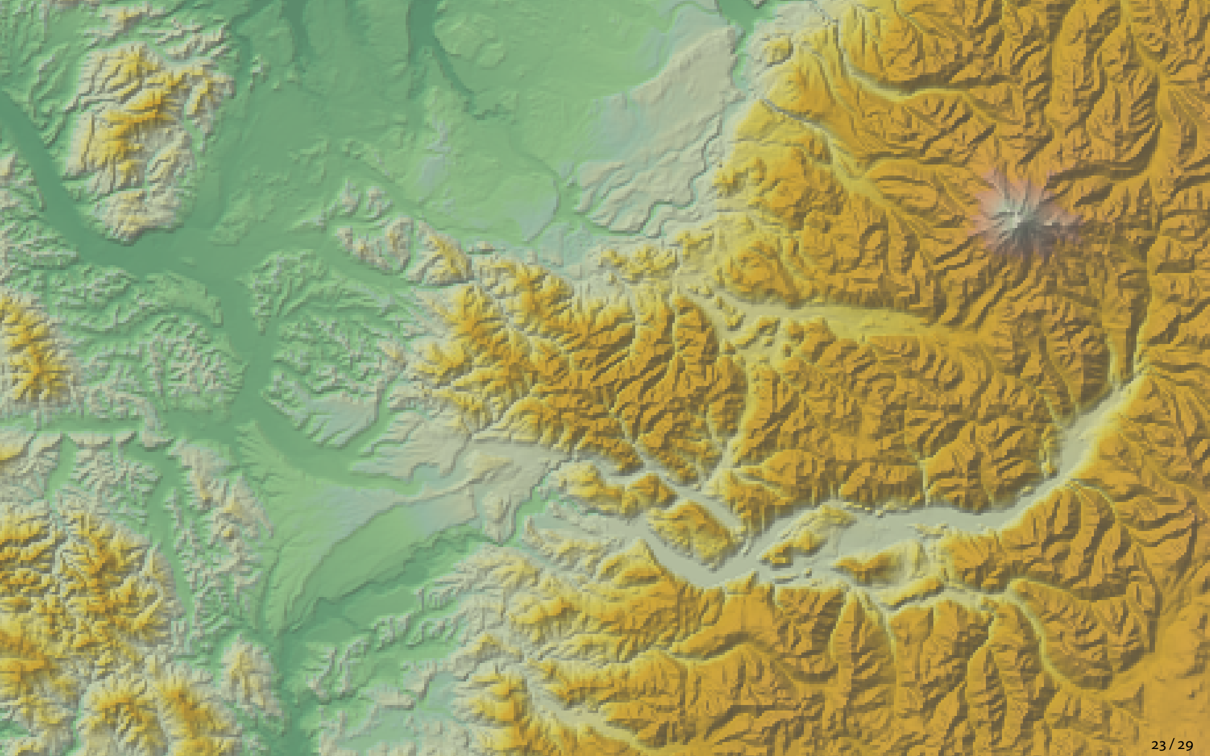
- But it is NP-hard to decide whether this bound can be achieved.

Theorem (Attali, B, Devillers, Glisse, Lieutier 2013)

Sublevel set simplification in $\mathbb{R}^3$ is NP-hard.







Topological simplification of functions

Consider the following problem:

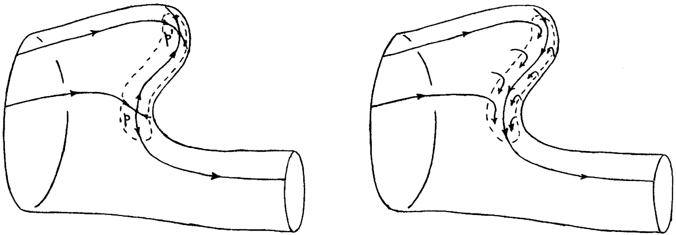
Problem (Topological simplification)

Given a function f and a real number $\delta \geq 0$, find a function f_δ subject to $\|f_\delta - f\|_\infty \leq \delta$ with the minimal number of critical points.

Persistence vs Morse theory

Morse theory (smooth or discrete):

- Relates critical points to homology of sublevel sets
- Provides a method for *canceling* pairs of critical points

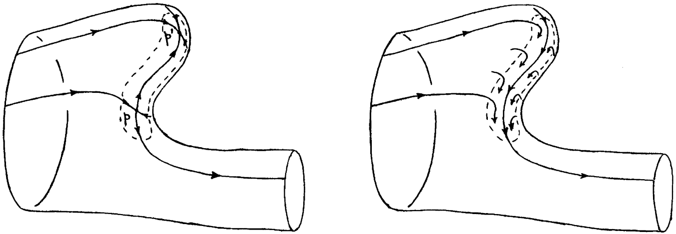


(from Milnor: *Lectures on the h-cobordism theorem*, 1965)

Persistence vs Morse theory

Morse theory (smooth or discrete):

- Relates critical points to homology of sublevel sets
- Provides a method for *canceling* pairs of critical points



(from Milnor: *Lectures on the h-cobordism theorem*, 1965)

Persistent homology:

- Relates homology of different sublevel set
- Identifies pairs of critical points (birth and death of homological feature) and quantifies their *persistence*

Combining persistence and discrete Morse theory

By stability of persistence barcodes:

Proposition

The intervals in the barcode of f with persistence $> 2\delta$ provide a lower bound on the number of critical points of any function g with $\|g - f\|_\infty \leq \delta$.

Combining persistence and discrete Morse theory

By stability of persistence barcodes:

Proposition

The intervals in the barcode of f with persistence $> 2\delta$ provide a lower bound on the number of critical points of any function g with $\|g - f\|_\infty \leq \delta$.

Theorem (B, Lange, Wardetzky, 2011)

Let f be a function on a surface and let $\delta > 0$.
Then canceling all persistence pairs with persistence $\leq 2\delta$ yields a function f_δ with

$$\|f_\delta - f\|_\infty \leq \delta$$

that achieves the lower bound on the number of critical points.

Combining persistence and discrete Morse theory

By stability of persistence barcodes:

Proposition

The intervals in the barcode of f *with persistence* $> 2\delta$ *provide a lower bound on the number of critical points of any function* g *with* $\|g - f\|_\infty \leq \delta$.

Theorem (B, Lange, Wardetzky, 2011)

Let f *be a function on a surface and let* $\delta > 0$.
Then canceling all persistence pairs with persistence $\leq 2\delta$ *yields a function* f_δ *with*

$$\|f_\delta - f\|_\infty \leq \delta$$

that achieves the lower bound on the number of critical points.

- Does not generalize to higher-dimensional manifolds!

Computation

Vietoris–Rips complexes

Consider a finite metric space (X, d).
The *Vietoris–Rips complex* is the simplicial complex

$$\mathrm{Rips}_t(X) = \{S \subseteq X \mid \operatorname{diam} S \leq t\}$$

- all edges with pairwise distance $\leq t$
- all possible higher simplices

Vietoris–Rips complexes

Consider a finite metric space (X, d).
The *Vietoris–Rips complex* is the simplicial complex

$$\mathrm{Rips}_t(X) = \{S \subseteq X \mid \operatorname{diam} S \leq t\}$$

- all edges with pairwise distance $\leq t$
- all possible higher simplices

For large t, $\mathrm{Rips}_t(X)$ is the full simplex with $n = X$ vertices

- Number of d-simplices is $\binom{n}{d+1}$
- Computation is one of the most important challenges in applied topology!

An example computation

Example data set:

- 192 points on $\mathbb{S}^2$
- persistent homology barcodes up to dimension 2
- over 56 mio. simplices in 3-skeleton

An example computation

Example data set:

- 192 points on $\mathbb{S}^2$
- persistent homology barcodes up to dimension 2
- over 56 mio. simplices in 3-skeleton

Some previous software:

- javaplex (Stanford): 3200 seconds, 12 GB
- Dionysus (Duke): 596 seconds, 3.4 GB
- DIPHA (IST Austria): 53 seconds, 5.5 GB
- GUDHI (INRIA): 46 seconds, 2.6 GB

An example computation

Example data set:

- 192 points on $\mathbb{S}^2$
- persistent homology barcodes up to dimension 2
- over 56 mio. simplices in 3-skeleton

Some previous software:

- javaplex (Stanford): 3200 seconds, 12 GB
- Dionysus (Duke): 596 seconds, 3.4 GB
- DIPHA (IST Austria): 53 seconds, 5.5 GB
- GUDHI (INRIA): 46 seconds, 2.6 GB

Demo: live.ripser.org

An example computation

Example data set:

- 192 points on $\mathbb{S}^2$
- persistent homology barcodes up to dimension 2
- over 56 mio. simplices in 3-skeleton

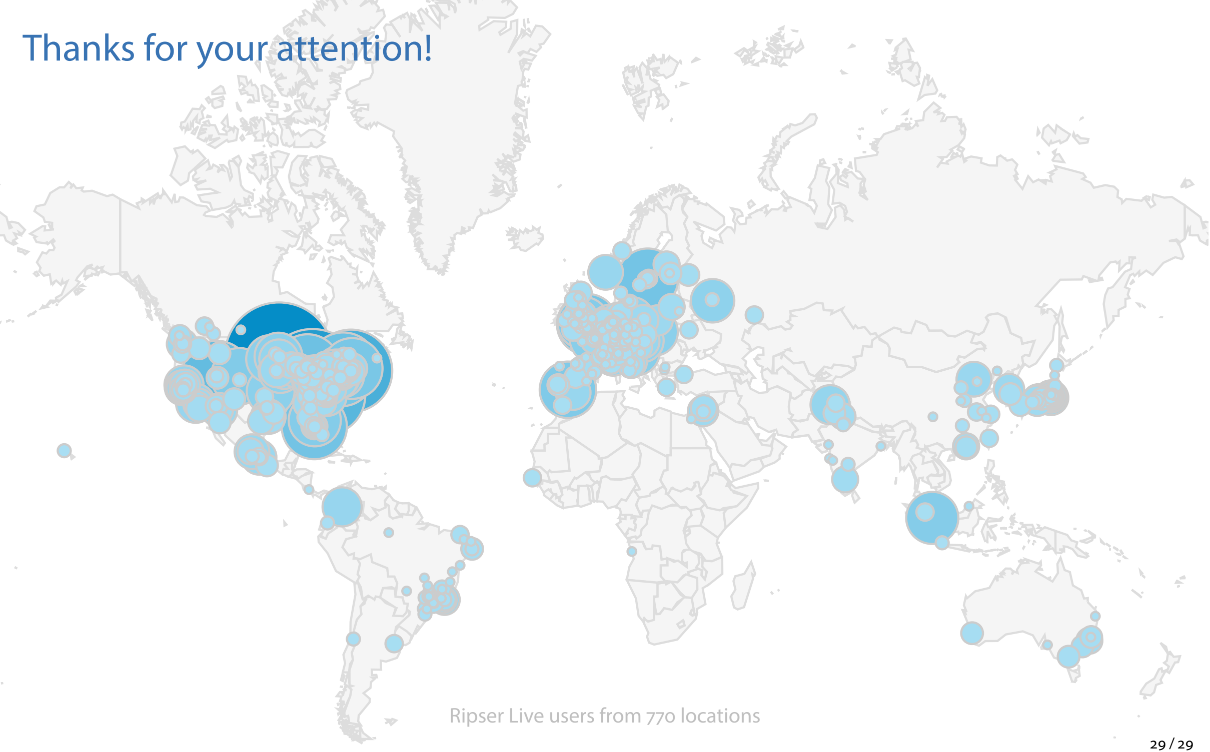
Some previous software:

- javaplex (Stanford): 3200 seconds, 12 GB
- Dionysus (Duke): 596 seconds, 3.4 GB
- DIPHA (IST Austria): 53 seconds, 5.5 GB
- GUDHI (INRIA): 46 seconds, 2.6 GB

Demo: live.ripser.org

- Ripser: 1.1 seconds, 196 MB

Thanks for your attention!



Ripser Live users from 770 locations