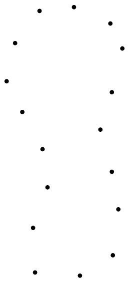
Geometric Complexes in Topological Data Analysis

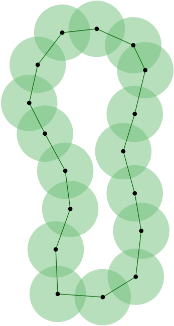
Fabian Roll

Technical University of Munich

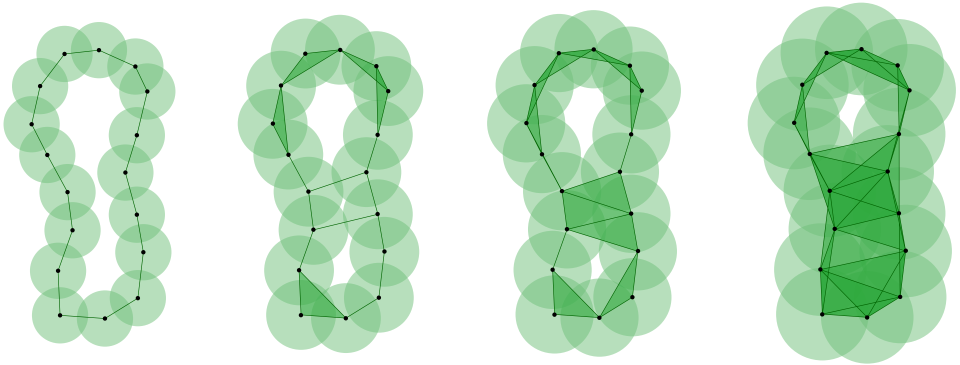
Persistent homology pipeline



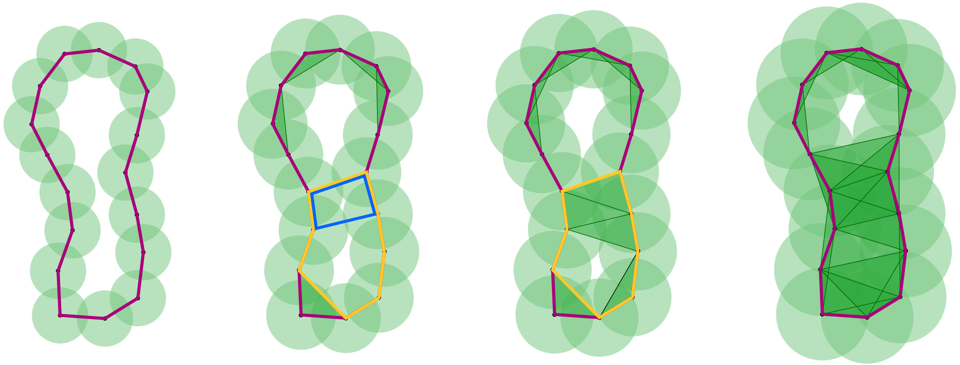
Persistent homology pipeline



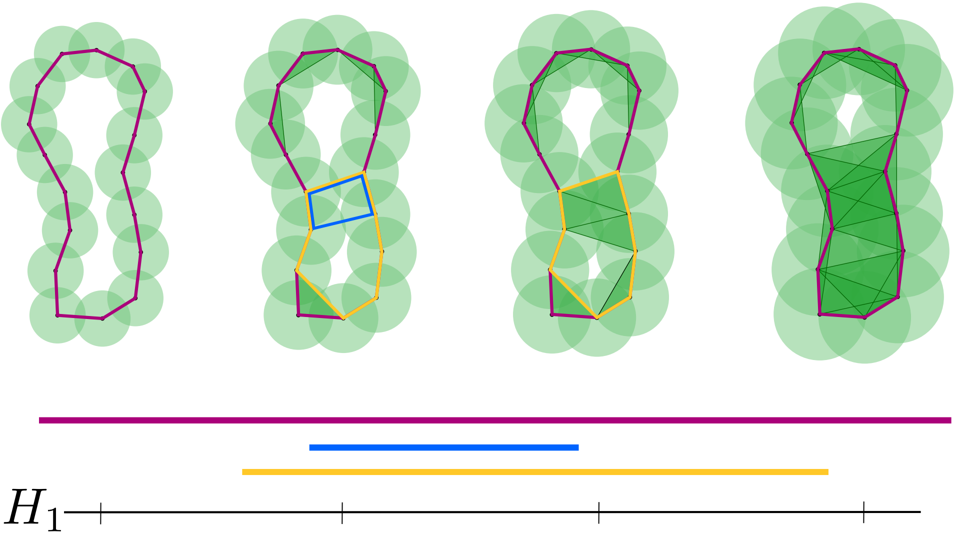
Persistent homology pipeline



Persistent homology pipeline



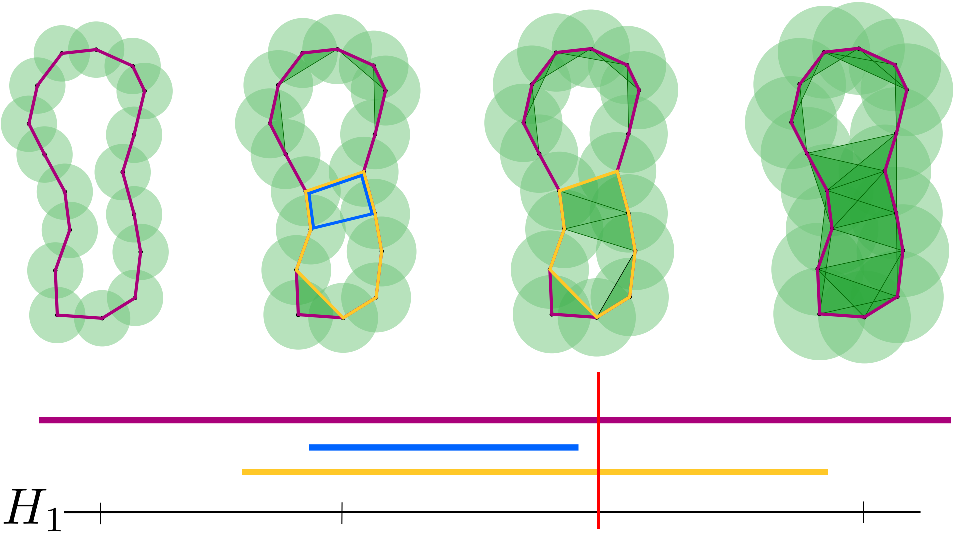
Persistent homology pipeline



Research question about geometric complexes

What are interesting topological properties and how do they interact with the persistent homology pipeline?

Nerve theorems



Nerve theorems - issues

- many original results use concepts and tools that are now obscure

Nerve theorems - issues

- many original results use concepts and tools that are now obscure
- wrongly stated nerve theorems and no proofs:

  Theorem 2.1 (Nerve Theorem [46, 300]). *Given a finite cover $\mathfrak{U}$ (open or closed) of a metric space M, the underlying space $|N(\mathfrak{U})|$ is homotopy equivalent to M if every non-empty intersection $\cap_{i=0}^{k} U_{\alpha_i}$ of cover elements is homotopy equivalent to a point, that is, contractible.*

Nerve theorems - issues

- many original results use concepts and tools that are now obscure
- wrongly stated nerve theorems and no proofs:

Theorem 2.1 (Nerve Theorem [46, 300]). *Given a finite cover $\mathfrak{U}$ (open or ~~closed~~) of a metric space M, the underlying space $|N(\mathfrak{U})|$ is homotopy equivalent to M if every non-empty intersection $\cap_{i=0}^{k} U_{\alpha_i}$ of cover elements is homotopy equivalent to a point, that is, contractible.*

Nerve theorems - issues

- many original results use concepts and tools that are now obscure
- wrongly stated nerve theorems and no proofs:

  Theorem 2.1 (Nerve Theorem [46, 300]). *Given a finite cover $\mathfrak{U}$ (open or ~~closed~~) of a metric space M, the underlying space $|N(\mathfrak{U})|$ is homotopy equivalent to M if every non-empty intersection $\cap_{i=0}^{k} U_{\alpha_i}$ of cover elements is homotopy equivalent to a point, that is, contractible.*

- no general-purpose and complete discussion of this topic in the literature that is tailored to the use cases in topological data analysis (e.g., functorial closed cover nerve theorems)

Nerve theorems - contributions

- provide a thorough overview to (functorial) nerve theorems for open/closed covers

$$\begin{tikzcd} {|\mathrm{Nrv}(\mathcal{U}_r)|} \arrow[r] \arrow[d, "\simeq"] \arrow[dr, phantom, "\circlearrowright ?"] & {|\mathrm{Nrv}(\mathcal{U}_l)|} \arrow[d, "\simeq"'] \\ X_r \arrow[r] & X_l \end{tikzcd}$$

U. Bauer, M. Kerber, F. Roll, and A. Rolle
A unified view on the functorial nerve theorem and its variations
Expositiones Mathematicae, 2023. doi:10.1016/j.exmath.2023.04.005

Nerve theorems - contributions

- provide a thorough overview to (functorial) nerve theorems for open/closed covers

$$\begin{tikzcd}
{|\mathrm{Nrv}(\mathcal{U}_r)|} \arrow[r] \arrow[d, "\simeq"] \arrow[dr, phantom, "\circlearrowright ?"] & {|\mathrm{Nrv}(\mathcal{U}_l)|} \arrow[d, "\simeq"'] \\
X_r \arrow[r] & X_l
\end{tikzcd}$$

- use model categories to establish a "unified" nerve theorem

U. Bauer, M. Kerber, F. Roll, and A. Rolle
A unified view on the functorial nerve theorem and its variations
Expositiones Mathematicae, 2023. doi:10.1016/j.exmath.2023.04.005

Nerve theorems - contributions

- provide a thorough overview to (functorial) nerve theorems for open/closed covers

$$\begin{tikzcd} |\mathrm{Nrv}(\mathcal{U}_r)| \arrow[r] \arrow[d, "\simeq"] \arrow[dr, phantom, "\circlearrowright ?"] & |\mathrm{Nrv}(\mathcal{U}_l)| \arrow[d, "\simeq"'] \\ X_r \arrow[r] & X_l \end{tikzcd}$$

- use model categories to establish a "unified" nerve theorem
  - Quillen (1967): formalizes similarities between homotopy theory and homological algebra

U. Bauer, M. Kerber, F. Roll, and A. Rolle
A unified view on the functorial nerve theorem and its variations
Expositiones Mathematicae, 2023. doi:10.1016/j.exmath.2023.04.005

Nerve theorems - contributions

- provide a thorough overview to (functorial) nerve theorems for open/closed covers

$$
\begin{tikzcd}
|\mathrm{Nrv}(\mathcal{U}_r)| \arrow[r] \arrow[d, "\simeq"] \arrow[dr, phantom, "\circlearrowright ?"] & |\mathrm{Nrv}(\mathcal{U}_l)| \arrow[d, "\simeq"'] \\
X_r \arrow[r] & X_l
\end{tikzcd}
$$

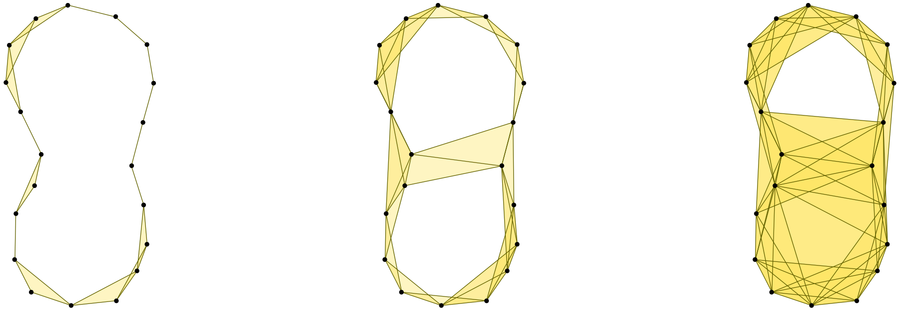
- use model categories to establish a "unified" nerve theorem
  - Quillen (1967): formalizes similarities between homotopy theory and homological algebra
- give various counterexamples that show the non-obvious subtleties of nerve theorems

U. Bauer, M. Kerber, F. Roll, and A. Rolle
A unified view on the functorial nerve theorem and its variations
Expositiones Mathematicae, 2023. doi:10.1016/j.exmath.2023.04.005

Vietoris–Rips complexes

Definition. Let X be a metric space. The Vietoris-Rips complex at scale r is the simplicial complex

$$\mathrm{Rips}_r(X) = \{S \subseteq X \text{ finite} \mid S \neq \emptyset,\ \operatorname{diam} S \leq r\}.$$



Vietoris–Rips complexes

Applications

- For $r \to 0$: Used by Leopold Vietoris (1927) to extend homology theory to metric spaces.

Vietoris–Rips complexes

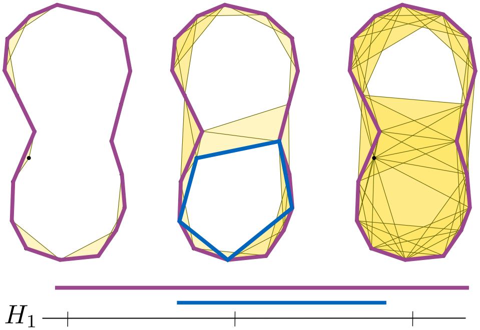
Applications

- For $r \to 0$: Used by Leopold Vietoris (1927) to extend homology theory to metric spaces.
- For $r \to \infty$: Used by Eliyahu Rips and Mikhael Gromov (1987) to study hyperbolic groups.

Vietoris–Rips complexes

Applications

- For $r \to 0$: Used by Leopold Vietoris (1927) to extend homology theory to metric spaces.
- For $r \to \infty$: Used by Eliyahu Rips and Mikhael Gromov (1987) to study hyperbolic groups.
- For all $r > 0$: Used in topological data analysis (nowadays) in the context of persistent homology:



Application of Vietoris–Rips persistent homology

COVID-19 genetic evolution data

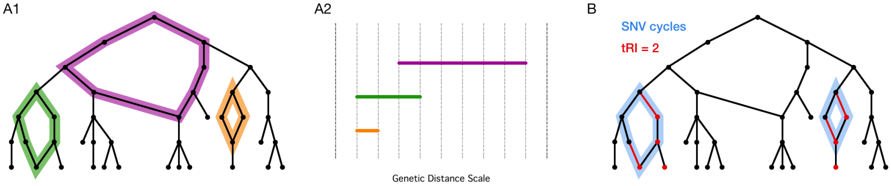


Figure 2. Topological data analysis quantifies convergent evolution. (A) Persistent homology detects reticulate events in viral evolution by means of a persistence barcode. Each bar in the barcode (A2) corresponds to a topological cycle in the reticulate phylogeny (A1). Bars at small genetic distance scales are expected to correspond mainly to homoplasies, while recombination events typically produce topological features at larger scales. (B) SNV cycles are topological cycles in the reticulate phylogeny for which adjacent sequences differ by single nucleotide variations (SNV) only. Under the assumption of single substitutions per site, any SNV in

M. Bleher, L. Hahn, M. Neumann, J. A. Patino-Galindo, M. Carriere, U. Bauer, R. Rabadan, and A. Ott
Topological data analysis identifies emerging adaptive mutations in SARS-CoV-2
Preprint, arXiv:2106.07292, 2023

Application of Vietoris–Rips persistent homology

COVID-19 genetic evolution data

covid data (≈ 15000 points)	Ripser's runtime
ordered chronologically	1 day

U. Bauer
Ripser: efficient computation of Vietoris–Rips persistence barcodes
Journal of Applied and Computational Topology, doi:10.1007/s41468-021-00071-5, 2021

Application of Vietoris–Rips persistent homology

COVID-19 genetic evolution data

covid data (≈ 15000 points)	Ripser's runtime
ordered chronologically	1 day
ordered reversed chronologically	2 min

U. Bauer
Ripser: efficient computation of Vietoris–Rips persistence barcodes
Journal of Applied and Computational Topology, doi:10.1007/s41468-021-00071-5, 2021

Rips contractibility lemma

Theorem (Rips, Gromov 1987). Let X be a δ-hyperbolic geodesic metric space. Then $\mathrm{Rips}_t(X)$ is contractible for all $t \geq 4\delta$.

Rips contractibility lemma

Theorem (Rips, Gromov 1987). Let X be a δ-hyperbolic geodesic metric space. Then $\mathrm{Rips}_t(X)$ is contractible for all $t \geq 4\delta$.

We address two questions:

1. What about non-geodesic spaces? Finite metric spaces?
2. Connections to Ripser?

Generalized contractibility lemma

Theorem (Bauer, R). Let X be a δ-hyperbolic metric space. Then there is a discrete gradient encoding the collapses

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow \{*\}$$

for all $u > t \geq 4\delta + 2\nu$, where ν is the geodesic defect of X.

U. Bauer, F. Roll
Gromov Hyperbolicity, Geodesic Defect, and Apparent Pairs in Vietoris-Rips Filtrations
SoCG 2022. doi:10.4230/LIPIcs.SoCG.2022.15

Generalized contractibility lemma

Theorem (Bauer, R). Let X be a δ-hyperbolic metric space. Then there is a discrete gradient encoding the collapses

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow \{*\}$$

for all $u > t \geq 4\delta + 2\nu$, where ν is the geodesic defect of X.

- Applies directly to the use case in geometric group theory: X is a hyperbolic group.

U. Bauer, F. Roll
Gromov Hyperbolicity, Geodesic Defect, and Apparent Pairs in Vietoris-Rips Filtrations
SoCG 2022. doi:10.4230/LIPIcs.SoCG.2022.15

Generalized contractibility lemma

Theorem (Bauer, R). Let X be a δ-hyperbolic metric space. Then there is a discrete gradient encoding the collapses

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow \{*\}$$

for all $u > t \geq 4\delta + 2\nu$, where ν is the geodesic defect of X.

- Applies directly to the use case in geometric group theory: X is a hyperbolic group.
- We use discrete Morse theory for infinite complexes in the proof.

U. Bauer, F. Roll
Gromov Hyperbolicity, Geodesic Defect, and Apparent Pairs in Vietoris-Rips Filtrations
SoCG 2022. doi:10.4230/LIPIcs.SoCG.2022.15

Generalized contractibility lemma

Theorem (Bauer, R). Let X be a δ-hyperbolic metric space. Then there is a discrete gradient encoding the collapses

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow \{*\}$$

for all $u > t \geq 4\delta + 2\nu$, where ν is the geodesic defect of X.

- Applies directly to the use case in geometric group theory: X is a hyperbolic group.
- We use discrete Morse theory for infinite complexes in the proof.
- Connections to Ripser?

U. Bauer, F. Roll
Gromov Hyperbolicity, Geodesic Defect, and Apparent Pairs in Vietoris-Rips Filtrations
SoCG 2022. doi:10.4230/LIPIcs.SoCG.2022.15

Generalized contractibility lemma

Collapsing Rips complexes of trees

Let X be the path length metric space for a finite weighted tree $T = (V, E)$.

Generalized contractibility lemma

Collapsing Rips complexes of trees

Let X be the path length metric space for a finite weighted tree $T = (V, E)$.

Theorem (Bauer, R). If X is totally ordered in a *compatible way*, the apparent pairs gradient induces the collapses

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow T_t$$

for every $u > t > 0$ such that no edge $e \in E$ has length $l(e) \in (t, u]$.

Generalized contractibility lemma

Collapsing Rips complexes of trees

Let X be the path length metric space for a finite weighted tree $T = (V, E)$.

Theorem (Bauer, R). If X is totally ordered in a *compatible way*, the apparent pairs gradient induces the collapses

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow T_t$$

for every $u > t > 0$ such that no edge $e \in E$ has length $l(e) \in (t, u]$.

- In this case, Ripser computes the persistent homology without a single matrix operation.
  - Heuristically explains its outstanding performance on genetic distances (close to tree metrics).

Generalized contractibility lemma

Collapsing Rips complexes of trees

Let X be the path length metric space for a finite weighted tree $T = (V, E)$.

Theorem (Bauer, R). If X is totally ordered in a *compatible way*, the apparent pairs gradient induces the collapses

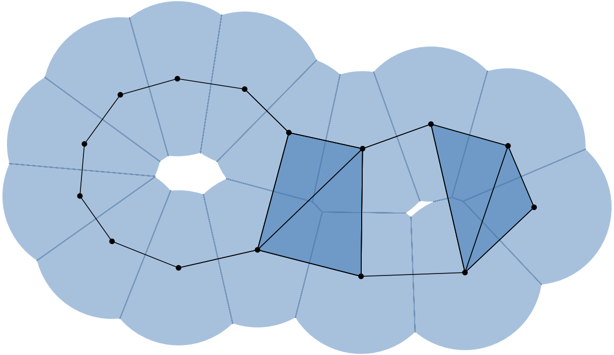
$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow T_t$$

for every $u > t > 0$ such that no edge $e \in E$ has length $l(e) \in (t, u]$.

- In this case, Ripser computes the persistent homology without a single matrix operation.
  - Heuristically explains its outstanding performance on genetic distances (close to tree metrics).
- We use symbolic pertubation of edge lengths in the proof.
  - If X is generic, the diameter function is a generalized discrete Morse function.

Delaunay complexes

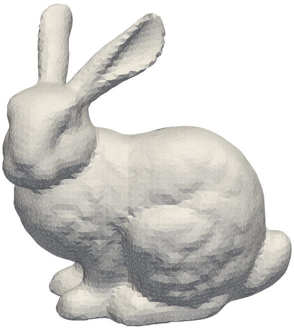
Definition. The *Delaunay complex* $\mathrm{Del}_r(X)$, or *alpha complex*, of $X \subseteq \mathbb{R}^d$ is the nerve of the cover by closed Voronoi balls of radius r centered at points in X.



Wrap

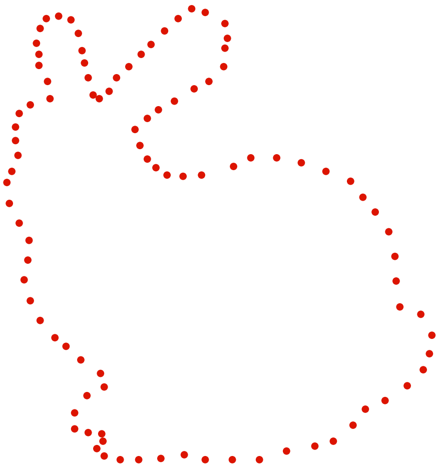
- Originally introduced by Edelsbrunner (1995) as a subcomplex of the Delaunay triangulation for surface reconstruction, using flow lines associated to Euclidean distance functions
- Redeveloped using discrete Morse theory (Forman 1998) by Bauer & Edelsbrunner (2014/17)

Delaunay complex



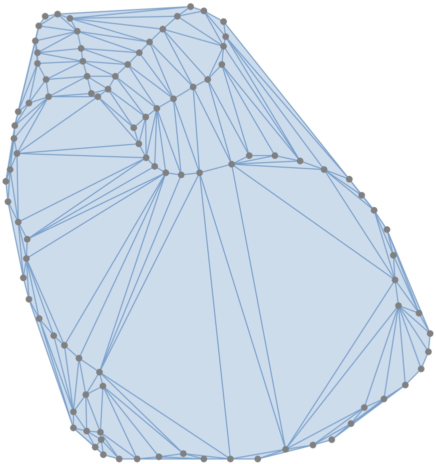
Wrap complex

Wrap complexes



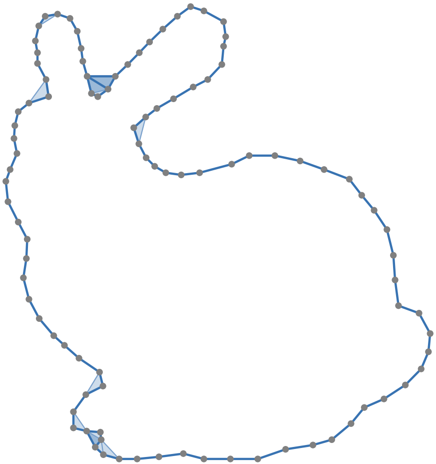
Point cloud

Wrap complexes



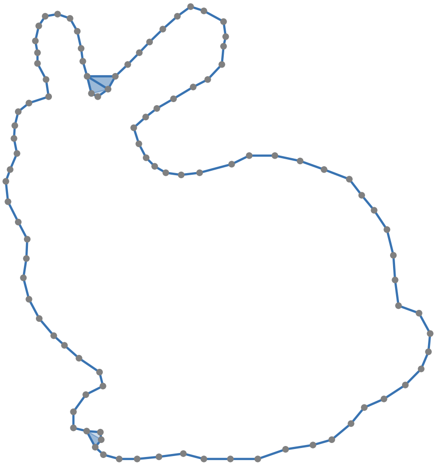
Delaunay triangulation

Wrap complexes



Delaunay complex

Wrap complexes

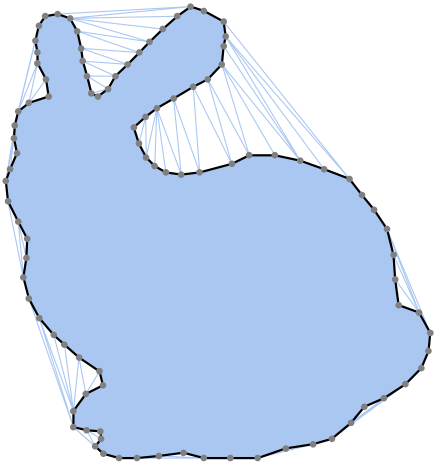


Wrap complex

Wrap complexes

Can we do better?

Exhaustively reduced cycles



Reduction process

Algebraic gradient flows and persistent homology

Loose ends in the literature:

- Cohen-Steiner–Lieutier–Vuillamy (2022) consider shape reconstruction from point clouds

Algebraic gradient flows and persistent homology

Loose ends in the literature:

- Cohen-Steiner–Lieutier–Vuillamy (2022) consider shape reconstruction from point clouds
  - using lexicographically minimal cycles $\min_{c \in C_*}(z + \partial c)$

Algebraic gradient flows and persistent homology

Loose ends in the literature:

- Cohen-Steiner–Lieutier–Vuillamy (2022) consider shape reconstruction from point clouds
  - using lexicographically minimal cycles $\min_{c \in C_*}(z + \partial c)$
  - and its connection to the persistence computation (lex. minimal $\Leftrightarrow$ exhaustively reduced)

Algebraic gradient flows and persistent homology

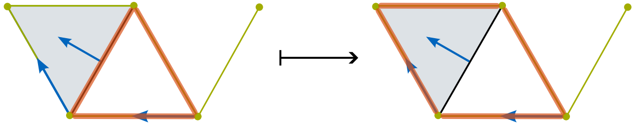
Loose ends in the literature:

- Cohen-Steiner–Lieutier–Vuillamy (2022) consider shape reconstruction from point clouds
  - using lexicographically minimal cycles $\min_{c \in C_*}(z + \partial c)$
  - and its connection to the persistence computation (lex. minimal $\Leftrightarrow$ exhaustively reduced)
  - using a slightly different total order on simplices refining the minimum enclosing radius function

Algebraic gradient flows and persistent homology

Loose ends in the literature:

- Cohen-Steiner–Lieutier–Vuillamy (2022) consider shape reconstruction from point clouds
  - using lexicographically minimal cycles $\min_{c \in C_*}(z + \partial c)$
  - and its connection to the persistence computation (lex. minimal $\Leftrightarrow$ exhaustively reduced)
  - using a slightly different total order on simplices refining the minimum enclosing radius function
- Forman (1998) defines a *flow* $C_*(K) \to C_*(K)$ associated to a discrete gradient



Algebraic gradient flows and persistent homology

We interpret persistent homology in terms of algebraic Morse theory:

- persistence pairs form an algebraic gradient

U. Bauer, F. Roll,
Wrapping Cycles in Delaunay Complexes: Bridging Persistent Homology and Discrete Morse Theory
SoCG 2024. doi:10.4230/LIPIcs.SoCG.2024.15

Algebraic gradient flows and persistent homology

We interpret persistent homology in terms of algebraic Morse theory:

- persistence pairs form an algebraic gradient
- exhaustive Matrix reduction corresponds to gradient flow

U. Bauer, F. Roll,
Wrapping Cycles in Delaunay Complexes: Bridging Persistent Homology and Discrete Morse Theory
SoCG 2024. doi:10.4230/LIPIcs.SoCG.2024.15

Algebraic gradient flows and persistent homology

We interpret persistent homology in terms of algebraic Morse theory:

- persistence pairs form an algebraic gradient
- exhaustive Matrix reduction corresponds to gradient flow
- the lexicographically minimal cycles are invariant under the algebraic gradient flow

U. Bauer, F. Roll,
Wrapping Cycles in Delaunay Complexes: Bridging Persistent Homology and Discrete Morse Theory
SoCG 2024. doi:10.4230/LIPIcs.SoCG.2024.15

Algebraic gradient flows and persistent homology

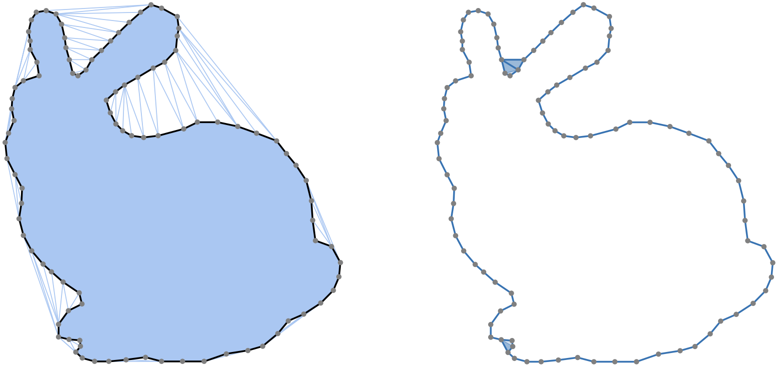
We interpret persistent homology in terms of algebraic Morse theory:

- persistence pairs form an algebraic gradient
- exhaustive Matrix reduction corresponds to gradient flow
- the lexicographically minimal cycles are invariant under the algebraic gradient flow
- connects to generalized discrete Morse theory, and hence to the Wrap complex, through gradient refinements (by apparent pairs)

U. Bauer, F. Roll,
Wrapping Cycles in Delaunay Complexes: Bridging Persistent Homology and Discrete Morse Theory
SoCG 2024. doi:10.4230/LIPIcs.SoCG.2024.15

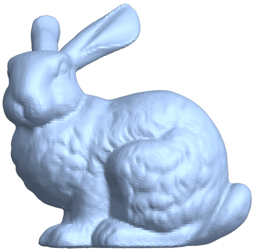
Minimal cycles and Wrap complexes

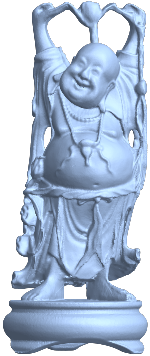
Theorem (Bauer, R). Let $X \subset \mathbb{R}^d$ be a finite subset in general position and let $r \in \mathbb{R}$. Then the lexicographically minimal cycles of $\mathrm{Del}_r(X)$, with respect to the Delaunay-lexicographic order on the simplices, are supported on $\mathrm{Wrap}_r(X)$.

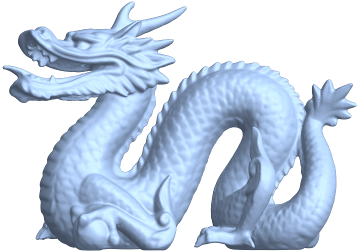


Point cloud reconstruction with most persistent features

The lexicographically minimal cycle, with respect to the Delaunay-lexicographic order on the simplices, corresponding to the interval in the persistence barcode of the Delaunay filtration with the largest death/birth ratio:







```
$ docker build -o output github.com/fabian-roll/wrappingcycles
```

Summary

- Used model categories to establish a "unified" framework to (functorial) nerve theorems, and provided a complete discussion of this topic tailored to the needs in topological data analysis.

Summary

- Used model categories to establish a "unified" framework to (functorial) nerve theorems, and provided a complete discussion of this topic tailored to the needs in topological data analysis.
- Generalized Rips contractibility lemma and explained, in a special case, why the geometry and combinatorics of the input strongly influences the persistence computation time.

Summary

- Used model categories to establish a "unified" framework to (functorial) nerve theorems, and provided a complete discussion of this topic tailored to the needs in topological data analysis.
- Generalized Rips contractibility lemma and explained, in a special case, why the geometry and combinatorics of the input strongly influences the persistence computation time.
- Interpreted persistence in terms of algebraic Morse theory and established a strong connection between a Morse-theoretic and a homological approach to shape reconstruction.