Gromov hyperbolicity, geodesic defect, and apparent pairs in Vietoris-Rips filtrations

Fabian Roll (TUM)

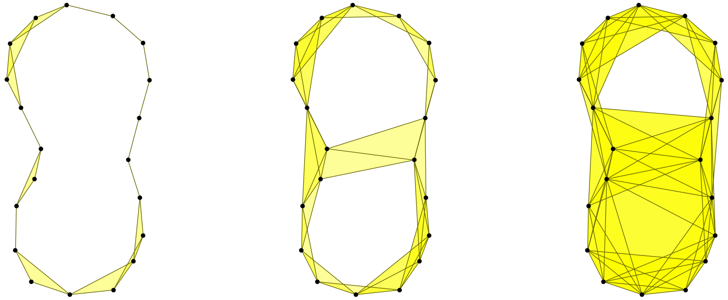
TGDA Seminar
September 27, 2022

joint work with Ulrich Bauer

The Vietoris–Rips complex

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

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



The Vietoris–Rips complex

Applications

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

The Vietoris–Rips complex

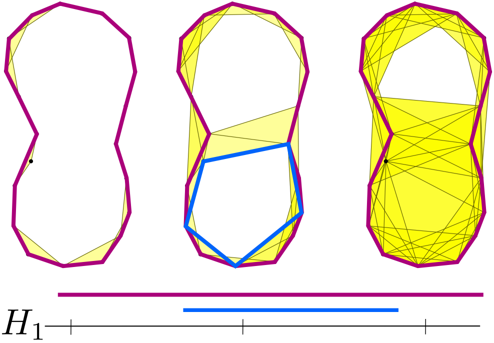
Applications

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

The Vietoris–Rips complex

Applications

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



Application of Vietoris–Rips persistent homology

COVID-19 genetic evolution data

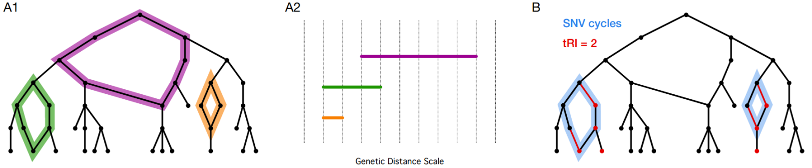


Figure 2. Topological data analysis quantifies convergent evolution. (A) Persistent homology detects reticulate events in viral evolution by means of a persistence barcode. Each bar in the barcode (A2) corresponds to a topological cycle in the reticulate phylogeny (A1). Bars at small genetic distance 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

M. Bleher, L. Hahn, 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, 2021

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

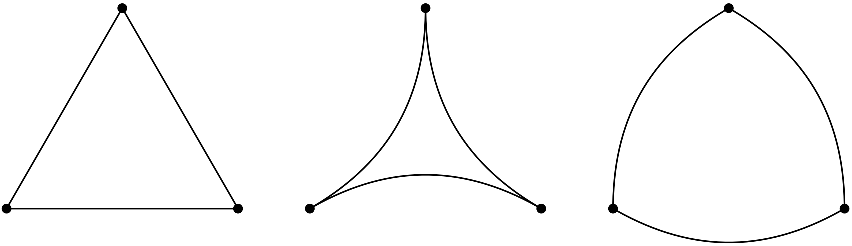
COVID-19 genetic evolution data

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

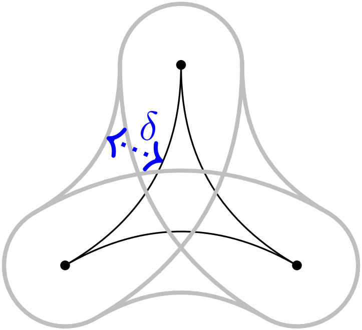
Gromov-hyperbolicity



euclidean triangle hyperbolic triangle spherical triangle

Rips contractibility lemma

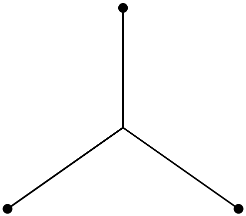
Gromov-hyperbolicity



hyperbolic triangle

Rips contractibility lemma

Gromov-hyperbolicity



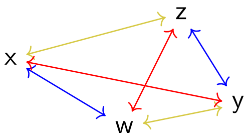
hyperbolic triangle

Rips contractibility lemma

Gromov-hyperbolicity

Definition (four-point condition). A metric space X is (Gromov) δ-*hyperbolic* if for all four points $w, x, y, z \in X$

$$d(x,w) + d(y,z) \leq \max\{d(x,y) + d(z,w), d(x,z) + d(y,w)\} + 2\delta$$

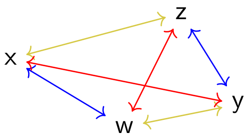


Rips contractibility lemma

Gromov-hyperbolicity

Definition (four-point condition). A metric space X is (Gromov) δ-*hyperbolic* if for all four points $w, x, y, z \in X$

$$d(x,w) + d(y,z) \leq \max\{d(x,y) + d(z,w), d(x,z) + d(y,w)\} + 2\delta$$



Example. finite metric spaces, trees are 0-hyperbolic, hyperbolic plane, ...

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 $t \geq 4\delta$.

Any hyperbolic group G

- is finitely generated and finitely presented.
- admits an Eilenberg–MacLane space $K(G,1)$ with finitely many cells in each dimension.

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 $t \geq 4\delta$.

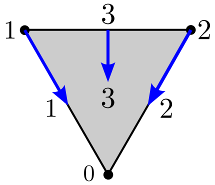
Any hyperbolic group G

- is finitely generated and finitely presented.
- admits an Eilenberg–MacLane space $K(G, 1)$ with finitely many cells in each dimension.

We address two questions:

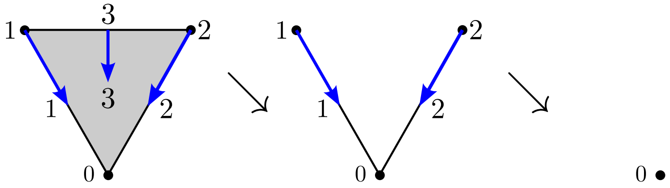
1. What about non-geodesic spaces? Finite metric spaces?
2. Connections to Ripser?

Discrete Morse theory (Forman 1998)



- Discrete Morse function $K \to \mathbb{R}$ with discrete gradient.

Discrete Morse theory (Forman 1998)

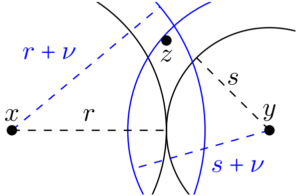


- Discrete Morse function $K \to \mathbb{R}$ with discrete gradient.
- They induce collapses that preserve the homotopy type.

Generalized contractibility lemma

The geodesic defect

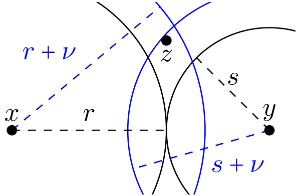
Definition (Bonk, Schramm 2000). The metric space X is *ν-geodesic* if for all $x, y \in X$ and $r, s \geq 0$ with $r + s = d(x, y)$ there exists $z \in X$ with $d(x, z) \leq r + \nu$ and $d(y, z) \leq s + \nu$.



Generalized contractibility lemma

The geodesic defect

Definition (Bonk, Schramm 2000). The metric space X is ν-*geodesic* if for all $x, y \in X$ and $r, s \geq 0$ with $r + s = d(x, y)$ there exists $z \in X$ with $d(x, z) \leq r + \nu$ and $d(y, z) \leq s + \nu$.



- $\nu \geq \frac{1}{2} \inf_{x \neq y} d(x, y)$
- X is r-geodesic if it is an r-dense subset of a geodesic metric space

Generalized contractibility lemma

Theorem (Bauer, R). Let X be a finite δ-hyperbolic metric space. Then there exists 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

Generalized contractibility lemma

Theorem (Bauer, R). Let X be a finite δ-hyperbolic metric space. Then there exists 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

- Vietoris–Rips persistent homology of trees (0-hyperbolic) is concentrated in degree zero.

Generalized contractibility lemma

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

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

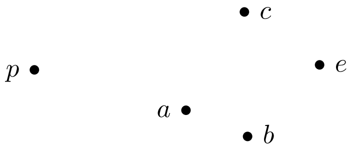
for all $u > t \geq 4\delta + 2\nu$, where ν is the geodesic defect of X

- Vietoris–Rips persistent homology of trees (0-hyperbolic) is concentrated in degree zero.
- Connections to Ripser?

Generalized contractibility lemma

Sketch of proof

Let $t \geq 4\delta + 2\nu$.

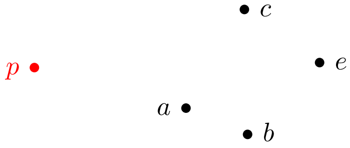


Generalized contractibility lemma

Sketch of proof

Let $t \geq 4\delta + 2\nu$.

- Sort X according to the distance to *any* point $p \in X$

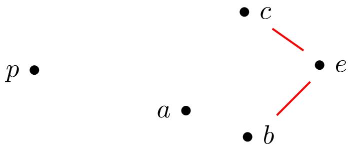


Generalized contractibility lemma

Sketch of proof

Let $t \geq 4\delta + 2\nu$.

- Sort X according to the distance to *any* point $p \in X$
- Consider the largest point e and its cofaces in $\mathrm{Rips}_t(X)$



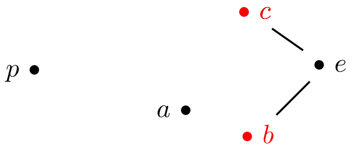
Generalized contractibility lemma

Sketch of proof

Let $t \geq 4\delta + 2\nu$.

- Sort X according to the distance to *any* point $p \in X$
- Consider the largest point e and its cofaces in $\mathrm{Rips}_t(X)$

Goal: Show that the link of e is collapsible



Generalized contractibility lemma

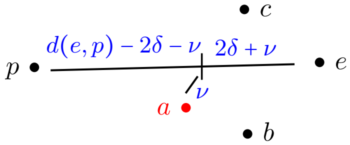
Sketch of proof

Let $t \geq 4\delta + 2\nu$.

- Sort X according to the distance to *any* point $p \in X$
- Consider the largest point e and its cofaces in $\mathrm{Rips}_t(X)$

Goal: Show that the link of e is collapsible

- Find $a \in X$: $d(a,p) \leq d(e,p) - 2\delta$ and $d(a,e) \leq 2\delta + 2\nu \leq t - 2\delta$



Generalized contractibility lemma

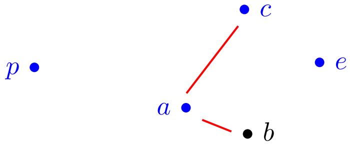
Sketch of proof

Let $t \geq 4\delta + 2\nu$.

- Sort X according to the distance to *any* point $p \in X$
- Consider the largest point e and its cofaces in $\mathrm{Rips}_t(X)$

Goal: Show that the link of e is collapsible

- Find $a \in X$: $d(a, p) \leq d(e, p) - 2\delta$ and $d(a, e) \leq 2\delta + 2\nu \leq t - 2\delta$
- Use the four-point condition to estimate $d(a, b), d(a, c) \leq t$



Generalized contractibility lemma

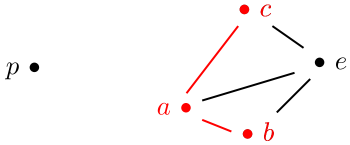
Sketch of proof

Let $t \geq 4\delta + 2\nu$.

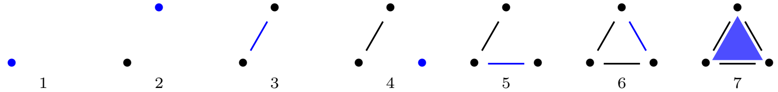
- Sort X according to the distance to *any* point $p \in X$
- Consider the largest point e and its cofaces in $\mathrm{Rips}_t(X)$

Goal: Show that the link of e is collapsible

- Find $a \in X$: $d(a,p) \leq d(e,p) - 2\delta$ and $d(a,e) \leq 2\delta + 2\nu \leq t - 2\delta$
- Use the four-point condition to estimate $d(a,b), d(a,c) \leq t$
- The link of e is a cone with apex a and $\mathrm{Rips}_t(X) \searrow \mathrm{Rips}_t(X \setminus \{e\})$



Computing persistent homology



	1	2	3	4	5	6	7
1			1		1		
2			1			1	
3							1
4					1	1	
5							1
6							1
7							

R

$= D \cdot$

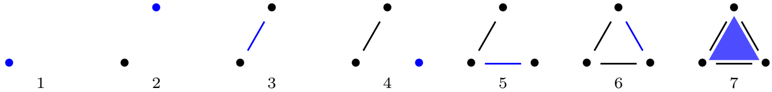
	1	2	3	4	5	6	7
1	1						
2		1					
3			1				
4				1			
5					1		
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1		
2			1			1	
3							1
4					1	1	
5							1
6							1
7							

R

$= D \cdot$

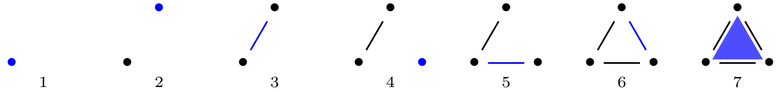
	1	2	3	4	5	6	7
1	1						
2		1					
3			1				
4				1			
5					1		
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1	1	
2			1			1	
3							1
4					1	0	
5							1
6							1
7							

R

$= D \cdot$

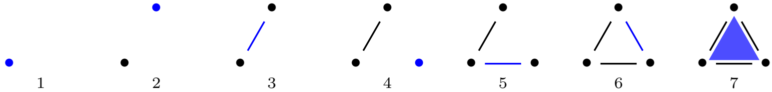
	1	2	3	4	5	6	7
1	1						
2		1					
3			1				
4				1			
5					1	1	
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1	1	
2			1			1	
3							1
4					1		
5							1
6							1
7							

R

$= D \cdot$

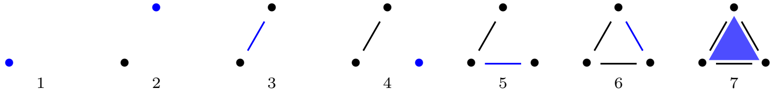
	1	2	3	4	5	6	7
1	1						
2		1					
3			1				
4				1			
5					1	1	
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1	0	
2			1			0	
3							1
4					1		
5							1
6							1
7							

R

$= D \cdot$

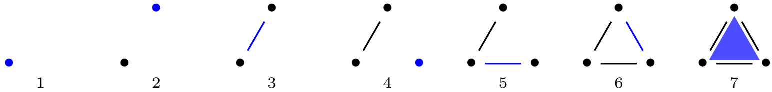
	1	2	3	4	5	6	7
1	1						
2		1					
3			1			1	
4				1			
5					1	1	
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1		
2			1				
3							1
4					1		
5							1
6							1
7							

R

$= D \cdot$

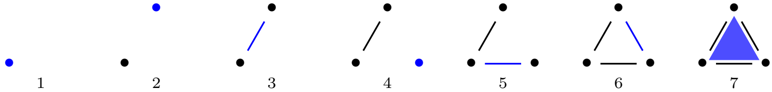
	1	2	3	4	5	6	7
1	1						
2		1					
3			1			1	
4				1			
5					1	1	
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1		
2			1				
3							1
4					1		
5							1
6							1
7							

R

$= D \cdot$

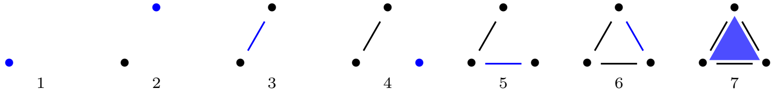
	1	2	3	4	5	6	7
1	1						
2		1					
3			1			1	
4				1			
5					1	1	
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\}$

Computing persistent homology



	1	2	3	4	5	6	7
1			1		1		
2			1				
3							1
4					1		
5							1
6							1
7							

R

$= D \cdot$

	1	2	3	4	5	6	7
1	1						
2		1					
3			1			1	
4				1			
5					1	1	
6						1	
7							1

V

Algorithm:

- while $\exists i < j$ with $\operatorname{pivot} R_i = \operatorname{pivot} R_j$
  - add R_i to R_j and V_i to V_j
- Barcode $\{[\operatorname{pivot} R_i, i) \mid \operatorname{pivot} R_i \neq 0\} = \{[2,3), [4,5), [6,7)\}$

Computing persistent homology

The diam-lexicographic filtration

We use the *lexicographic refinement* of the Vietoris–Rips filtration:

- choose a total order on the vertices
- order simplices by diameter
- order simplices with the same diameter lexicographically

Computing persistent homology

Apparent pairs

Ripser uses the following construction for a computational shortcut:

Definition. In a simplexwise filtration $(K_i = \{\sigma_1, \dots, \sigma_i\})_i$, a pair of simplices (σ, τ) is an *apparent pair* if

- σ latest proper face of τ, and
- τ is the earliest proper coface of σ.

Computing persistent homology

Apparent pairs

Ripser uses the following construction for a computational shortcut:

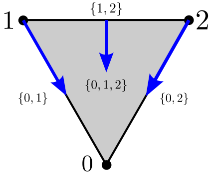
Definition. In a simplexwise filtration $(K_i = \{\sigma_1, \ldots, \sigma_i\})_i$, a pair of simplices (σ, τ) is an *apparent pair* if

- σ latest proper face of τ, and
- τ is the earliest proper coface of σ.

Lemma. If (σ_i, σ_j) is an apparent pair, then $[i, j)$ is an interval in the persistence barcode.

Apparent pairs

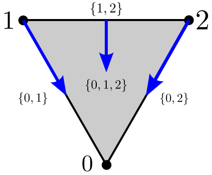
Lemma. The apparent pairs form a discrete gradient.



Apparent pairs

Discrete Morse theory

Lemma. The apparent pairs form a discrete gradient.



Remark. Kahle (2009) uses a specific apparent pairs gradient to study random Vietoris–Rips complexes above the thermodynamic limit.

Collapsing Rips complexes of trees

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

Collapsing Rips complexes of trees

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

- order points away from an arbitrarily chosen root

Collapsing Rips complexes of trees

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

- order points away from an arbitrarily chosen root
- T_t subforest on V with all edges in E of length at most t

Collapsing Rips complexes of trees

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

- order points away from an arbitrarily chosen root
- T_t subforest on V with all edges in E of length at most t

Theorem (Bauer, R). If X is ordered in a compatible way, the apparent pairs gradient induces a sequence of 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 $l(e) \in (t, u]$.

Collapsing Rips complexes of trees

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

- order points away from an arbitrarily chosen root
- T_t subforest on V with all edges in E of length at most t

Theorem (Bauer, R). If X is ordered in a compatible way, the apparent pairs gradient induces a sequence of 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 $l(e) \in (t, u]$.

- Ripser computes the persistent homology of X without a single column operation.

Collapsing Rips complexes of trees

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

- order points away from an arbitrarily chosen root
- T_t subforest on V with all edges in E of length at most t

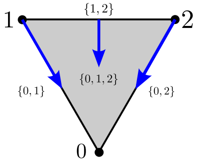
Theorem (Bauer, R). If X is ordered in a compatible way, the apparent pairs gradient induces a sequence of collapses

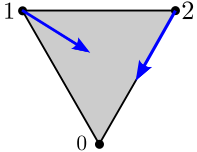
$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \searrow T_t$$

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

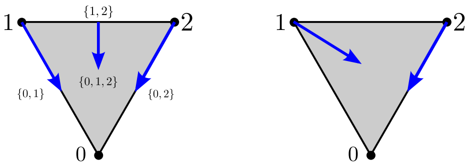
- Ripser computes the persistent homology of X without a single column operation.
- Explains Ripser's outstanding performance on genetic distances.

Generalized discrete Morse theory





Generalized discrete Morse theory

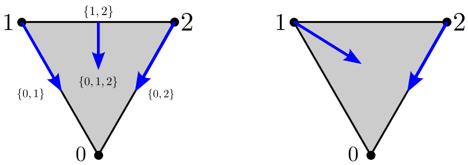


Definition. A monotone $f\colon K \to \mathbb{R}$ is a *generalized discrete Morse function* if there exists a partition of K, called the *discrete gradient* of f,

$$W = \{[\sigma, \tau] = \{\eta \mid \sigma \subseteq \eta \subseteq \tau\}\}$$

such that for $\sigma \subseteq \tau$ it is equivalent $f(\sigma) = f(\tau) \Leftrightarrow \sigma, \tau \in I \in W$.

Generalized discrete Morse theory



Definition. A monotone $f\colon K \to \mathbb{R}$ is a *generalized discrete Morse function* if there exists a partition of K, called the *discrete gradient* of f,

$$W = \{[\sigma, \tau] = \{\eta \mid \sigma \subseteq \eta \subseteq \tau\}\}$$

such that for $\sigma \subseteq \tau$ it is equivalent $f(\sigma) = f(\tau) \Leftrightarrow \sigma, \tau \in I \in W$.

Example (Bauer, Edelsbrunner 2016). For finite $X \subseteq \mathbb{R}^d$ in general spherical position the Čech radius function is a generalized discrete Morse function with its discrete gradient determined by smallest circumspheres.

Generalized discrete Morse theory

Refine W to another discrete gradient

$$\widetilde{W} = \{(\psi \setminus \{v\}, \psi \cup \{v\}) \mid \psi \in [\rho, \phi] \in W,\ v = \min(\phi \setminus \rho)\}$$

by doing a minimal vertex refinement on each interval.

Generalized discrete Morse theory

Refine W to another discrete gradient

$$\widetilde{W} = \{(\psi \smallsetminus \{v\}, \psi \cup \{v\}) \mid \psi \in [\rho, \phi] \in W,\ v = \min(\phi \smallsetminus \rho)\}$$

by doing a minimal vertex refinement on each interval.

Lemma. The zero persistence apparent pairs with respect to the f-lexicographic order are precisely the gradient pairs of $\widetilde{W}$.

Collapsing Rips complexes of trees

Some proof details

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

Collapsing Rips complexes of trees
Some proof details

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

If X is generic (pairwise distances are distinct)

Collapsing Rips complexes of trees

Some proof details

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

If X is generic (pairwise distances are distinct)

- $\mathrm{diam}\colon \mathrm{Cl}(V) \to \mathbb{R}$ is a generalized discrete Morse function

Collapsing Rips complexes of trees

Some proof details

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

If X is generic (pairwise distances are distinct)

- $\mathrm{diam}\colon \mathrm{Cl}(V) \to \mathbb{R}$ is a generalized discrete Morse function
- only the edges E are critical and therefore the diameter function induces the collapses

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

Collapsing Rips complexes of trees

Some proof details

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

If X is generic (pairwise distances are distinct)

- $\mathrm{diam}\colon \mathrm{Cl}(V) \to \mathbb{R}$ is a generalized discrete Morse function
- only the edges E are critical and therefore the diameter function induces the collapses

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

- for *any* total order on V the apparent pairs gradient for the lexicographically refined Vietoris–Rips filtration also induces these

Collapsing Rips complexes of trees

Some proof details

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

If X is arbitrary

Collapsing Rips complexes of trees

Some proof details

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

If X is arbitrary

- for a *compatible* total order on V a symbolic perturbation scheme on the edges is induced, establishing the generic situation locally

Collapsing Rips complexes of trees

Some proof details

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

If X is arbitrary

- for a *compatible* total order on V a symbolic perturbation scheme on the edges is induced, establishing the generic situation locally
- This suffices to prove the existence of collapses

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

Collapsing Rips complexes of trees

Some proof details

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

If X is arbitrary

- for a *compatible* total order on V a symbolic perturbation scheme on the edges is induced, establishing the generic situation locally
- This suffices to prove the existence of collapses

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

- A detailed analysis shows that the apparent pairs gradient for the lexicographically refined Vietoris–Rips filtration induces these

Summary

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

Summary

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

- Extended the Contractibility Lemma to finite metric spaces and made it filtration compatible.

Summary

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

- Extended the Contractibility Lemma to finite metric spaces and made it filtration compatible.
- Identified a subclass of metric spaces for which the persistent homology computation is very efficient.