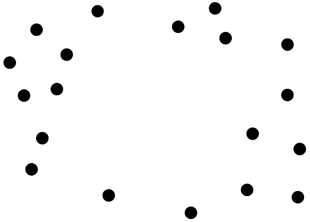
Stop interpolation Topologically

Luciano Melodia

Agenda

Introduction

Topology matters
Guess the manifold



This set of points forms some shape, I would say the shape of an **annulus**.
We want to augment the dataset and accurately represent its shape.
How can we guarantee to maintain shape during interpolation?

Hypothesis

Two main assumptions

Hypothesis 1: Manifold assumption

A point set X lies on a *topological* or even *smooth manifold* $\mathbb{M}$, having a family of continuous/smooth *coordinate systems* to describe it, with $\dim \mathbb{M} \ll \dim X$.

Hypothesis 2: Meaning of shape

The topology of the sublevel sets of a manifold $\mathbb{M}$ underlying a data set can be used to distinguish data up to some equivalence relation.

Simplices

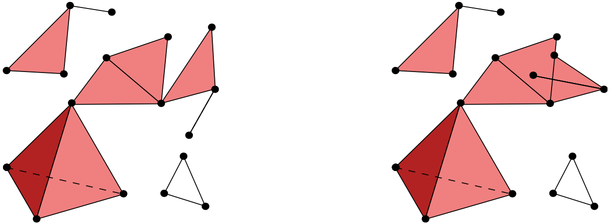
Given a set $X = \{x_0, \cdots, x_k\} \subset \mathbb{R}^d$ of $k+1$ points that do not lie on a hyperplane with dimension less than d, the k-dimensional simplex σ spanned by X is the set of convex combinations, such that

$$\sum_{i=0}^{k} \lambda_i x_i \quad \text{with} \quad \sum_{i=0}^{k} \lambda_i = 1 \quad \text{and} \quad \lambda_i \geq 0.$$

Simplicial complexes

A **simplicial complex** K is a collection of simplices such that:

- Any face of a simplex of K is a simplex of K.
- Any intersection of two simplices in K is either empty or a common face of both.



Valid Invalid

Simplicial complexes upon point sets

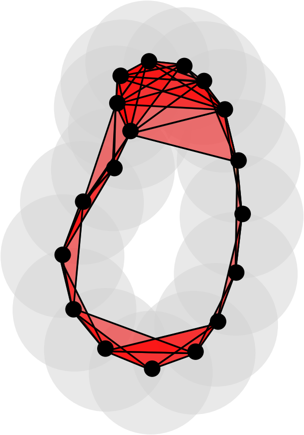
How do we capture the topology of the underlying space?

- Associate a simplicial complex to the point set.
- There is an isomorphism from simplicial to singular homology.
- We study the simplicial complex to study the underlying topological space.

Question: Which (simplicial) complex is a good choice?

Simplicial complexes upon point sets

Spanning the Čech complex



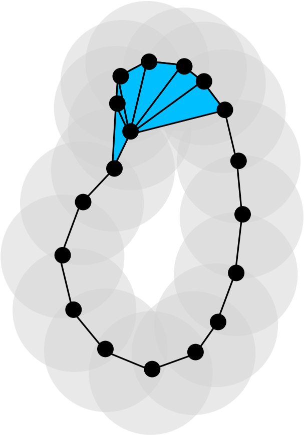
Let the radius $r \geq 0$ be a real number and $B(x, r) = \{y \in \mathbb{R}^d \mid ||x - y|| \leq r\}$ the closed ball centered around $x \in X \subseteq \mathbb{R}^d$. The **Čech complex** for a finite point set X is defined as

$$\check{\text{C}}\text{ech}(X, r) = \left\{ U \subseteq X \mid \bigcap_{x \in U} B(x, r) = \emptyset \right\}.$$

On the left one can see a high dimensional Čech complex projected onto the plane.

Simplicial complexes upon point sets

Spanning the Delaunay complex



The **Voronoi cell** around a point is defined as

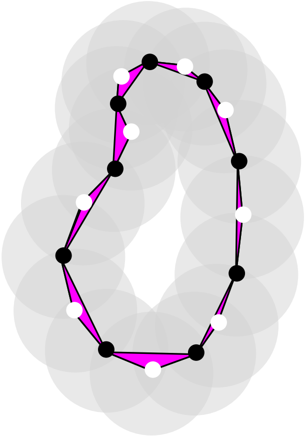
$$\mathrm{Vor}(x) = \left\{ y \in \mathbb{R}^d \;\middle|\; \|y - x\| \leq \|y - z\|, \text{for all } z \in X \right\}.$$

The **Voronoi ball** of x with respect to X is defined as the intersection of the **Voronoi cell** with the closed ball of given radius around this point $\mathrm{VorBall}(x, r) = B(x, r) \cap Vor(x)$. Thus we define

$$\mathrm{Del}(X, r) = \left\{ U \subseteq X \mid \bigcap_{x \in U} \mathrm{VorBall}(x, r) \neq \emptyset \right\}.$$

Simplicial complexes upon point sets

Spanning the Witness complex



We call the subsets $W \subset \mathbb{R}^d$ witnesses and $L \subset \mathbb{R}^d$ landmarks, respectively.
We use $L \subset W$. We say that w is witnessed by σ if $\|w - p\| \leq \|w - q\|$ for all $p \in \sigma$ and $q \in L \setminus \sigma$. The Witness complex $\mathrm{Wit}(L, W)$ consists of all simplices σ, such that any simplex $\sigma' \subseteq \sigma$ has a witness in W.

All points are within W and the white bullets are in L.

Topological inference
Investigating the homology groups

- The homology groups of a fully spanned simplicial complex are trivial.
- We need to capture local properties of data to some extend.

Solution: We use the notion of persistence for homology groups.

Persistent Homology

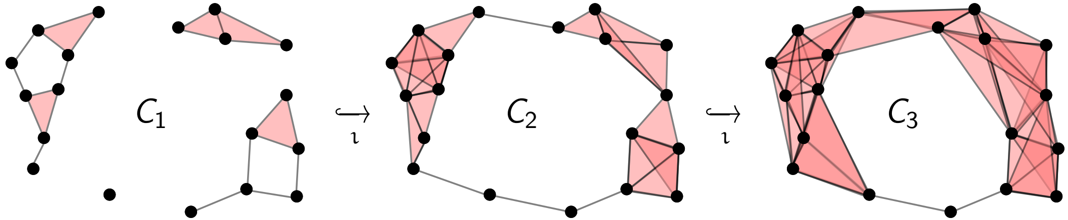
An outline to persistent homology

A brief description

Persistent homology is the study of k-**dimensional abelian groups** attached to a topological space. They are studied algebraically and behave *similarly* to the topological space itself under change revealing properties of the original space they are attached to by their *invariants*.

The connection of abelian groups by a boundary homomorphism allows one to investigate all k-dimensions at once. The induced object is a chain complex.

Chain complex
An example



Chain complex

How does it work?

Homology theory is the study of properties of an object X, such as a set of points, by means of homological (commutative) algebra. One assigns to X a sequence of abelian groups/modules $C_0, C_1, \cdots$ which are connected by homomorphisms $\partial_{k+1} : C_{k+1} \to C_k$ such that $\mathrm{im}\partial_{k+1} \subseteq \ker \partial_k$. Such a structure is called **chain complex**.

$$0 \xrightarrow{\partial_{k+1}} C_k \xrightarrow{\partial_k} C_{k-1} \xrightarrow{\partial_{k-1}} C_{k-2} \xrightarrow{\partial_{k-2}} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} 0$$

Cycle and boundary groups

These subgroups are normal with respect to C_k

The k**th cycle group** Z_k is defined as

$$Z_k = \ker \partial_k = \{c \in C_k \mid \partial_k c = \emptyset\}.$$

An element of this group is called k-**cycle**.

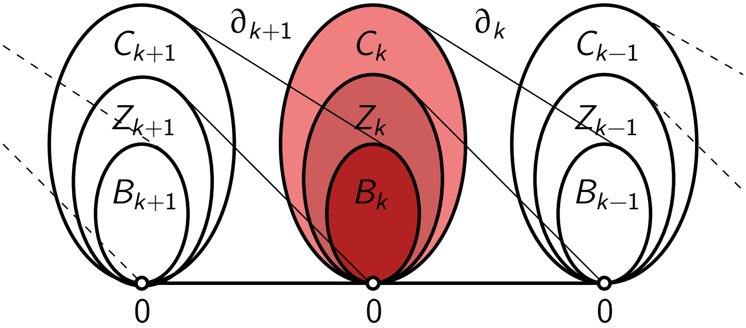
The k**th boundary group** B_k is defined as

$$B_k = \text{im } \partial_{k+1} = \{c \in C_k \mid \exists d \in C_{k+1} : c = \partial_{k+1} d\}.$$

An element of this group is called k-**boundary**.

Putting them into context

Illustration of the chain complex



Following Zomorodian, Edelsbrunner and many more ...

Homology groups

The kth **homology group** of a simplicial complex is defined as

$$H_k(K) = \frac{\ker \partial_k C_k(K)}{\text{im } \partial_{k+1} C_{k+1}(X)}.$$

Intuitively, the kernel of the boundary homomorphism of the kth chain group gives all k-cycles, thus the cycle group, from which we quotient out all elements of the kth boundary group, i.e.

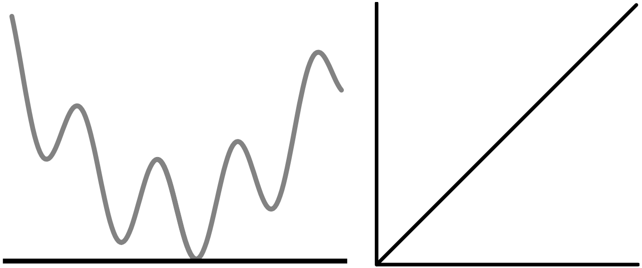
$$H_k(K) = Z_k(K)/B_k(K).$$

Betti numbers

The kth **Betti** number β_k is the rank of the kth homology group $H_k(K)$ of the topological space K.

We'll track the betti numbers to track the amount of holes along the filtration.

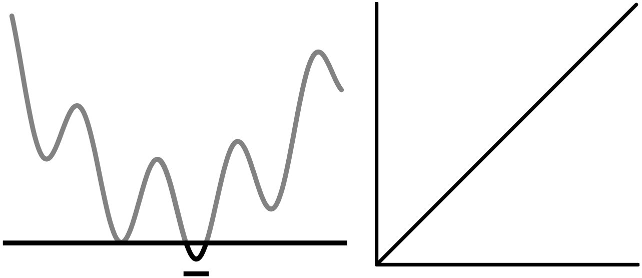
Persistent homology
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

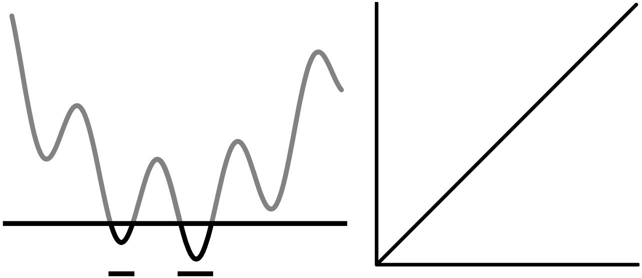
Persistent homology

An example



Following Bastian Rieck: An Introduction to Persistent Homology.

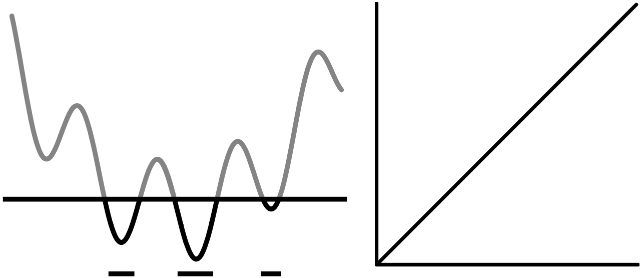
Persistent homology
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

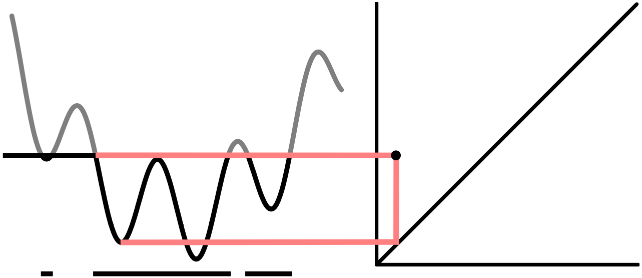
Persistent homology

An example



Following Bastian Rieck: An Introduction to Persistent Homology.

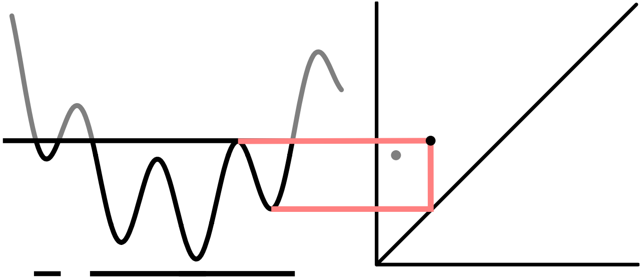
Persistent homology
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

Persistent homology

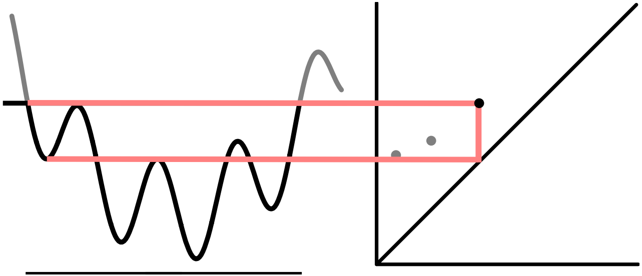
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

Persistent homology

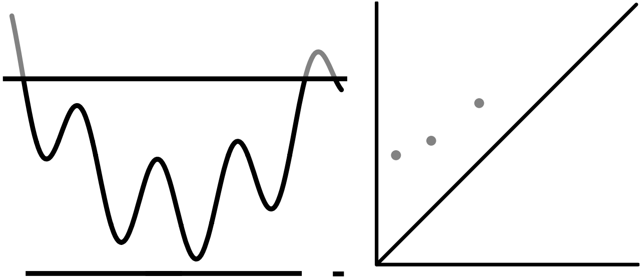
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

Persistent homology

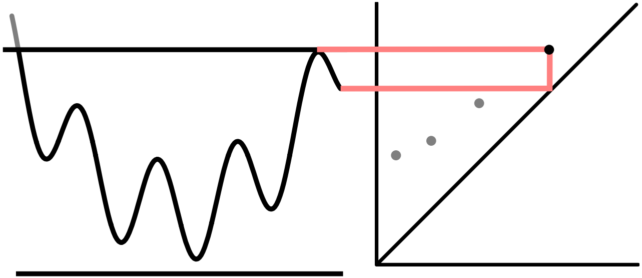
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

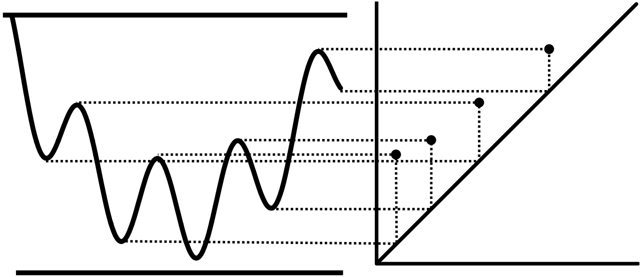
Persistent homology

An example



Following Bastian Rieck: An Introduction to Persistent Homology.

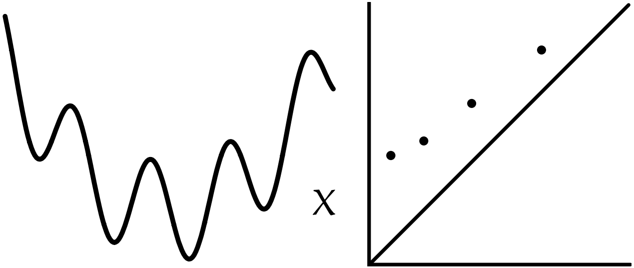
Persistent homology
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

Persistent homology

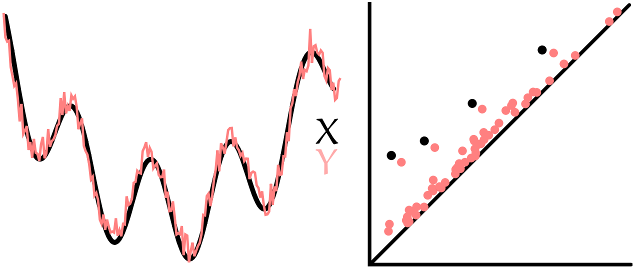
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

Persistence stability

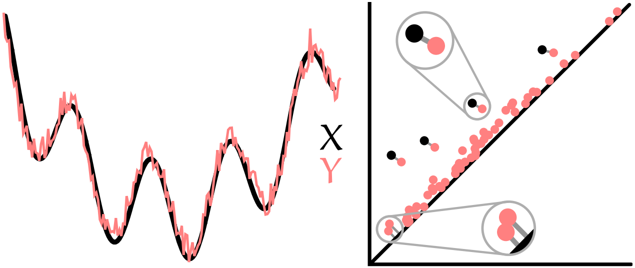
An example



Following Bastian Rieck: An Introduction to Persistent Homology.

Persistence stability

An example

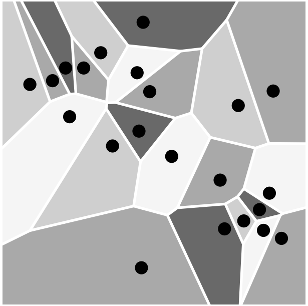


Following Bastian Rieck: An Introduction to Persistent Homology.

$$W_p(X, Y) = \sqrt{\inf_{\varphi : X \to Y} \sum_{x \in X} \|x - \varphi(x)\|_\infty^p}$$

Natural neighbor interpolation

Step 1: Voronoi diagram
Natural neighbor interpolation example

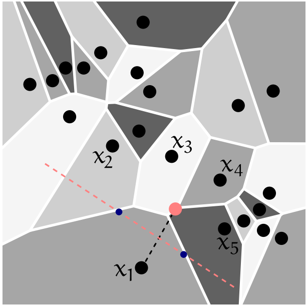


The **Voronoi Diagram** is spanned over the point cloud embedded into some Euclidean space. Technically one does not calculate the Voronoi diagram but the dual **Delaunay triangulation**.

$$\mathrm{Vordgm} = \{x \in X \mid d(x, \lambda_i) \leq d(x, \lambda_j)\}.$$

Step 2: Add a point to the convex hull

We need to compute the weights per region

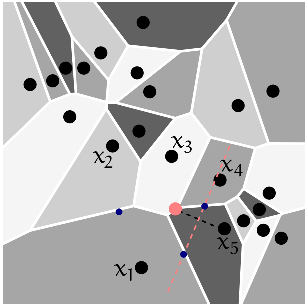


We insert a point into the convex hull of the points $X = \{x_1, \cdots, x_n\}$. The selected position of the new point follows an uniform distribution. The point is added to the set forming the convex hull:

$$\mathrm{conv}X = \left\{ \sum_{i=1}^{n} \lambda_i x_i \;\middle|\; x_i \in X, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0 \right\}.$$

Step 3: Iterative construction of Voronoi regions

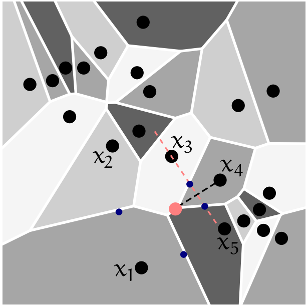
Construction of the new regions



Iteratively, the points of the Voronoi region are connected with a straight line. The perpendicular that bisects the line provides the two intersections to the neighboring Voronoi regions.

Step 3: Iterative construction of Voronoi regions

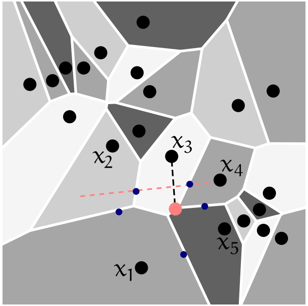
Construction of the new regions



Iteratively, the points of the Voronoi region are connected with a straight line. The perpendicular that bisects the line provides the two intersections to the neighboring Voronoi regions.

Step 3: Iterative construction of Voronoi regions

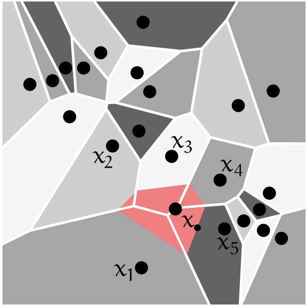
Construction of the new regions



Iteratively, the points of the Voronoi region are connected with a straight line. The perpendicular that bisects the line provides the two intersections to the neighboring Voronoi regions.

Step 3: Iterative construction of Voronoi regions

Construction of the new regions

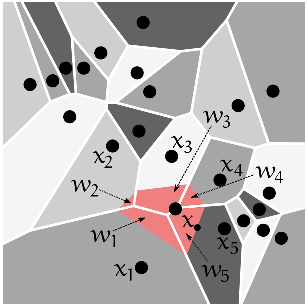


The weights w_i are the *stolen* regions of the new Voronoi region of the added point from all other Voronoi regions.

They can be calculated during construction.

Coordinate shift of the added point

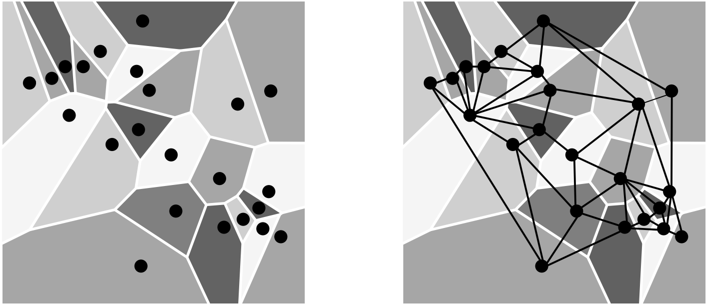
Weighting by Voronoi regions



The algorithm re-weights the **coordinates of a new point** within the convex hull of a set of points by the change of Voronoi regions relative to the Voronoi regions without the added point:

$$\hat{x}^{\bullet} = \sum_{i=1}^{n} w_i x_i^{\bullet}. \tag{1}$$

Duality of Delaunay complex and Voronoi diagram



The Delaunay complex is spanned, connecting each point to its neighbors.

It is always well defined and does not create simplicial complexes which can only be embedded into higher dimensions than the point set.

The Collapsing Sequence

The collapsing sequence
Theorem of simplicial collapse

A familiy of simplices σ of a non-empty finite subset of some set of simplices $\tilde{K}$ is an abstract simplicial complex if for every set $\sigma' \in \sigma$ and every non-empty subset $\sigma'' \subset \sigma'$ the set σ'' also belongs to σ.

Assume σ and σ' are two simplices of $\tilde{K}$, such that $\sigma \subset \sigma'$ and $\dim \sigma < \dim \sigma'$. The face σ' is called free, if it is a maximal face of $\tilde{K}$ and no other maximal face of $\tilde{K}$ contains σ.

The collapsing sequence
Theorem of simplicial collapse

Theorem

Let the simplicial collapse $\searrow$ denote the removal of all σ''-simplices, where $\sigma \subseteq \sigma'' \subseteq \sigma'$, with σ being a free face. If X is a finite set of points in general position in some metric space, then

$$\check{\mathrm{C}}\mathrm{ech}(X, r) \searrow \mathrm{Del}\check{\mathrm{C}}\mathrm{ech}(X, r) \searrow \mathrm{Del}(X, r) \searrow \mathrm{Wrap}(X, r).$$

Established by Bauer and Edelsbrunner: The Morse theory of Čech and Delaunay complexes.

Consequences

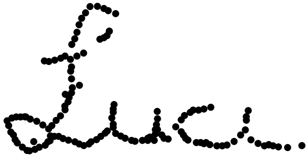
What do we learn?

- The Čech complex is the Nerve of a cover of a topological space.
- The Čech complex collapses to the Delaunay complex.
- The collapse establishes the simple homotopy-equivalence of both.
- Simple homotopy-equivalence implies homotopy-equivalence.
- We can use the Delaunay complex for PH up to simple homotopy-equivalence.

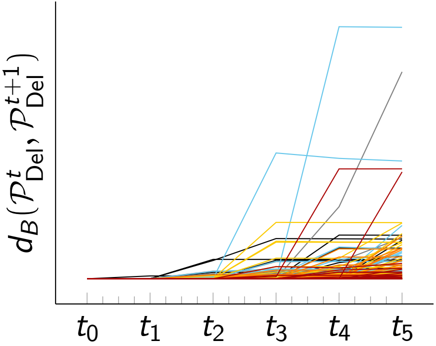
Results

Dataset description

- Investigation of 83 handwritings from different people.
- 45 signatures are recorded per user, which show the same letters.
- Each iteration we double the amount of data points.
- We insert the new points random uniformly within the convex hull.

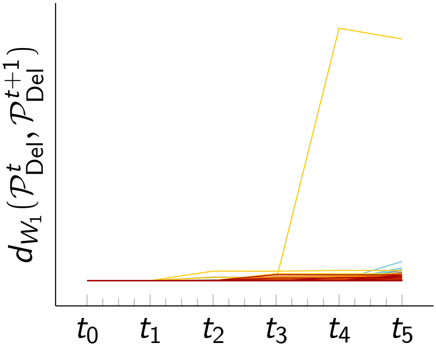


Bottleneck distance over Delaunay PH



In the diagram on the left, all signatures of a user are displayed in the same color. Each line corresponds to a independent signature.
We observe that the signatures are stable in terms of interpolation, since the colours in the diagram are also grouped together.

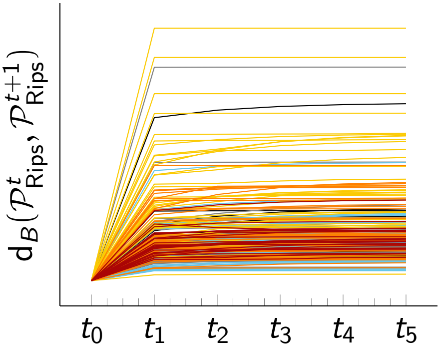
1-Wasserstein distance over Delaunay PH



1-Wasserstein metric for X and Y:

$$\sqrt{\inf_{\varphi : X \to Y} \sum_{x \in X} \|x - \varphi(x)\|_\infty}$$

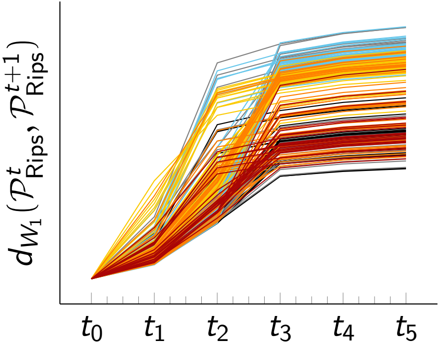
Bottleneck distance over Vietoris-Rips PH



Bottleneck metric:

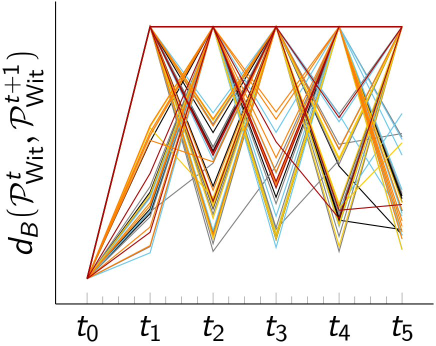
$$\sqrt{\inf_{\varphi: X \to Y} \sum_{x \in X} \|x - \varphi(x)\|_\infty^\infty}$$

1-Wasserstein distance over Vietoris-Rips PH



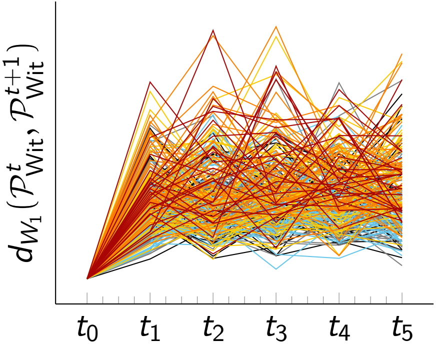
This simplicial complex is closely related to the Čech complex and provides a very accurate description of the topology.

1-Wasserstein distance over Witness PH



The Witness complex is constructed by selecting 5% of the data points as landmarks so that they are uniformly distributed over the set of points.
Since this variant is stochastic in nature and the landmarks are chosen anew for each iteration step, we obtain a zig-zag pattern.
The pattern indicates volatile representatives of homology groups.

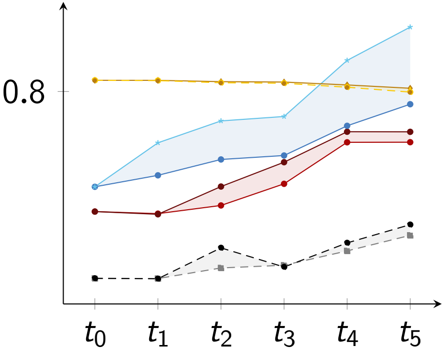
1-Wasserstein distance over Witness PH



The first Wasserstein metric behaves similarly to a 1-norm and is not as coarse as the Bottleneck metric, which behaves like a ∞-norm.

Summary

Summary statistics



With a hypothesis test, calculated using the Wasserstein metric, we have defined a stop criterion. To validate the stop criterion, we calculated elmentary statistics on the whole dataset during interpolation.

Upper curve: no stopping.
Lower curve: with stopping.

- ●: Variation.
- ●: Standard deviation.
- ●: Mean.
- ●: 1-Wasserstein metric.

Summary

What do we have achieved?

- Connection of Collapsing Theorem to Voronoi interpolation.
- A mathematical stopping condition for interpolation techniques.
- A hypothesis test for persistence diagrams.

Open problems?

- What if points do not lie in general position with respect to $\mathbb{R}^n$?
- Which embedding should be used for Voronoi interpolation?
- Does Voronoi interpolation preserve homology groups?

Thank you for your attention!

Got interested?
Drop me a line: luciano.melodia@fau.de.
Let's connect on GitHub: karhunenloeve.