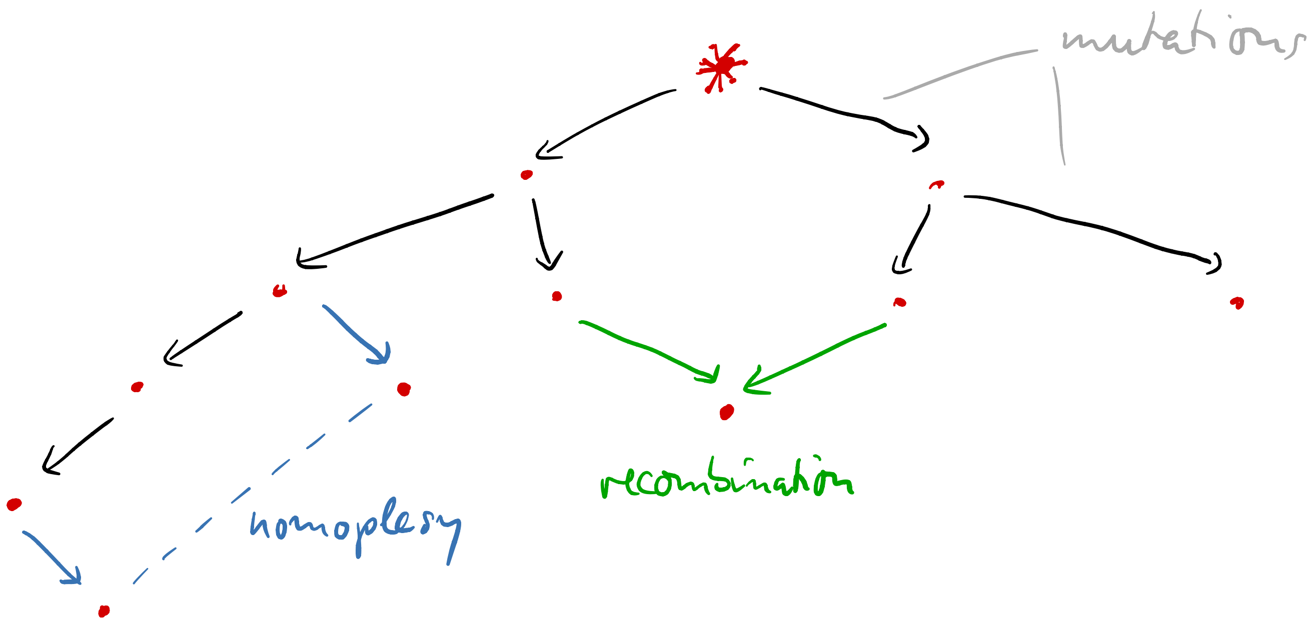
Gromov-hyperbolicity, geodesic defect & apparent pairs in Rips filtrations

Ulrich Bauer
TUM

joint w/ Fabian Roll

Funding: DFG (SFB/TR 109 Discretization in Geometry & Dynamics)

Motivation ①: Vietoris-Rips persistence of COVID-19 genetic evolution data



- Early detection of adaptive mutations (occurring independently multiple times)
- Computation faster than usual

New insights ⇝ huge speedup; without changes to code!
(~1 day ⇝ 2 min)

Motivation ②: The Rips Lemma

Lemma [Rips 198?; Gromov 1988]
Let X be a δ-hyperbolic geodesic space. Then

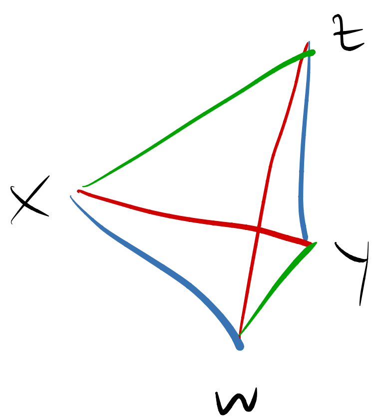
$$\mathrm{Rips}_t(X) \simeq \{*\} \quad \text{for } t \geq 4\delta.$$

- What about non-geodesic spaces?
- In particular, finite metric spaces?
- Connection to Ripser? Collapse by apparent pairs?

Gromov - hyperbolicity

<u>Definition</u> A metric space (X, d) is <u>δ-hyperbolic</u> if $\forall\, w, x, y, z \in X$,

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



- Other notions exist for geodesic spaces.
- A metric tree is 0-hyperbolic.

A Rips Lemma for non-geodesic spaces

<u>**Theorem**</u> (B, Roll)

Let X be a δ-hyperbolic metric space. Then

$$\mathrm{Rips}_t(X) \simeq \{*\} \quad \text{for } t \geq 4\delta + \nu,$$

where ν : <u>geodesic defect</u> of X.

Filtered Rips Lemma for finite X

Theorem (B, Roll)

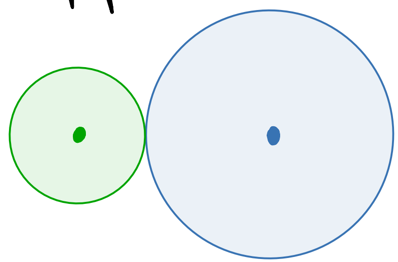
Let X be a δ-hyperbolic finite metric space with geod. defect ν. Then

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

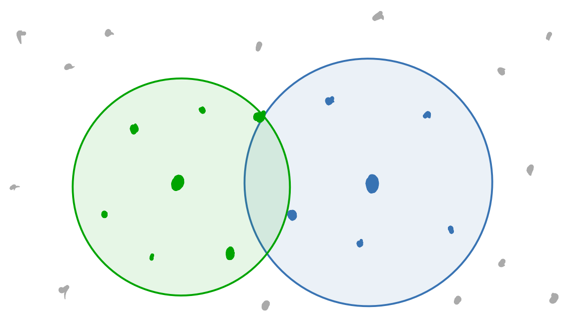
by a single discrete gradient for all $u \geq t \geq 4\delta + \nu$.

Disjoint balls

- In a geodesic space, balls $B_r(x)$ and $B_s(y)$ are disjoint iff $r + s \leq d(x, y)$.



- In general, the radii may be larger.



Geodesic defect

Definition The <u>geodesic defect</u> of a metric space (X, d) is

$$\sup\{r+s-d(x,y) \mid B_r(x)\cap B_s(y)=\emptyset\}.$$

See also:

- U. Lang, 2013 (idea used implicitly)
- P. Johannad, J. Jost 2020 (multiplicative local version)

Bounds on geodesic defect

ν : geodesic defect of finite (X, d)

- lower bound : $\nu \geq \min\limits_{x \neq y \in X} d(x, y)$
- upper bound : $\nu \leq 2 \cdot d_H(X, Z)$

  for $X \subseteq Z$ a geodesic ambient space

Possible ambient spaces for X :

- $\ell^\infty(X)$ (bounded functions $X \to \mathbb{R}$) $x \mapsto d_x$

  (Kuratowski embedding $e : X \to \ell^\infty(X)$)

- <u>tight span / injective envelope</u> : subspace of $\ell^\infty(X)$

Tight span

Given a finite metric space (X, d), the tight span is

$$E(X) := \{ f \in \ell^\infty(X) \mid f(x) + f(y) \geq d(x,y) \ \forall x, y \in X$$
$$\forall x \ \exists y : \ f(x) + f(y) = d(x,y) \}$$

- geodesic
- hyperconvex (metric balls have the Helly property)
  $\Updownarrow$ [Aronszajn – Panitchpakdi 1956]
- injective (in the category of metric spaces)
- contains the Kuratowski embedding $e(X)$
- minimal space with above properties
- contractible [Isbell 1962]

Density of Kuratowski embedding

Theorem (U. Lang 2013 + ε)

X δ-hyperbolic, with geodesic defect ν,

$e : X \hookrightarrow E$ Kuratowski embedding into tight span.

Then
$$d_H(e(X), E) \leq 2\delta + \frac{\nu}{2}.$$

Remark

- $\mathrm{Rips}_{2r}(X) = \check{\mathrm{C}}\mathrm{ech}_r(e(X), E)$ (hyperconvexity)
- For $r \geq 2\delta + \frac{\nu}{2}$: $\check{\mathrm{C}}\mathrm{ech}_r(e(X), E) \simeq E$ (nerve lemma)
- $E \simeq \{*\}$

Rips lemmas for non-geodesic spaces

Corollary $\mathrm{Rips}_t(X) \simeq \{*\}$ for $t \geq 4\delta + \nu$

not given by an explicit collapse

Extending the argument of Gromov–Rips:

Theorem (B, Roll) A single discrete gradient on $\Delta(X)$ induces the collapses

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

for all $u \geq t \geq 4\delta + \nu$.

VR complexes of trees & discrete Morse

Consider a <u>generic</u> finite metric tree $T = (X, E)$ (distinct pairwise distances).

- $\operatorname{diam} : \Delta(X) \to \mathbb{R}$ is a generalized discrete Morse function
- discrete gradient has non critical intervals
  $[e, \Delta_e]$
  where
  - $e \in \binom{V}{2} \setminus E$: a non-tree edge
  - Δ_e : unique max coface w/ diam = length(e)
- Only vertices X and tree edges E are critical

Consequences

- The discrete gradient induces collapses

$$\mathrm{Rips}_t(X) \searrow T_t \quad \text{← filtered tree} \qquad \forall t \in \mathbb{R}$$

and

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \qquad \text{for } (t, u] \cap d(E) = \emptyset$$

- For $t \geq \max\{\mathrm{length}(e) \mid e \in T\}$:

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

- Pers. homology concentrated in deg 0.

Non-generic tree metrics

Example: phylogenetic trees

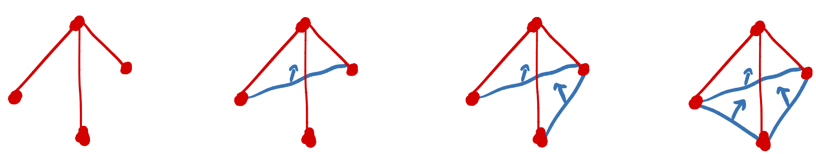
- diam is no longer Morse
- Non-tree edges still have unique max cofaces with same diameter
- There is still a compatible gradient:
  - indep. of choices (<u>canonical gradient</u>)
  - only X, E critical
  - coincides with diam. gradient for generic trees

Symbolic perturbation

- Choose total order on X
- Order edges lexicographically
- Perturb metric symbolically wrt edge order

$\rightsquigarrow$ gradient structure like a generic tree metric

- The canonical gradient (from before) refines any perturbed gradient



Apparent pairs

- K: simplexwise filtration (total order on K)

- (σ, τ) <u>apparent pair</u>:

  $\left.\begin{array}{l} \sigma \text{ is largest face of } \tau \\ \tau \text{ is smallest coface of } \sigma \end{array}\right\}$ in filt. order

$\rightsquigarrow$ persistence pairs & discrete gradient

- Lexicographically refined Rips filtration:
  order simplices by
  - diameter, then
  - lexicographically

Apparent pairs Rips Lemma for trees

Theorem (B, Roll) $T = (X, E)$ finite metric tree. X totally ordered away from root. Then the apparent pairs refine the perturbed gradient.

- Only X, E critical, all other simplices paired.
- This induces collapses

$$\mathrm{Rips}_t(X) \searrow T_t \qquad \forall t \in \mathbb{R}$$

and

$$\mathrm{Rips}_u(X) \searrow \mathrm{Rips}_t(X) \qquad \text{for } (t, u] \cap d(E) = \emptyset$$

Summary

Rips Lemmas beyond geodesic spaces

- geodesic defect
- filtration compatible
- collapses
- for trees: via apparent pairs

Explains Ripser's outstanding performance on evolution data (almost tree-like)