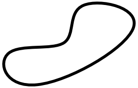
Algorithmic Aspects of Topological Data Analysis

Clément Maria

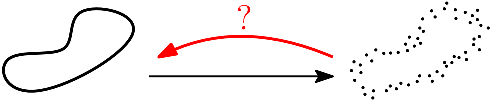
The University of Queensland, c.maria@uq.edu.au

Collège de France
Geometry Understanding in Higher Dimensions



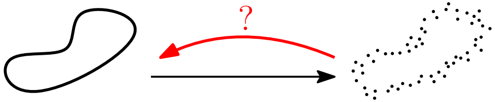
Topological Data Analysis

Topological Data Analysis



[Persistent homology '00]

Topological Data Analysis



[Persistent homology ’00]

1. Geometry

- Cech/Vietoris-Rips complexes,
- sparsification,
- zigzag persistence.

2. Combinatorics

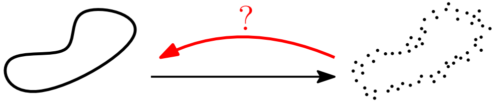
- discrete Morse theory,
- simplicial complex DS.

3. Algebra

- matrix optimisation,
- algebraic dualisation.

Focus on the topology inference problem.
$\longrightarrow$ using *persistent homology*.

Topological Data Analysis



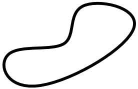
[Persistent homology '00]

1. Geometry
   - Cech/Vietoris-Rips complexes,
   - sparsification,
   - zigzag persistence.
2. Combinatorics
   - discrete Morse theory,
   - simplicial complex DS.
3. Algebra
   - matrix optimisation,
   - algebraic dualisation.

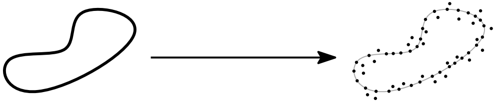
Focus on the topology inference problem.
$\longrightarrow$ using *persistent homology*.

Three (essentially) geometric constructions for persistent homology, with strong algorithmic consequences.

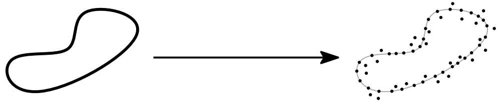
I/. Standard Persistent Homology



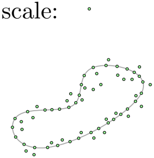
Persistent Homology and Topological Data Analysis



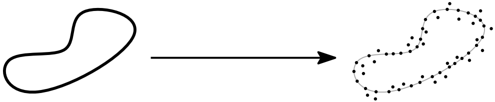
Persistent Homology and Topological Data Analysis



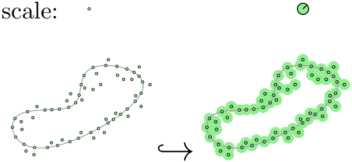
scale:



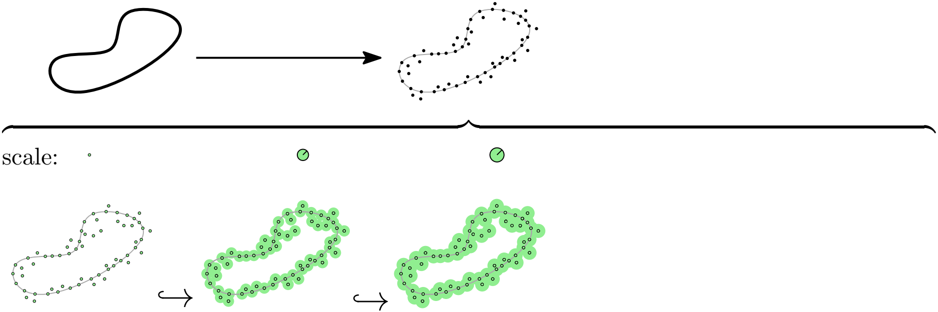
Persistent Homology and Topological Data Analysis



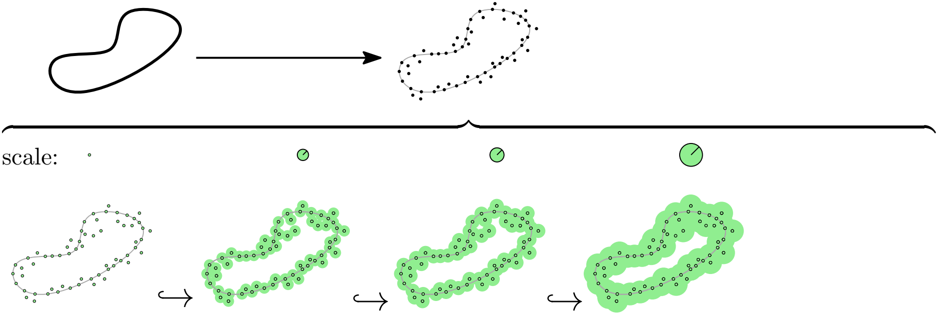
scale:



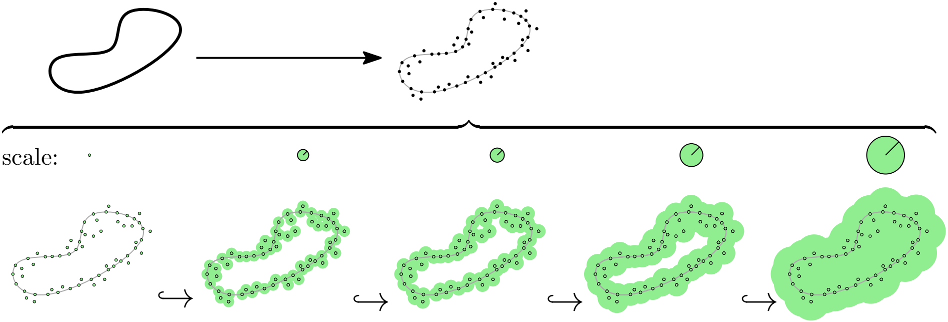
Persistent Homology and Topological Data Analysis



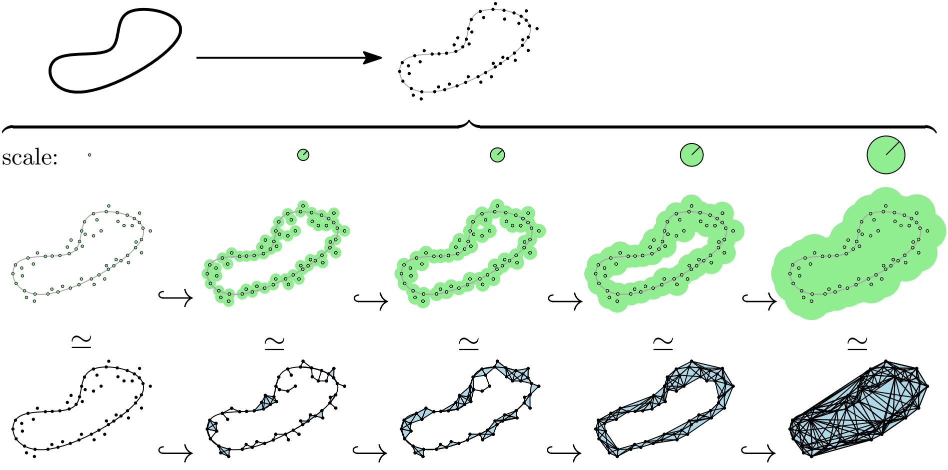
Persistent Homology and Topological Data Analysis



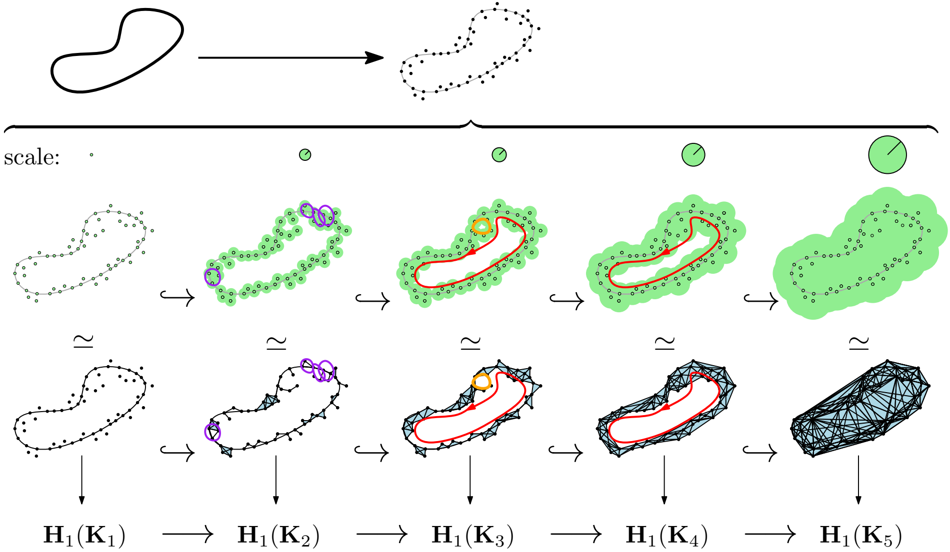
Persistent Homology and Topological Data Analysis



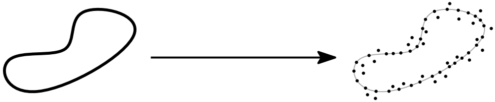
Persistent Homology and Topological Data Analysis



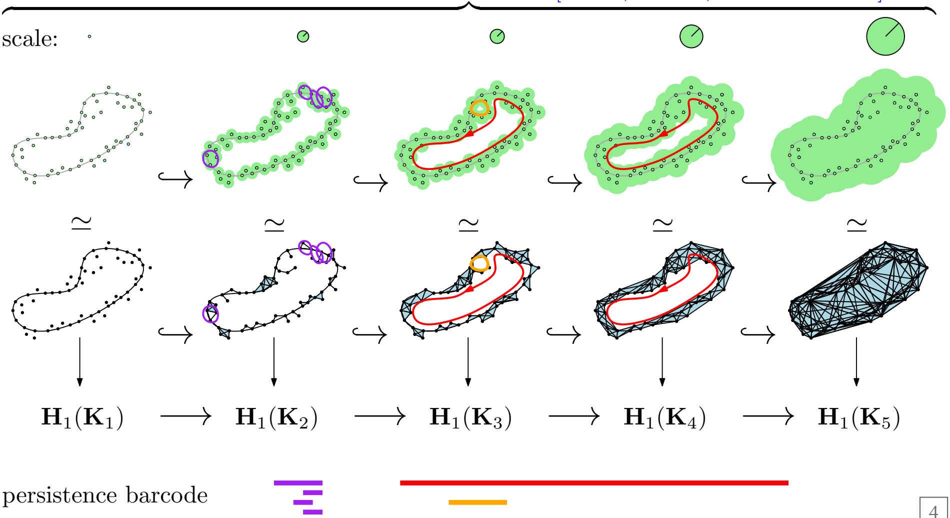
Persistent Homology and Topological Data Analysis



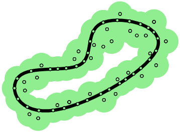
Persistent Homology and Topological Data Analysis



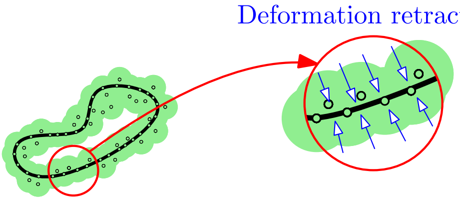
[Edelsbrunner, Letscher, Zomorodian '00]
[Carlsson, Zomorodian '04]
[Niyogi, Smale, Weinberger '08]
[Chazal, Oudot '08]
[Chazal, Lieutier '05-'09]
[Chazal, Lieutier, Cohen-Steiner '09] ...



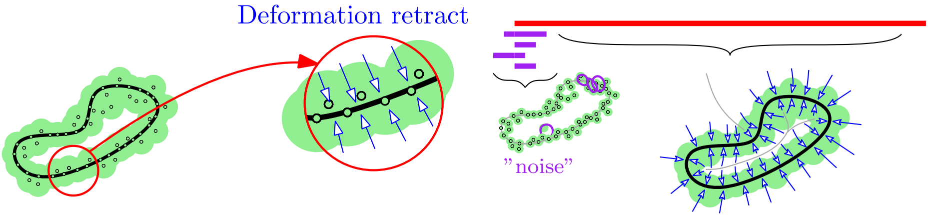
Topology Inference with Persistence



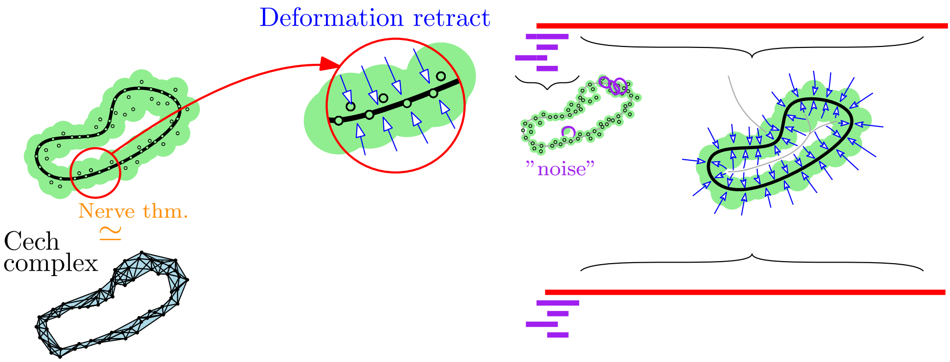
Topology Inference with Persistence



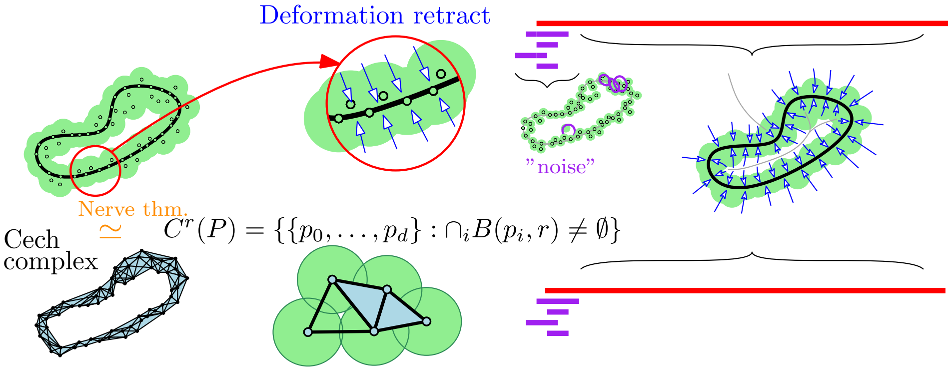
Topology Inference with Persistence



Topology Inference with Persistence



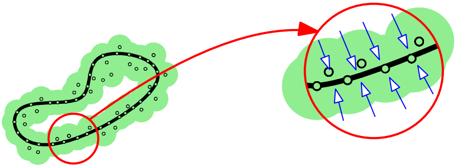
Topology Inference with Persistence

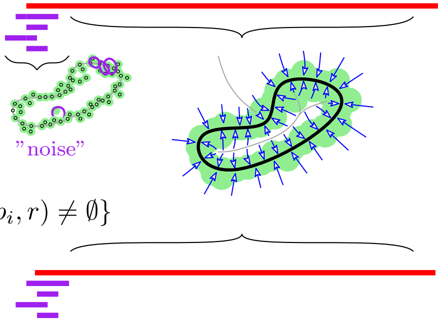


$$C^r(P) = \{\{p_0, \ldots, p_d\} : \cap_i B(p_i, r) \neq \emptyset\}$$

Topology Inference with Persistence

Deformation retract



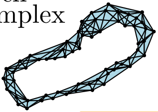


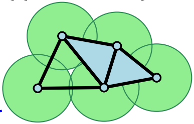
Nerve thm.

$\simeq$

$$C^r(P) = \{\{p_0, \ldots, p_d\} : \cap_i B(p_i, r) \neq \emptyset\}$$

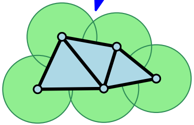
Cech complex





Relaxation

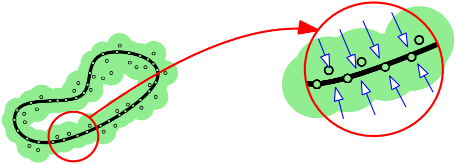
Rips complex

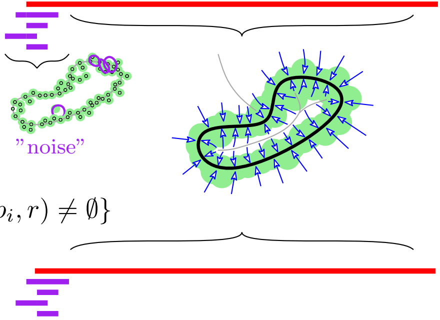


Only distance computations.

Topology Inference with Persistence

Deformation retract



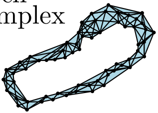


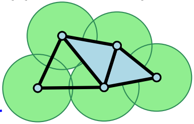
Nerve thm.

$\simeq$

Cech complex

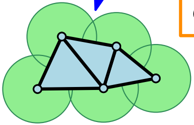
$C^r(P) = \{\{p_0, \ldots, p_d\} : \cap_i B(p_i, r) \neq \emptyset\}$





Relaxation

Rips complex

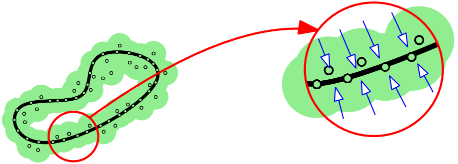


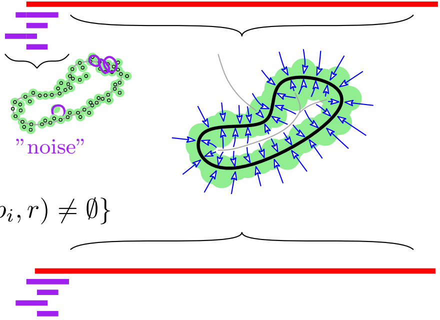
$C^r(P) \subseteq R^r(P)$

Only distance computations.

Topology Inference with Persistence

Deformation retract



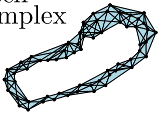


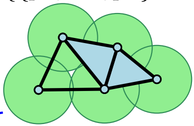
Nerve thm.

$\simeq$

Cech complex

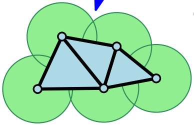
$$C^r(P) = \{\{p_0, \ldots, p_d\} : \cap_i B(p_i, r) \neq \emptyset\}$$





Relaxation

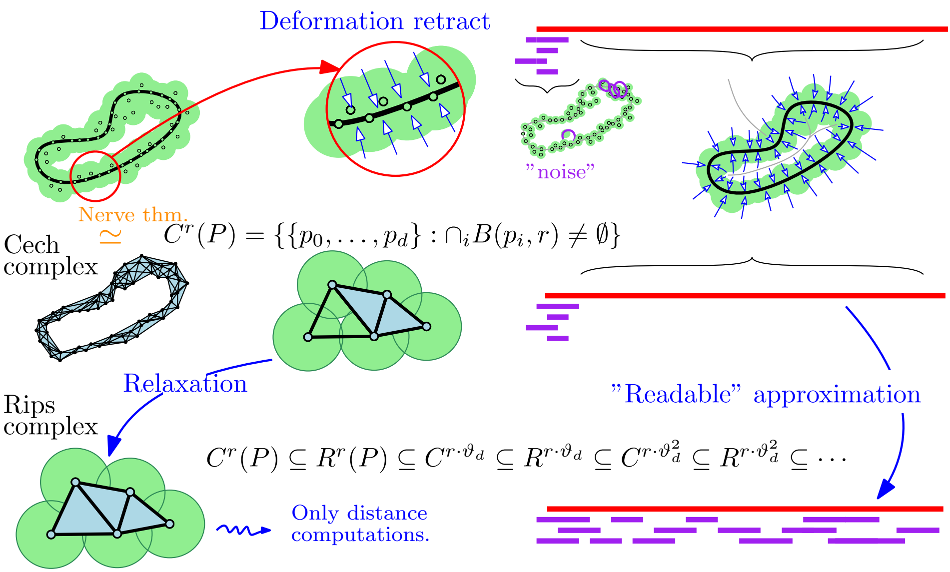
Rips complex



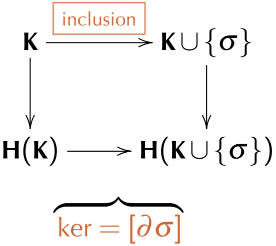
$$C^r(P) \subseteq R^r(P) \subseteq C^{r\cdot\vartheta_d} \subseteq R^{r\cdot\vartheta_d} \subseteq C^{r\cdot\vartheta_d^2} \subseteq R^{r\cdot\vartheta_d^2} \subseteq \cdots$$

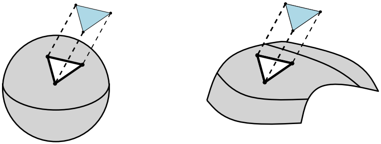
Only distance computations.

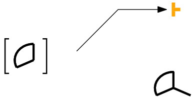
Topology Inference with Persistence



Algorithm for Persistent Homology via Inclusions



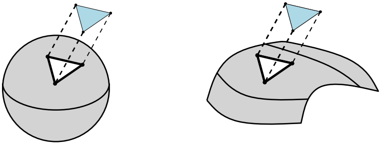


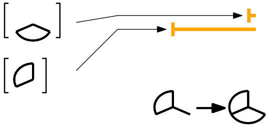


Algorithm for Persistent Homology via Inclusions

$$\begin{array}{ccc} \mathbf{K} & \xrightarrow{\text{inclusion}} & \mathbf{K} \cup \{\sigma\} \\ \downarrow & & \downarrow \\ \mathbf{H}(\mathbf{K}) & \longrightarrow & \mathbf{H}(\mathbf{K} \cup \{\sigma\}) \end{array}$$

$$\underbrace{\qquad\qquad\qquad}_{\ker = [\partial\sigma]}$$

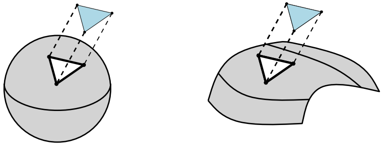


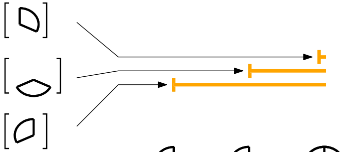


Algorithm for Persistent Homology via Inclusions

$$\begin{array}{ccc} K & \xrightarrow{\text{inclusion}} & K \cup \{\sigma\} \\ \downarrow & & \downarrow \\ H(K) & \longrightarrow & H(K \cup \{\sigma\}) \end{array}$$

$$\underbrace{\hphantom{H(K) \longrightarrow}}_{\ker = [\partial\sigma]}$$

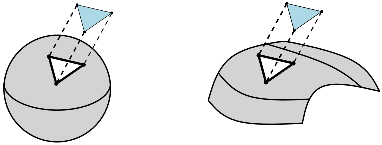


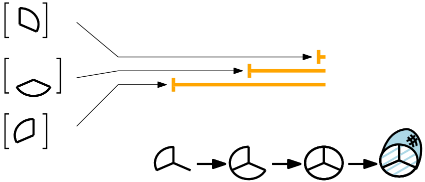


Algorithm for Persistent Homology via Inclusions

$$\begin{array}{ccc} \mathbf{K} & \xrightarrow{\text{inclusion}} & \mathbf{K} \cup \{\sigma\} \\ \downarrow & & \downarrow \\ \mathbf{H}(\mathbf{K}) & \longrightarrow & \mathbf{H}(\mathbf{K} \cup \{\sigma\}) \end{array}$$

$\text{ker} = [\partial\sigma]$

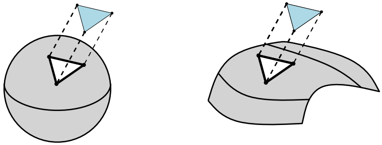


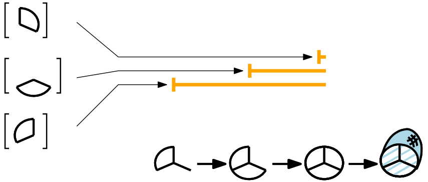


Algorithm for Persistent Homology via Inclusions

$$\begin{array}{ccc} \mathbf{K} & \xrightarrow{\text{inclusion}} & \mathbf{K}\cup\{\sigma\} \\ \downarrow & & \downarrow \\ \mathbf{H}(\mathbf{K}) & \longrightarrow & \mathbf{H}(\mathbf{K}\cup\{\sigma\}) \end{array}$$

$\ker = [\partial\sigma]$

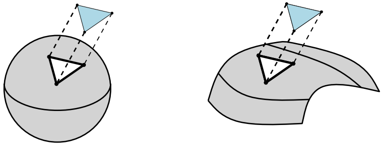


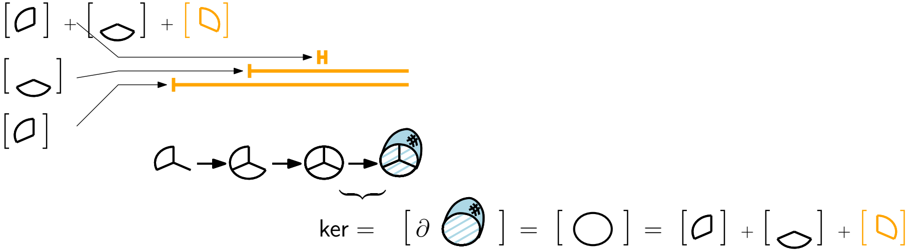


Algorithm for Persistent Homology via Inclusions

$$\begin{array}{ccc} \mathbf{K} & \xrightarrow{\text{inclusion}} & \mathbf{K} \cup \{\sigma\} \\ \downarrow & & \downarrow \\ \mathbf{H}(\mathbf{K}) & \longrightarrow & \mathbf{H}(\mathbf{K} \cup \{\sigma\}) \end{array}$$

$\text{ker} = [\partial \sigma]$

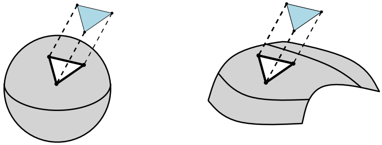




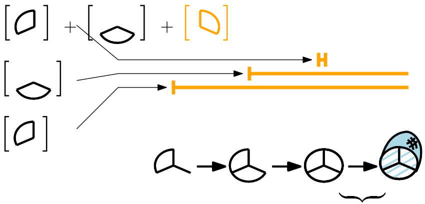
Algorithm for Persistent Homology via Inclusions

$$\begin{array}{ccc} \mathbf{K} & \xrightarrow{\text{inclusion}} & \mathbf{K} \cup \{\sigma\} \\ \downarrow & & \downarrow \\ \mathbf{H}(\mathbf{K}) & \longrightarrow & \mathbf{H}(\mathbf{K} \cup \{\sigma\}) \end{array}$$

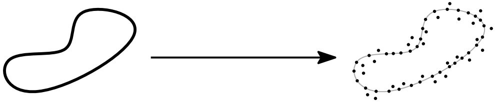
$\underbrace{\hphantom{\mathbf{H}(\mathbf{K}) \longrightarrow}}_{\ker = [\partial\sigma]}$



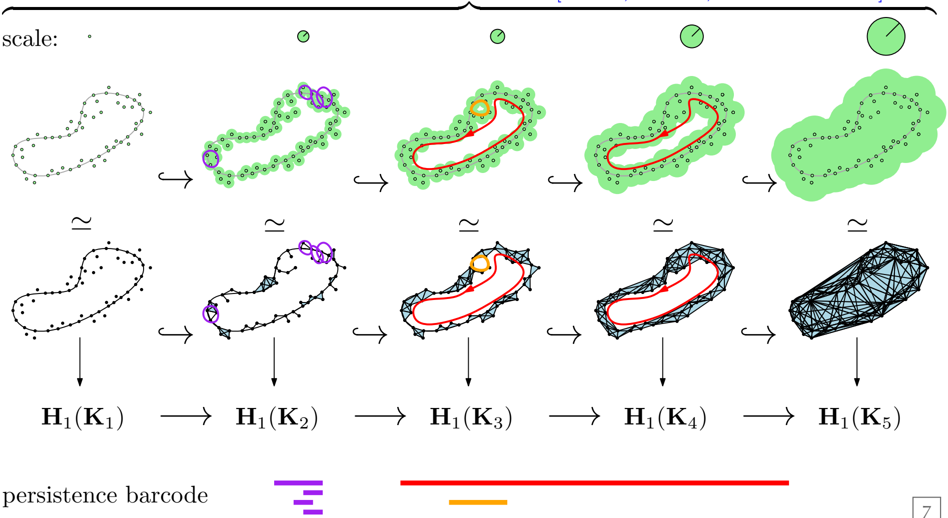
Only one bar modified per inclusion.



Persistent Homology and Topological Data Analysis



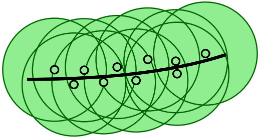
[Edelsbrunner, Letscher, Zomorodian '00]
[Carlsson, Zomorodian '04]
[Niyogi, Smale, Weinberger '08]
[Chazal, Oudot '08]
[Chazal, Lieutier '05-'09]
[Chazal, Lieutier, Cohen-Steiner '09] ...



II/. Sparsification in Persistence

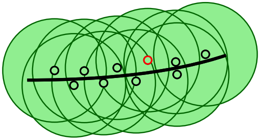
Sparsification

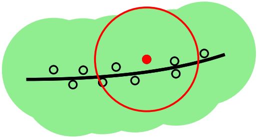
[Sheehy '12]



Sparsification

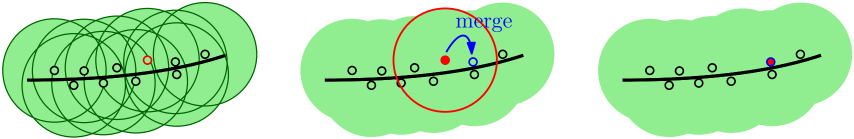
[Sheehy '12]





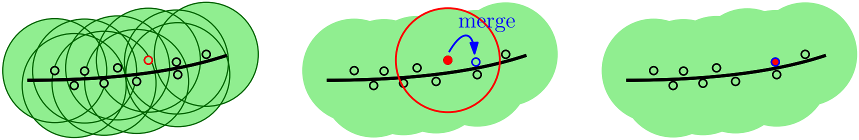
Sparsification

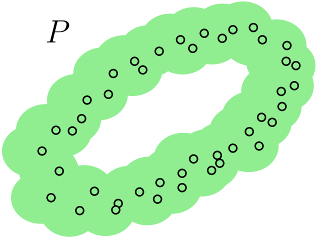
[Sheehy '12]



Sparsification

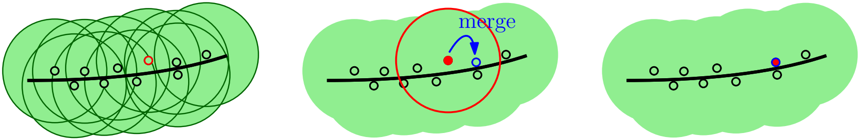
[Sheehy '12]

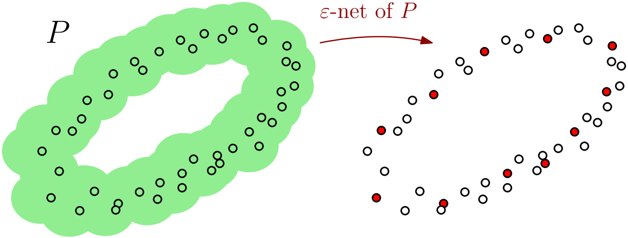




Sparsification

[Sheehy '12]

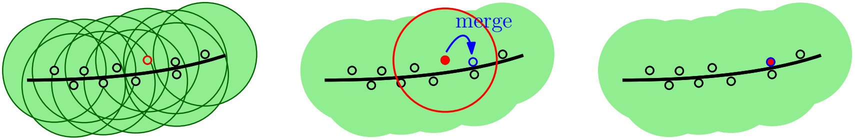


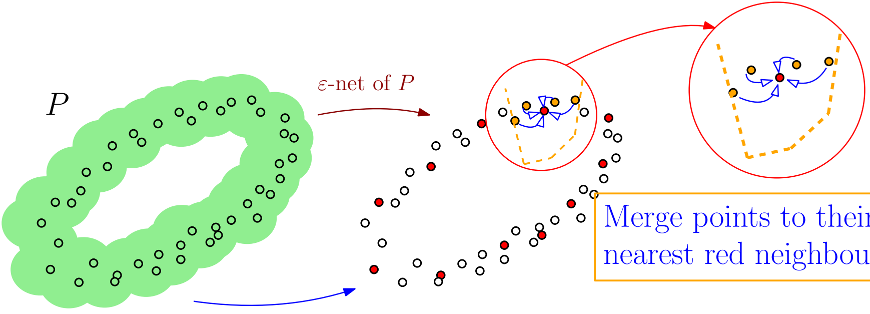


Select an ε-net of P

Sparsification

[Sheehy '12]



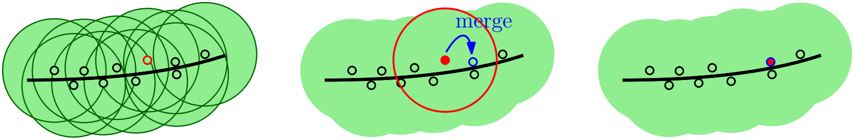


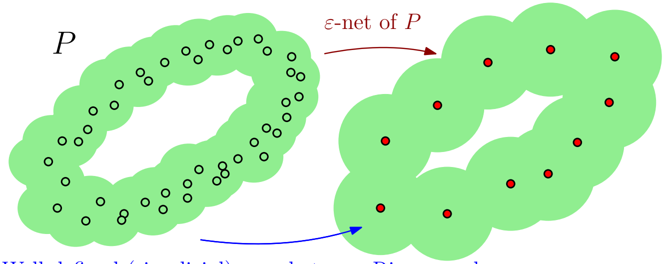
Merge points to their nearest red neighbour

Well-defined (simplicial) map between Rips complexes

Sparsification

[Sheehy '12]



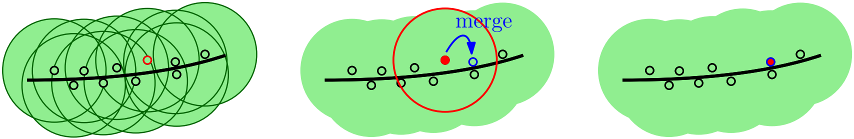


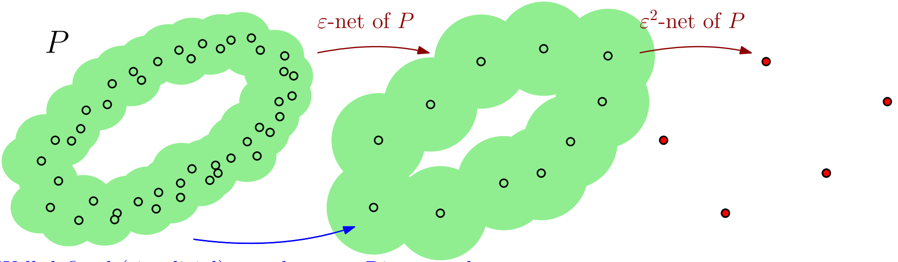
Increase scale

Well-defined (simplicial) map between Rips complexes

Sparsification

[Sheehy '12]



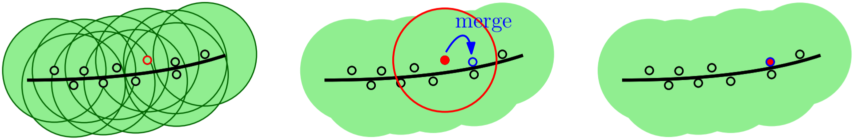


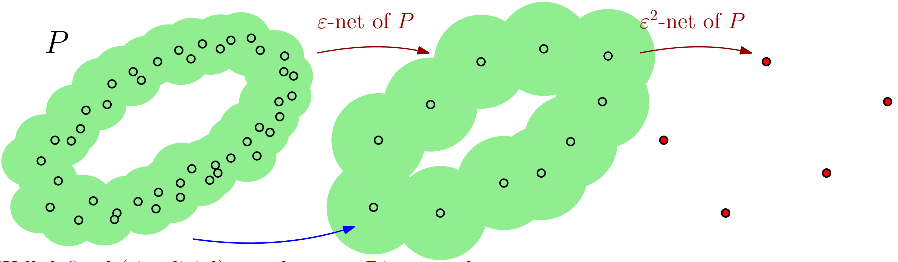
Well-defined (simplicial) map between Rips complexes

[Dey, Fan, Wang '14]

Sparsification

[Sheehy '12]





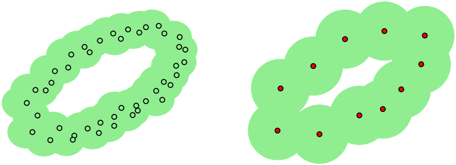
Well-defined (simplicial) map between Rips complexes

[Dey, Fan, Wang '14]

[Kerber, Sharathkumar '13] [Botnan, Spreeman '14] [Dey, Shi, Wang '16]

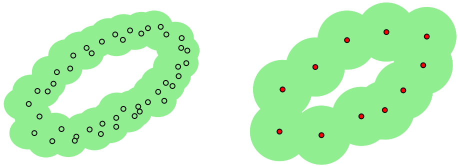
[Aruni, Kerber, Raghvendra '16] ...

Algorithm for Persistent Homology via Contractions



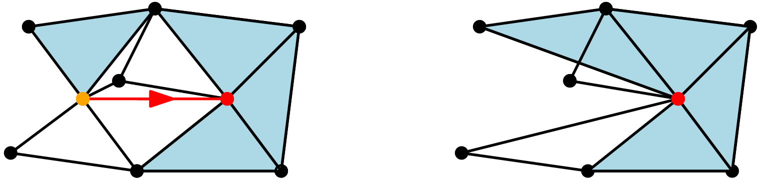
Increase scale $\longrightarrow$ inclusions $\mathbf{K} \xrightarrow{\text{inclusion}} \mathbf{K} \cup \{\sigma\}$

Algorithm for Persistent Homology via Contractions

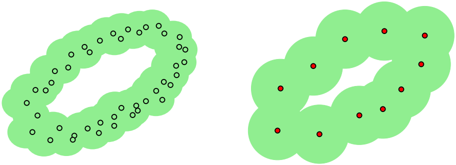


Increase scale $\longrightarrow$ inclusions $\mathbf{K} \xrightarrow{\text{inclusion}} \mathbf{K} \cup \{\sigma\}$

Merge points $\longrightarrow$ edge contraction:

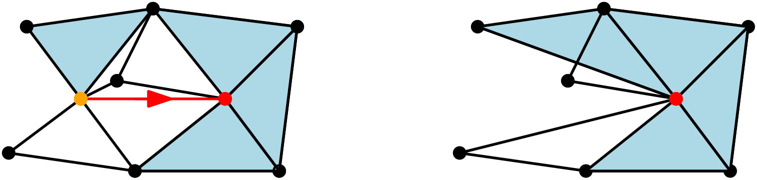


Algorithm for Persistent Homology via Contractions



Increase scale $\longrightarrow$ inclusions $\mathbf{K} \xrightarrow{\text{inclusion}} \mathbf{K} \cup \{\sigma\}$

Merge points $\longrightarrow$ edge contraction:

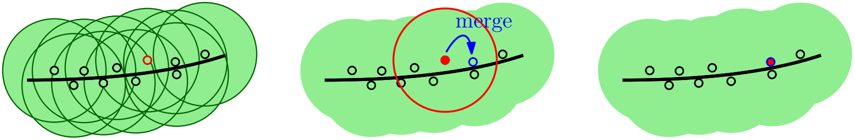


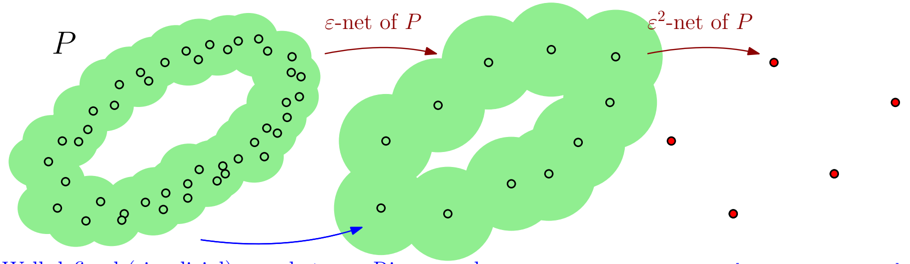
Several bars get 'destroyed' in the persistence barcode.

$\longrightarrow$ directed by the combinatorics/geometry.

Sparsification

[Sheehy '12]





Well-defined (simplicial) map between Rips complexes

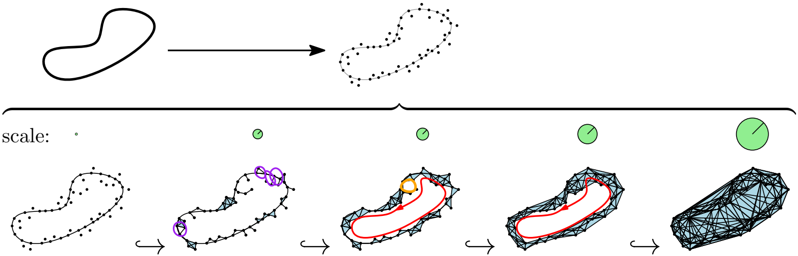
[Dey, Fan, Wang '14]

[Kerber, Sharathkumar '13] [Botnan, Spreeman '14] [Dey, Shi, Wang '16]

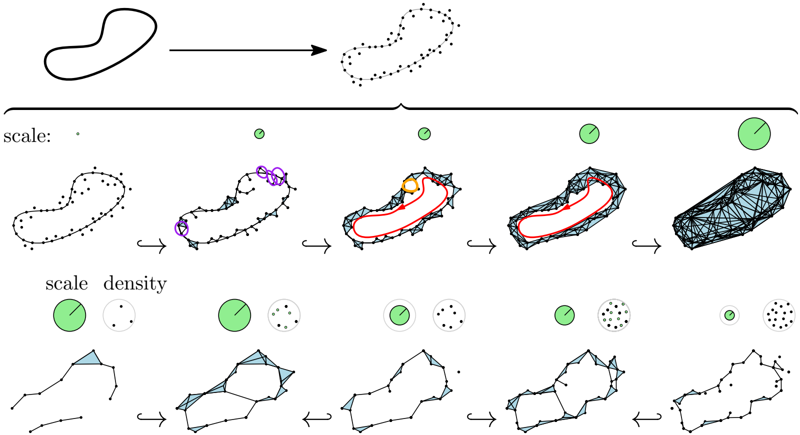
[Aruni, Kerber, Raghvendra '16] ...

III/. Zigzag Persistent Homology

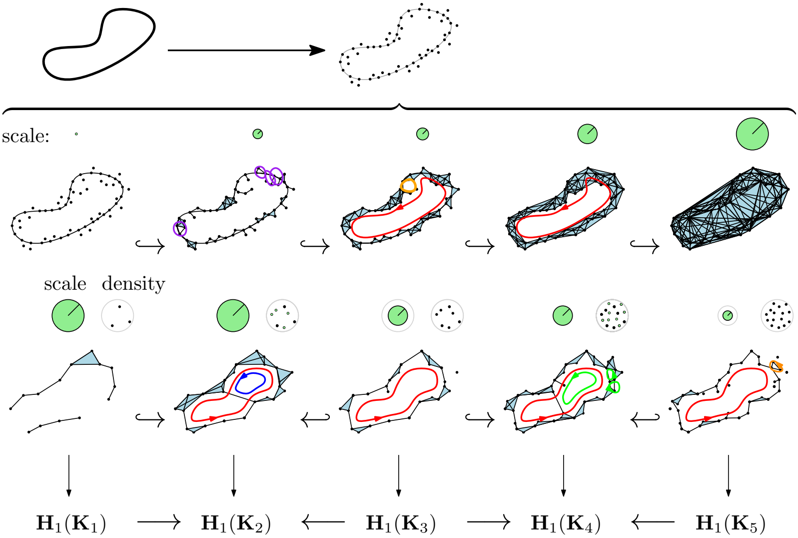
Zigzag Persistence for Topological Data Analysis



Zigzag Persistence for Topological Data Analysis

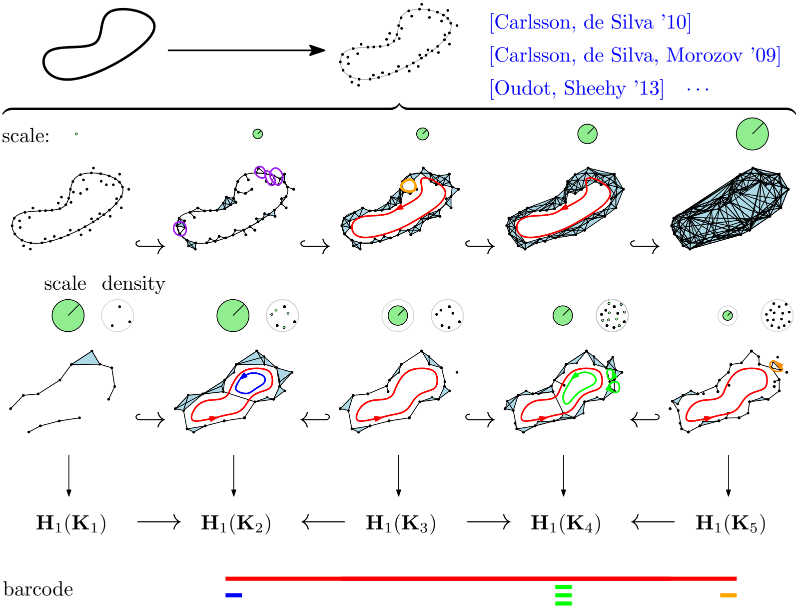


Zigzag Persistence for Topological Data Analysis



Zigzag Persistence for Topological Data Analysis

[Carlsson, de Silva '10]
[Carlsson, de Silva, Morozov '09]
[Oudot, Sheehy '13] ...



Zigzag Persistence Algorithm

We compute the persistent homology of:

$$\mathbf{K}_1 \longleftrightarrow \mathbf{K}_2 \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_i \overset{\sigma}{\longleftrightarrow} \mathbf{K}_{i+1} \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_{n-1} \longleftrightarrow \mathbf{K}_n$$

Zigzag Persistence Algorithm

We compute the persistent homology of:

$$\mathbf{K}_1 \longleftrightarrow \mathbf{K}_2 \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_i \overset{\sigma}{\longleftrightarrow} \mathbf{K}_{i+1} \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_{n-1} \longleftrightarrow \mathbf{K}_n$$

by maintaining a compatible homology basis for [M., Oudot '15 '16]

$$\underbrace{\mathbf{K}_1 \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_i = \mathbf{K}'_m}_{\mathbb{K}[1;i]} \xleftarrow{\tau_m} \mathbf{K}'_{m-1} \xleftarrow{\tau_{m-1}} \mathbf{K}'_{m-2} \xleftarrow{\tau_{m-2}} \cdots \xleftarrow{\tau_1} \emptyset$$

Zigzag Persistence Algorithm

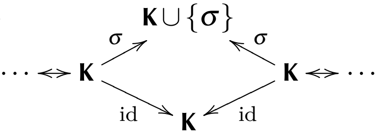
We compute the persistent homology of:

$$\mathbf{K}_1 \longleftrightarrow \mathbf{K}_2 \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_i \overset{\sigma}{\longleftrightarrow} \mathbf{K}_{i+1} \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_{n-1} \longleftrightarrow \mathbf{K}_n$$

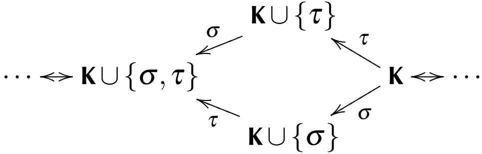
by maintaining a compatible homology basis for [M., Oudot '15 '16]

$$\underbrace{\mathbf{K}_1 \longleftrightarrow \cdots \longleftrightarrow \mathbf{K}_i = \mathbf{K}'_m}_{\mathbb{K}[1;i]} \xleftarrow{\tau_m} \mathbf{K}'_{m-1} \xleftarrow{\tau_{m-1}} \mathbf{K}'_{m-2} \xleftarrow{\tau_{m-2}} \cdots \xleftarrow{\tau_1} \emptyset$$

under

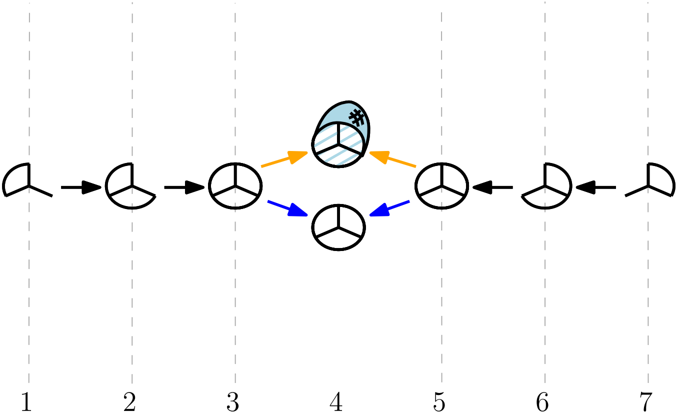


Arrow reflections if $\mathbf{K}_{\overline{k}} \xrightarrow{\sigma} \mathbf{K}_{\overline{k+1}}$ is forward.

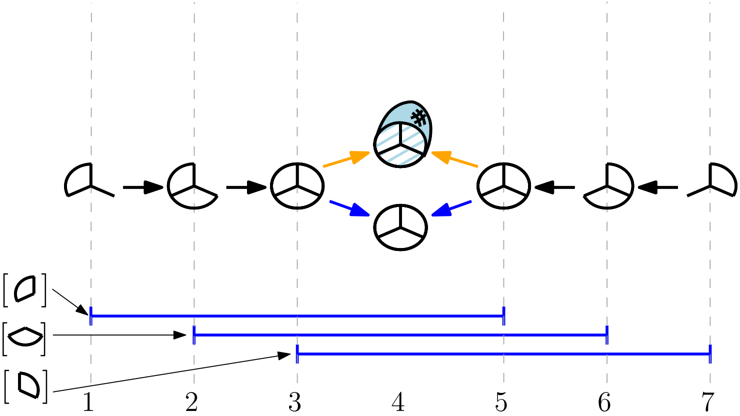


Arrow transpositions in sequence if $\mathbf{K}_{\overline{k}} \xleftarrow{\sigma} \mathbf{K}_{\overline{k+1}}$ is backward.

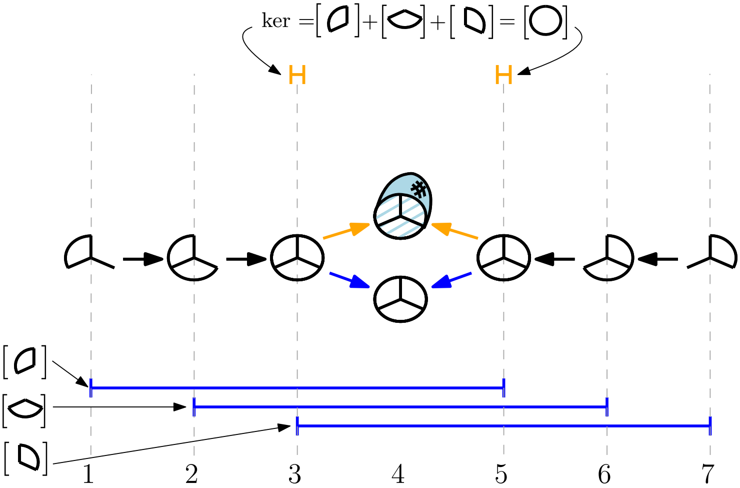
Zigzag Persistence Algorithm



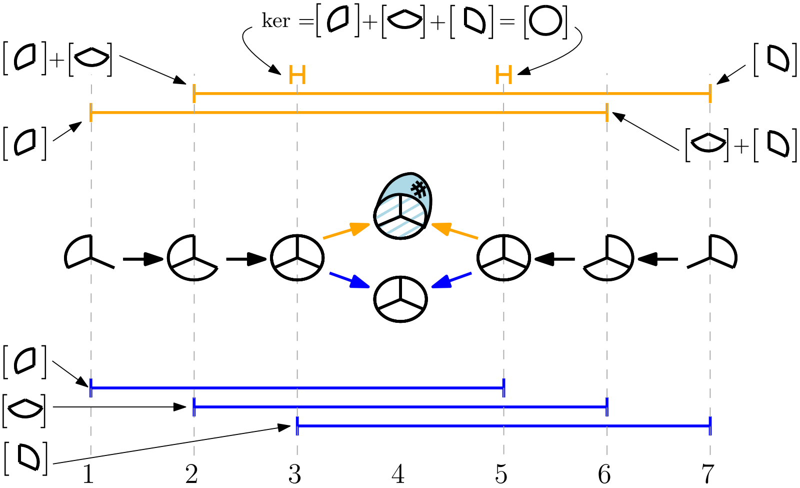
Zigzag Persistence Algorithm



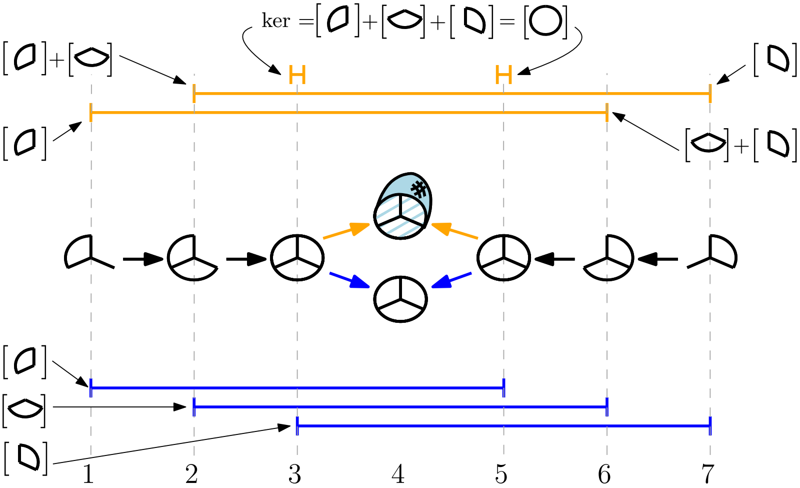
Zigzag Persistence Algorithm



Zigzag Persistence Algorithm



Zigzag Persistence Algorithm



An arbitrary number of bars change

⟶ directed by the topology/algebra.

Concluding Remarks

In Practice, a Toy Example

Standard persistence		Sparse persistence		Zigzag Persistence		
#$\mathbf{K}_{max}$	T	#$\mathbf{K}_{max}$	T	# arrows	#$\mathbf{K}_{max}$	T
$230 \cdot 10^6$	3147 sec.	7474	24 sec.	$2.4 \cdot 10^6$	50840	285 sec.

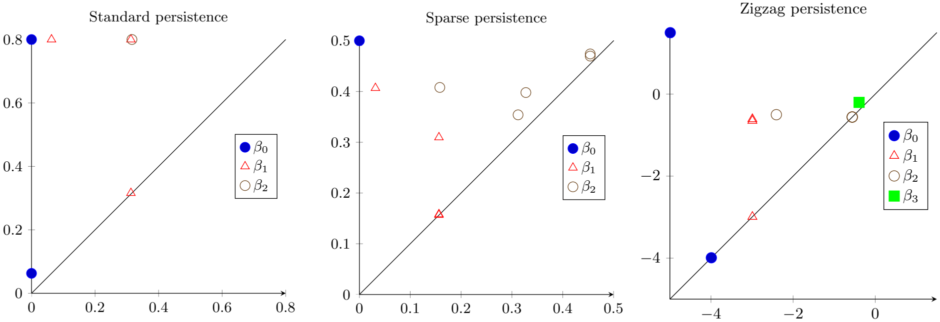


Figure: Best possible persistence diagrams obtained by the different methods on 2000 points sampling a torus wrapped around a (poorly sampled) 3-sphere in $\mathbb{R}^4$.

Conclusion & Perspectives

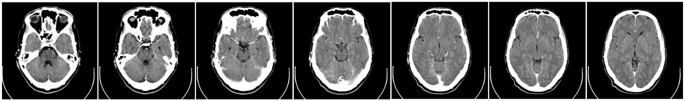
Three technologies to solve the inference problem:

$$\mathbf{K}_1 \hookrightarrow \mathbf{K}_2 \hookrightarrow \cdots \hookrightarrow \mathbf{K}_i \hookrightarrow \mathbf{K}_{i+1} \hookrightarrow \cdots \qquad \text{inclusions.}$$

$$\mathbf{K}_1 \longmapsto \mathbf{K}_2 \longmapsto \cdots \longmapsto \mathbf{K}_i \longmapsto \mathbf{K}_{i+1} \longmapsto \cdots \qquad \text{incl. \& contractions.}$$

$$\mathbf{K}_1 \longleftarrow \mathbf{K}_2 \longrightarrow \cdots \longleftarrow \mathbf{K}_i \longrightarrow \mathbf{K}_{i+1} \longleftarrow \cdots \qquad \text{incl. \& removals.}$$

Conclusion & Perspectives

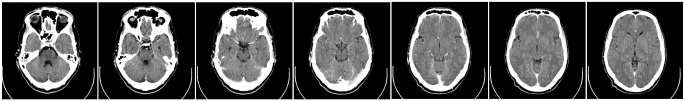


$\mathbf{K}_1 \hookrightarrow \mathbf{K}_2 \hookrightarrow \cdots \hookrightarrow \mathbf{K}_i \hookrightarrow \mathbf{K}_{i+1} \hookrightarrow \cdots$ inclusions.

$\mathbf{K}_1 \longmapsto \mathbf{K}_2 \longmapsto \cdots \longmapsto \mathbf{K}_i \longmapsto \mathbf{K}_{i+1} \longmapsto \cdots$ incl. & contractions.

$\mathbf{K}_1 \longleftarrow \mathbf{K}_2 \longrightarrow \cdots \longleftarrow \mathbf{K}_i \longrightarrow \mathbf{K}_{i+1} \longleftarrow \cdots$ incl. & removals.

Conclusion & Perspectives



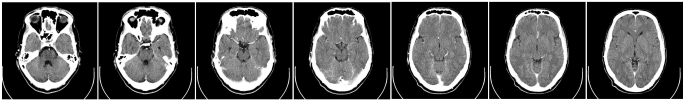
$\mathbf{K}_1 \hookrightarrow \mathbf{K}_2 \hookrightarrow \cdots \hookrightarrow \mathbf{K}_i \hookrightarrow \mathbf{K}_{i+1} \hookrightarrow \cdots$ inclusions.

 $\overset{?}{\hookrightarrow}$ 

$\mathbf{K}_1 \longmapsto \mathbf{K}_2 \longmapsto \cdots \longmapsto \mathbf{K}_i \longmapsto \mathbf{K}_{i+1} \longmapsto \cdots$ incl. & contractions.

$\mathbf{K}_1 \longleftarrow \mathbf{K}_2 \longrightarrow \cdots \longleftarrow \mathbf{K}_i \longrightarrow \mathbf{K}_{i+1} \longleftarrow \cdots$ incl. & removals.

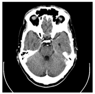
Conclusion & Perspectives



$\mathbf{K}_1 \hookrightarrow \mathbf{K}_2 \hookrightarrow \cdots \hookrightarrow \mathbf{K}_i \hookrightarrow \mathbf{K}_{i+1} \hookrightarrow \cdots$ inclusions.

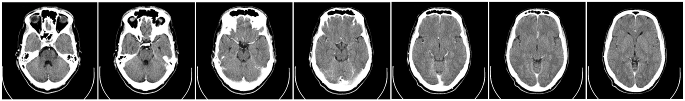
 $\overset{?}{\hookrightarrow}$ 

$\mathbf{K}_1 \longmapsto \mathbf{K}_2 \longmapsto \cdots \longmapsto \mathbf{K}_i \longmapsto \mathbf{K}_{i+1} \longmapsto \cdots$ incl. & contractions.

 $\overset{\text{deformation}}{\longmapsto}$ 

$\mathbf{K}_1 \longleftarrow \mathbf{K}_2 \longrightarrow \cdots \longleftarrow \mathbf{K}_i \longrightarrow \mathbf{K}_{i+1} \longleftarrow \cdots$ incl. & removals.

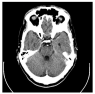
Conclusion & Perspectives



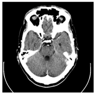
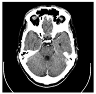
$\mathbf{K}_1 \hookrightarrow \mathbf{K}_2 \hookrightarrow \cdots \hookrightarrow \mathbf{K}_i \hookrightarrow \mathbf{K}_{i+1} \hookrightarrow \cdots$ inclusions.

 $\overset{?}{\hookrightarrow}$ 

$\mathbf{K}_1 \longmapsto \mathbf{K}_2 \longmapsto \cdots \longmapsto \mathbf{K}_i \longmapsto \mathbf{K}_{i+1} \longmapsto \cdots$ incl. & contractions.

 $\overset{\text{deformation}}{\longmapsto}$ 

$\mathbf{K}_1 \longleftarrow \mathbf{K}_2 \longrightarrow \cdots \longleftarrow \mathbf{K}_i \longrightarrow \mathbf{K}_{i+1} \longleftarrow \cdots$ incl. & removals.

 $\overset{\text{incl.}}{\longrightarrow}$  $\cup$  $\overset{\text{incl.}}{\longleftarrow}$