EMAp — Summer School On Data Science — 27/02/23

From Algebraic Topology to Data Analysis

Part III/III: Persistent homology

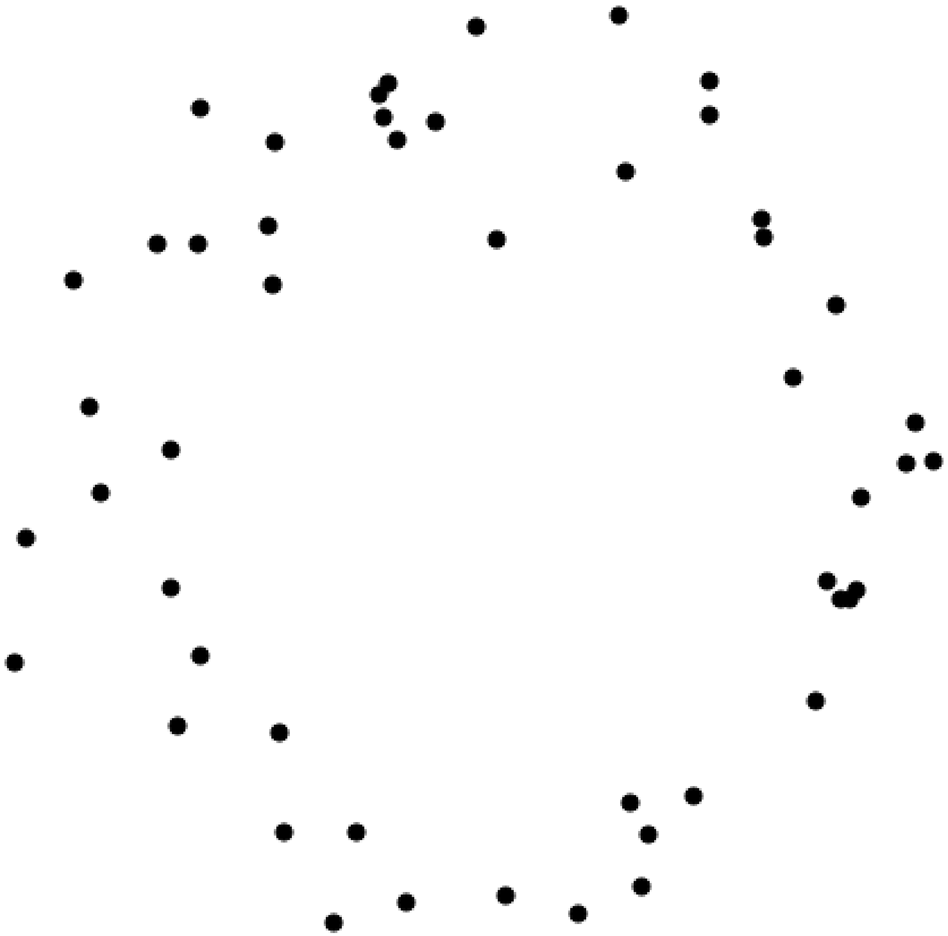
https://raphaeltinarrage.github.io

Last update: January 26, 2023

Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

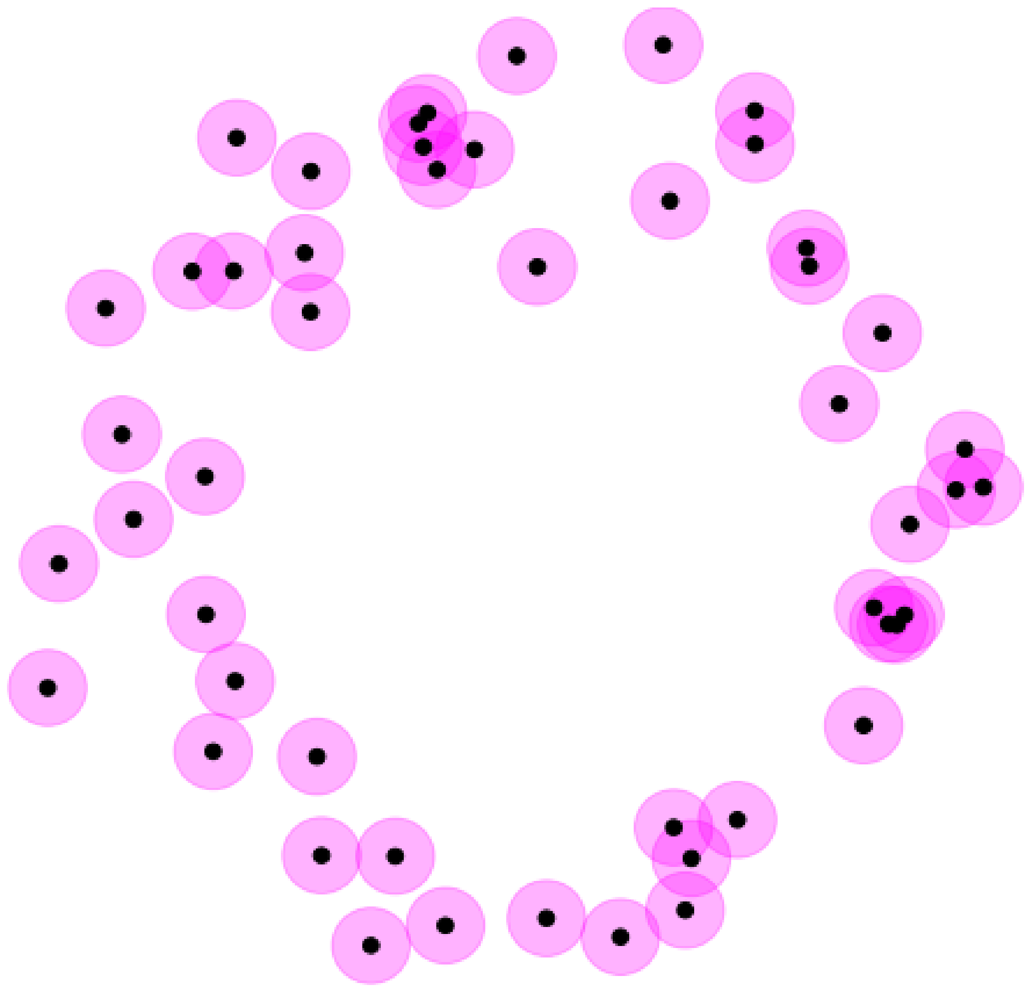
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

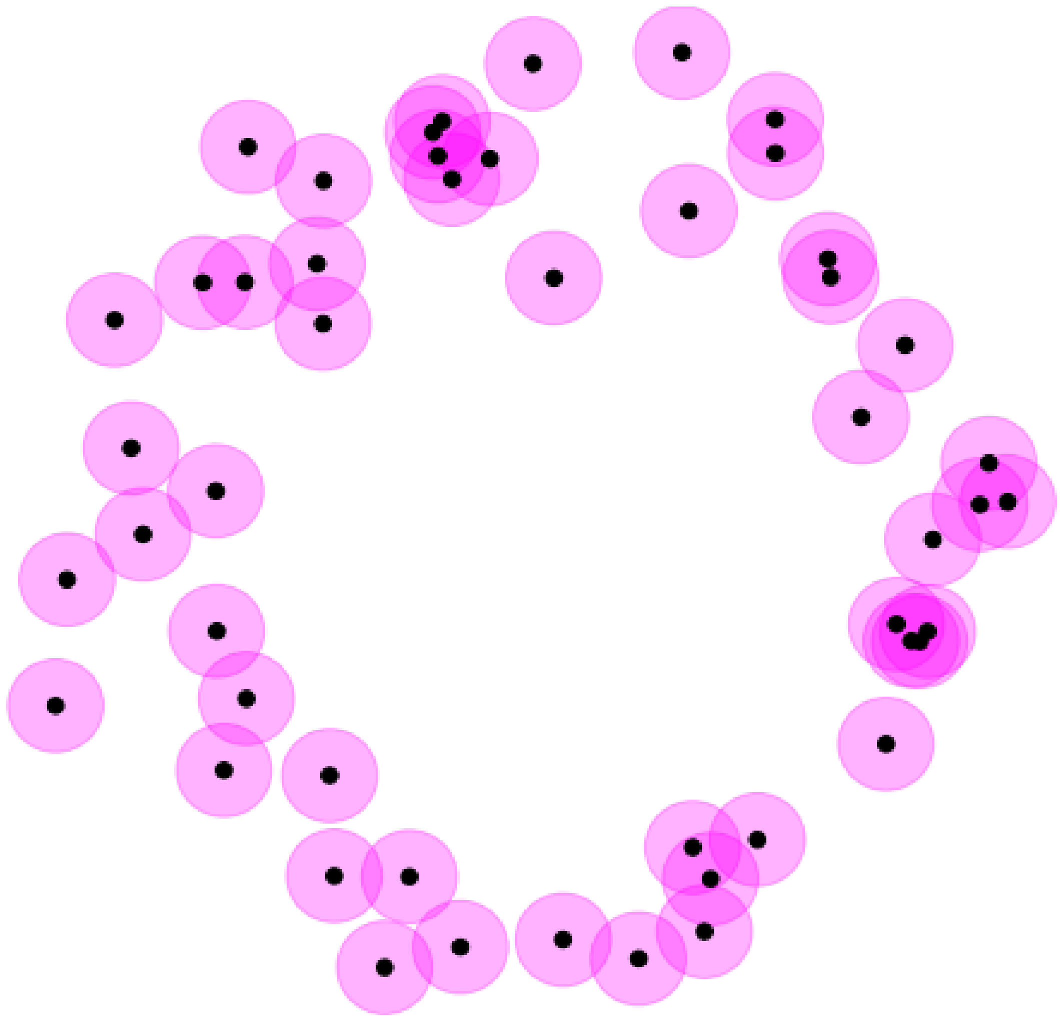
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

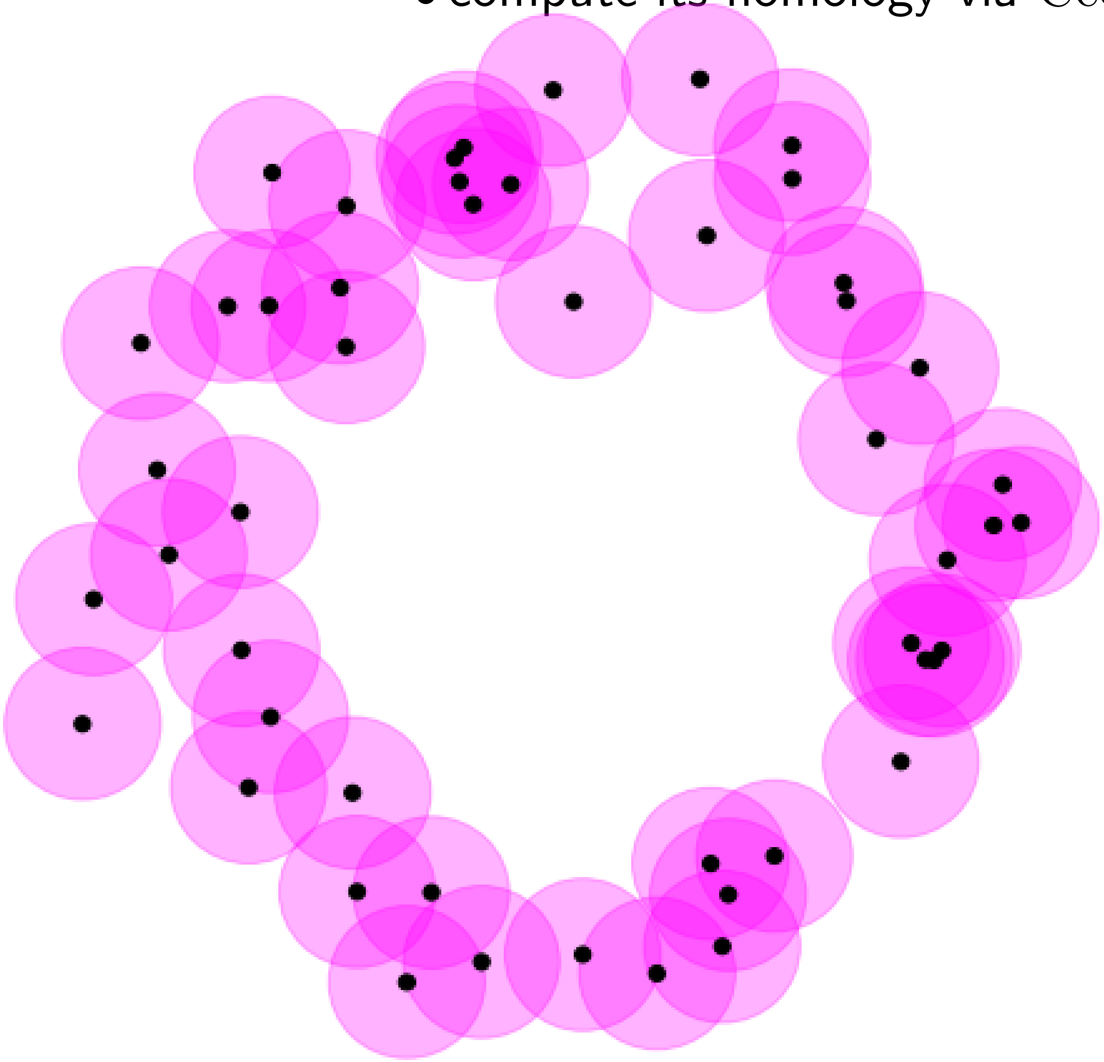
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

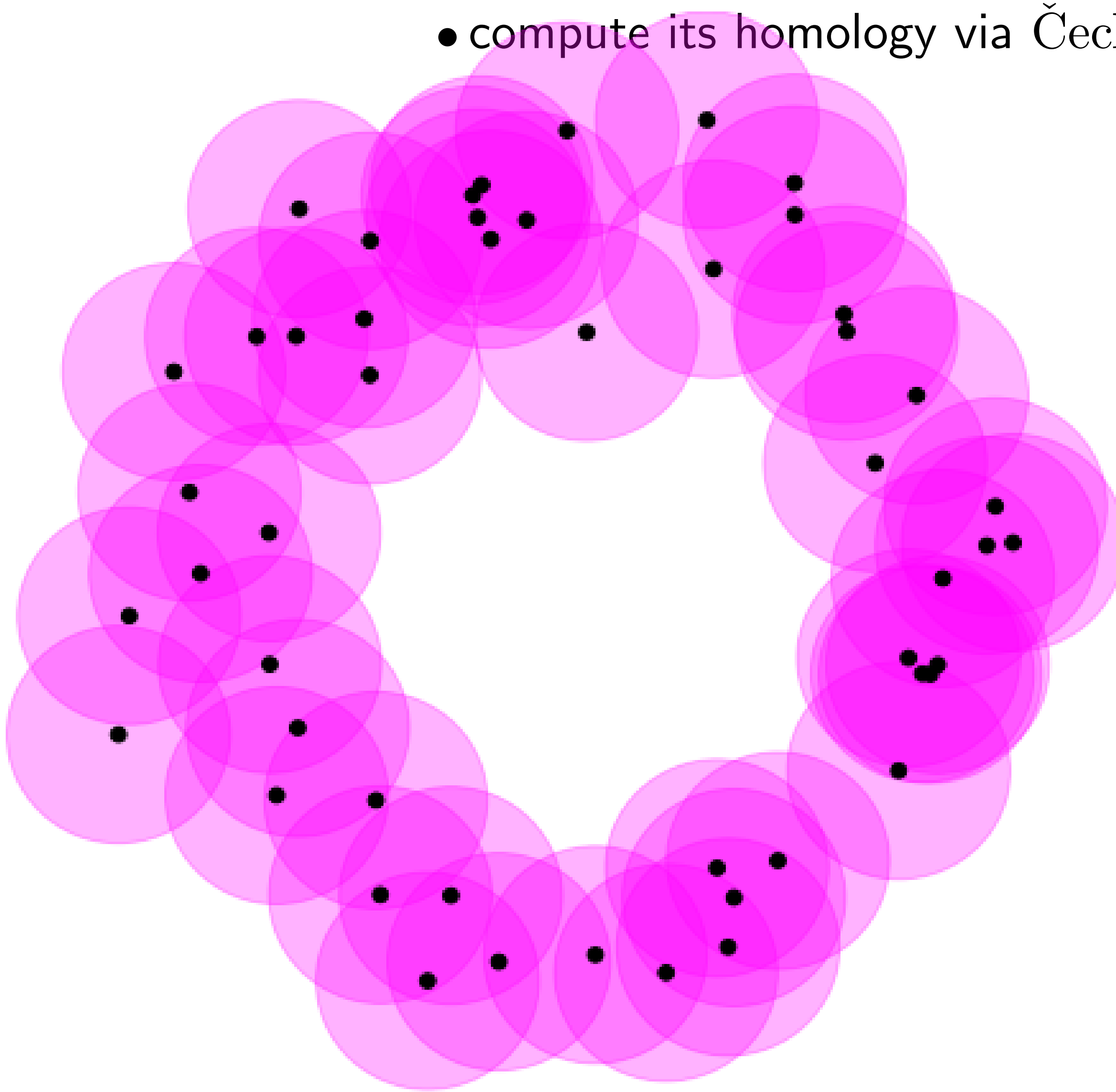
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

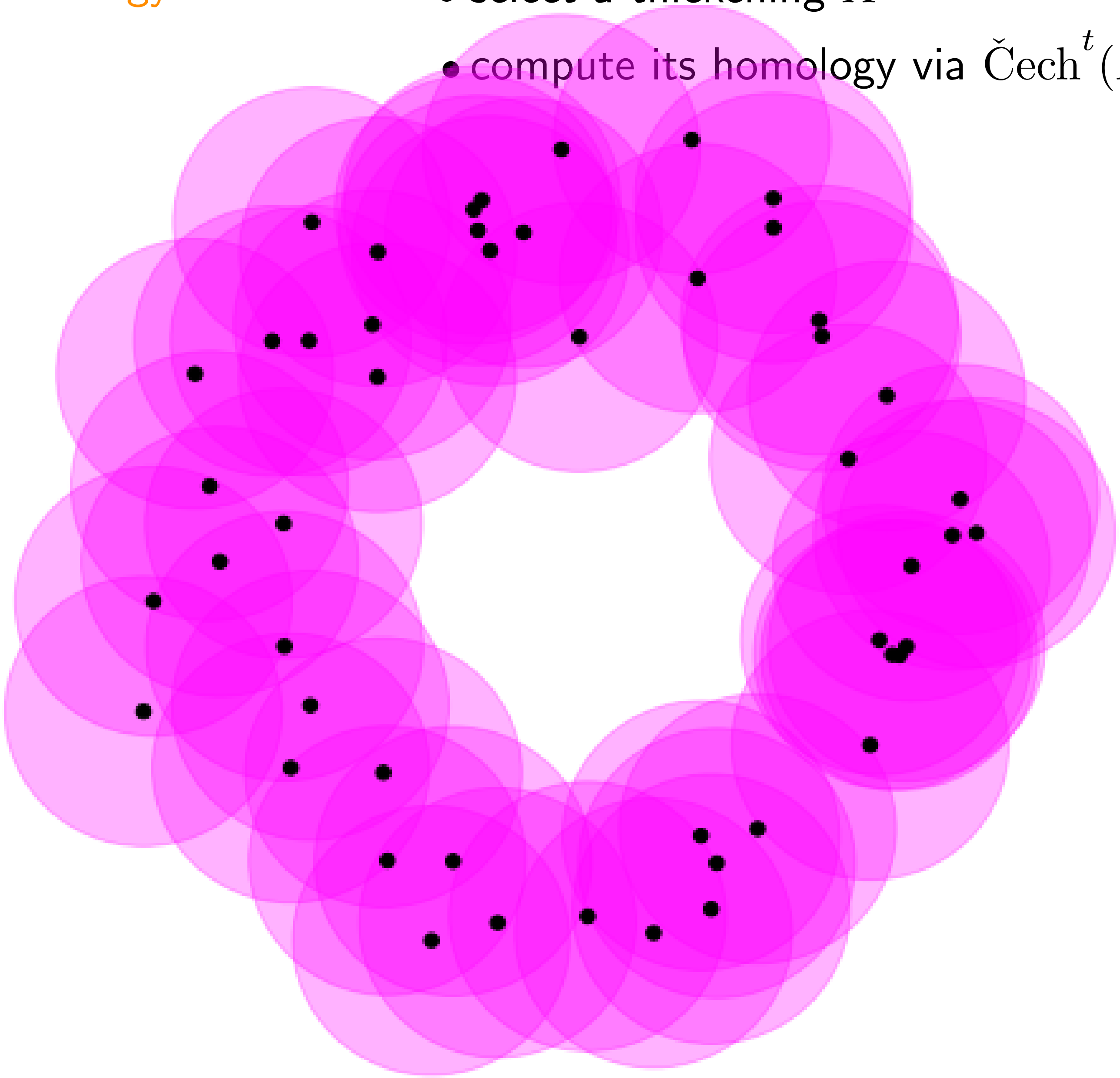
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

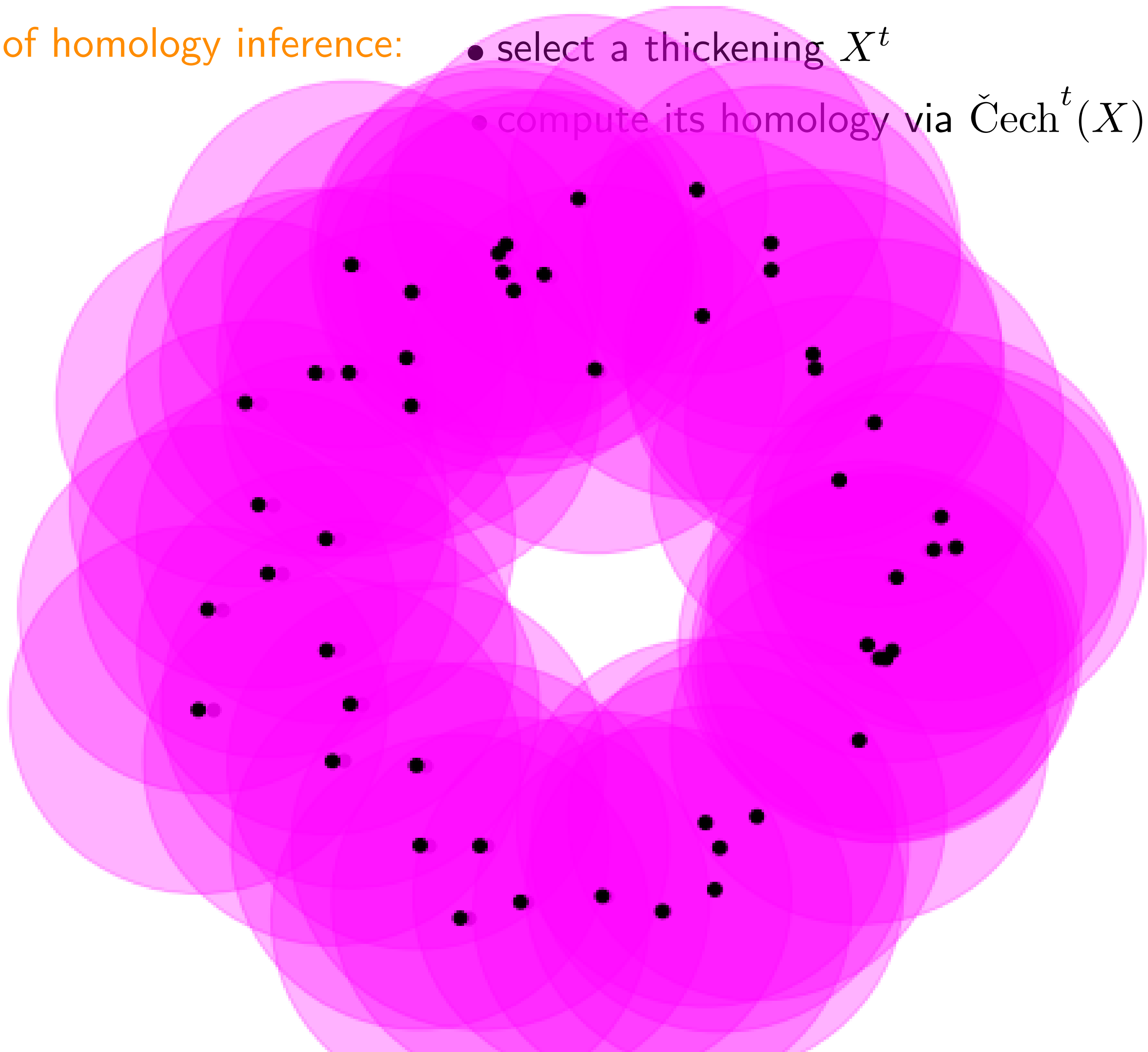
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

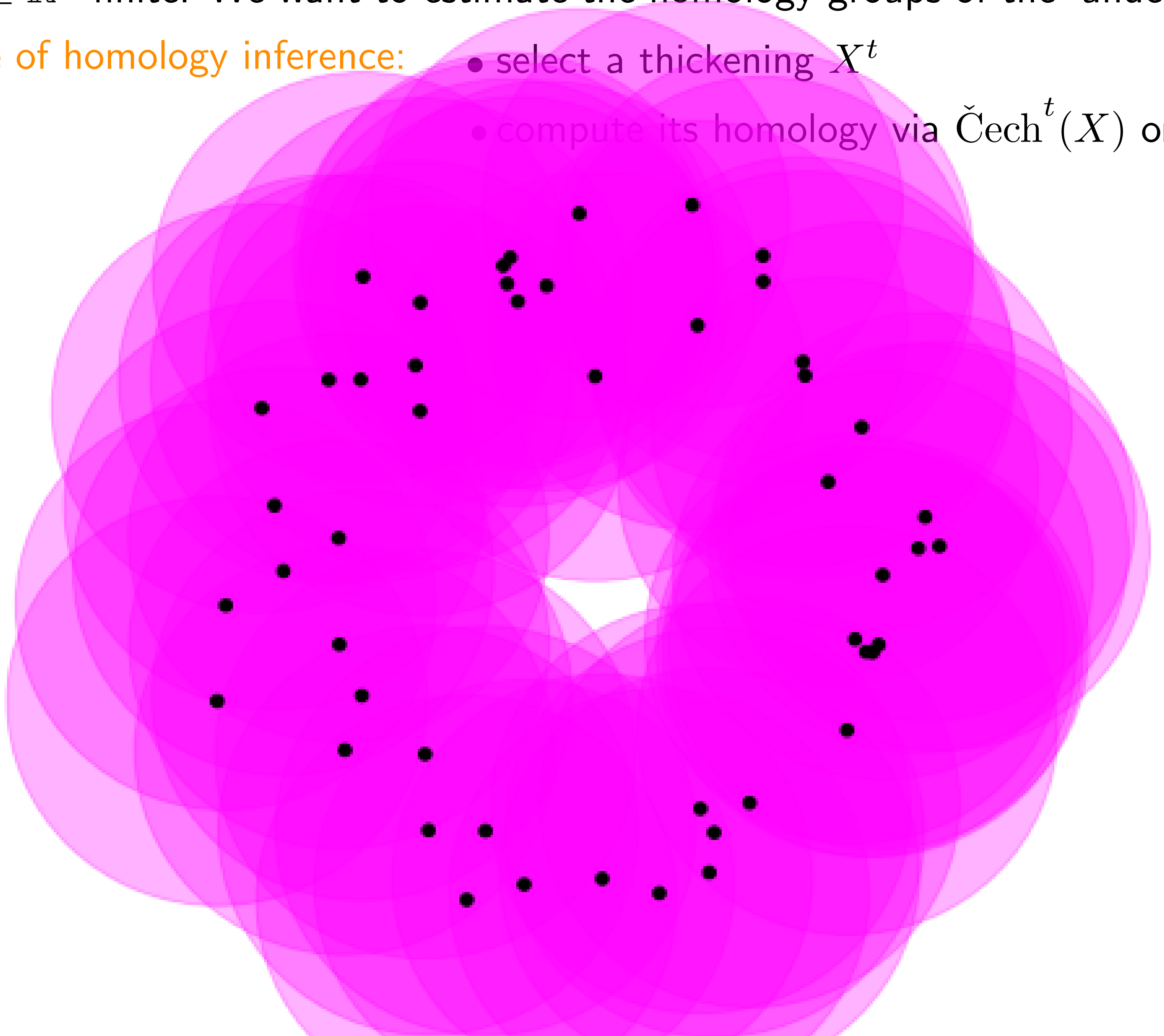
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

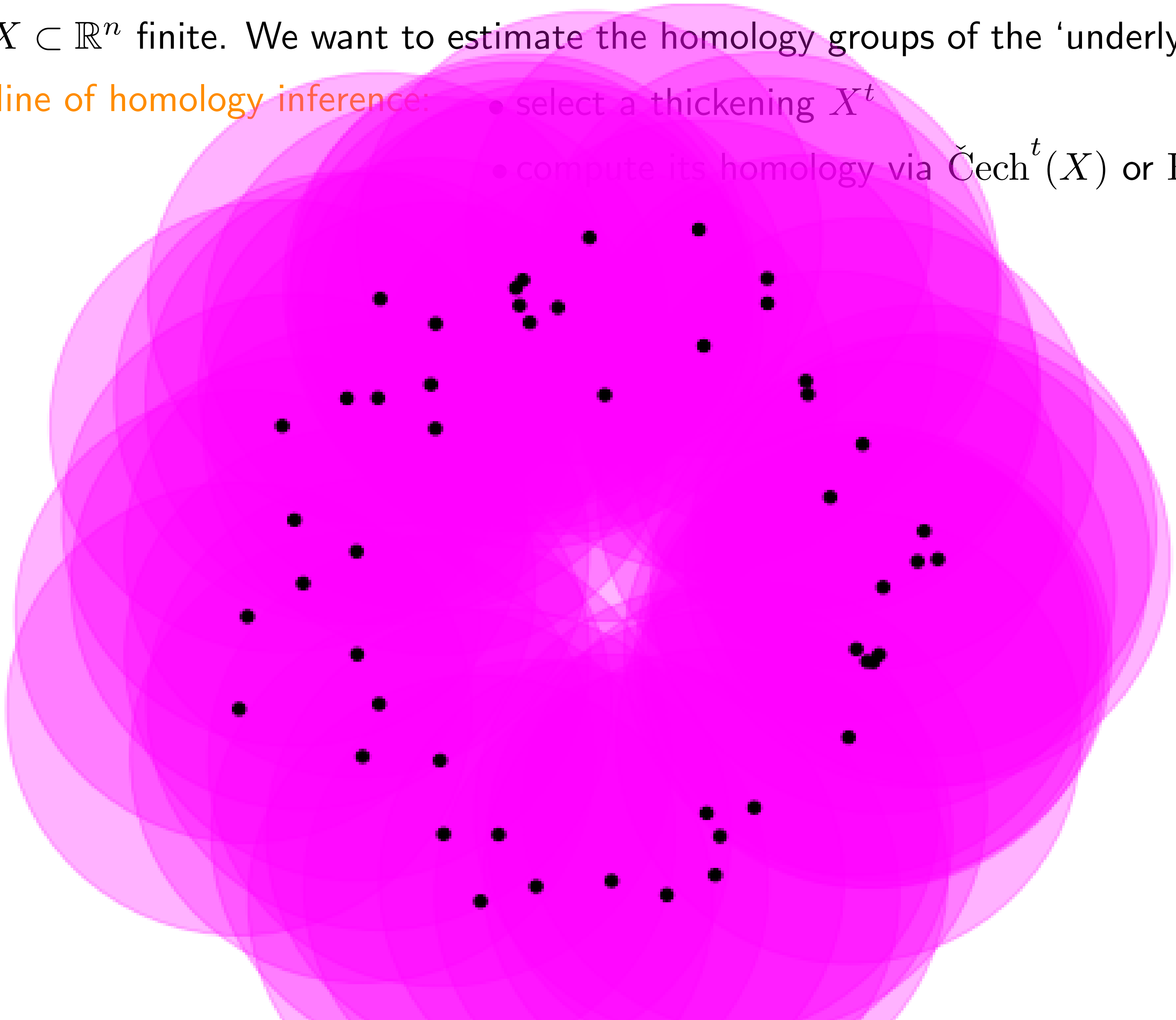
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:

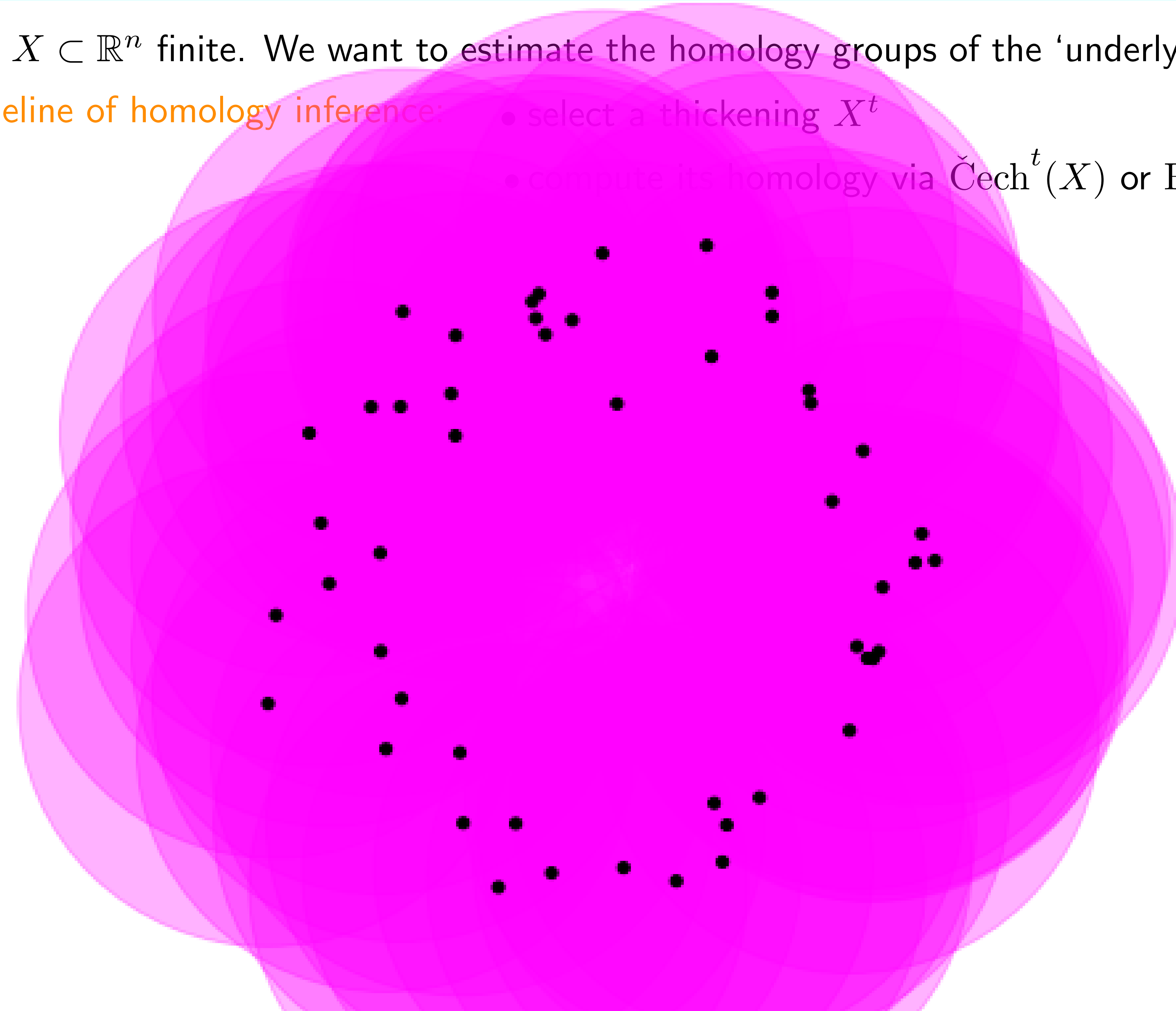
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

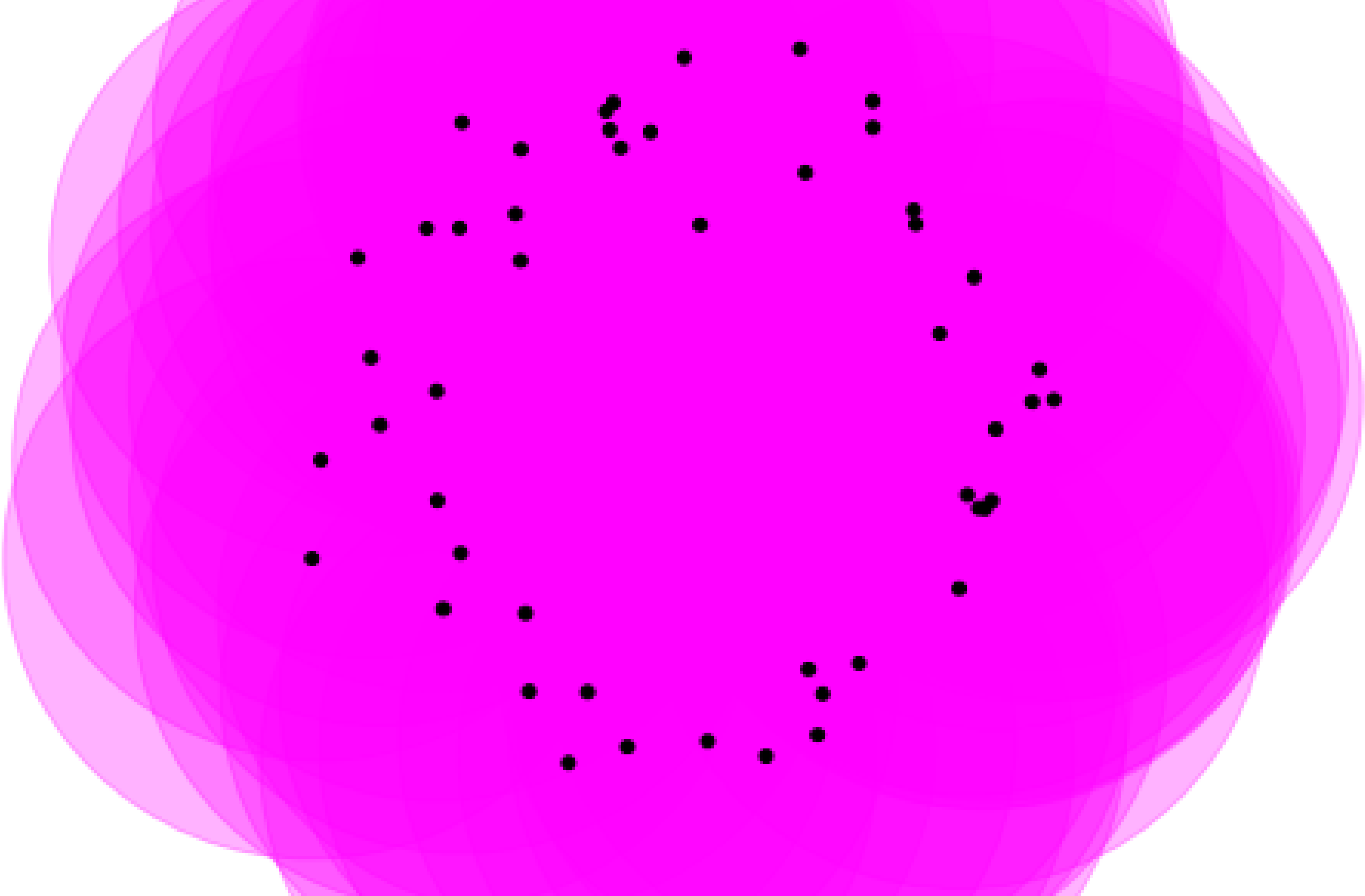
Pipeline of homology inference:

- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



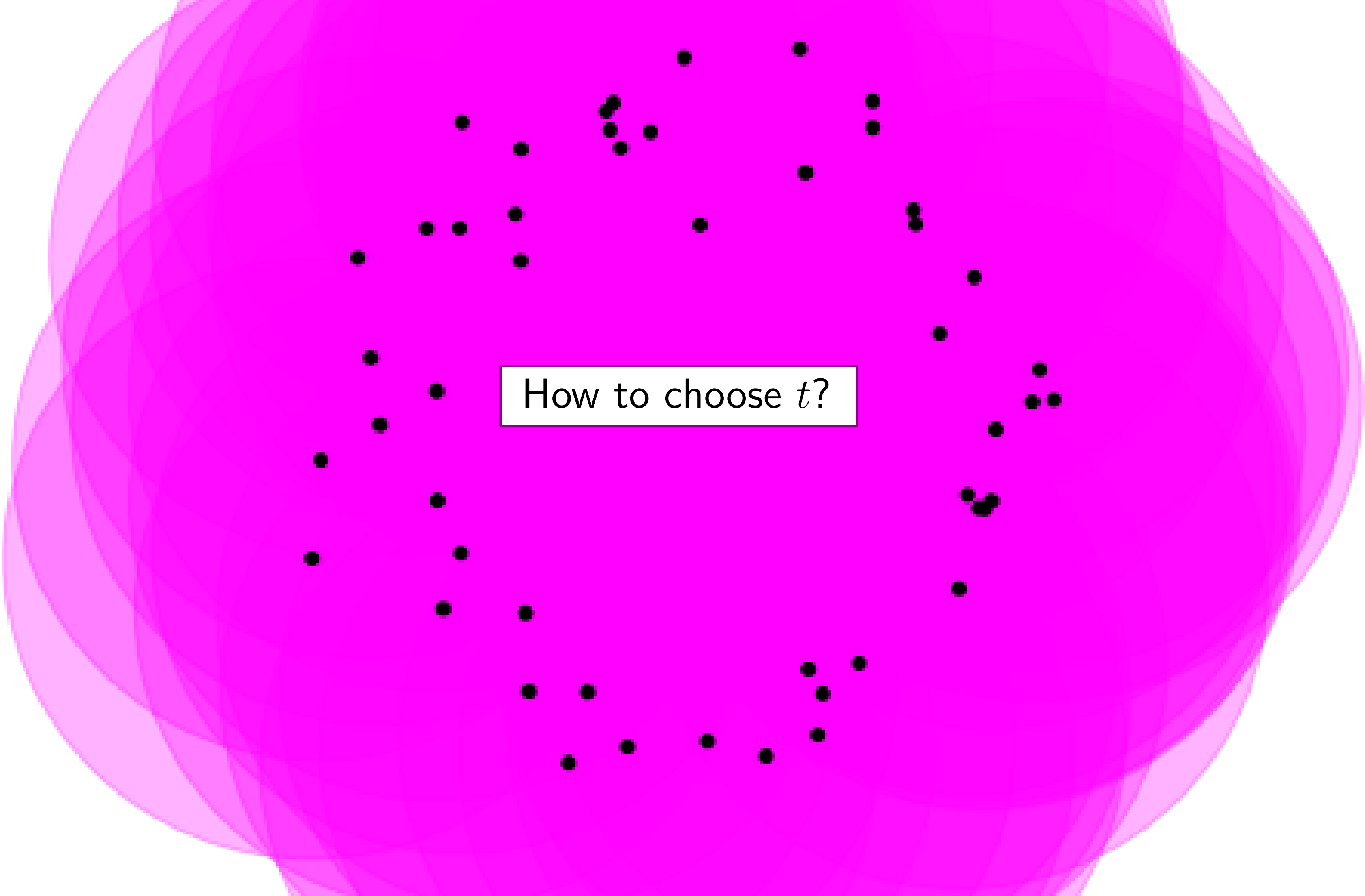
Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$

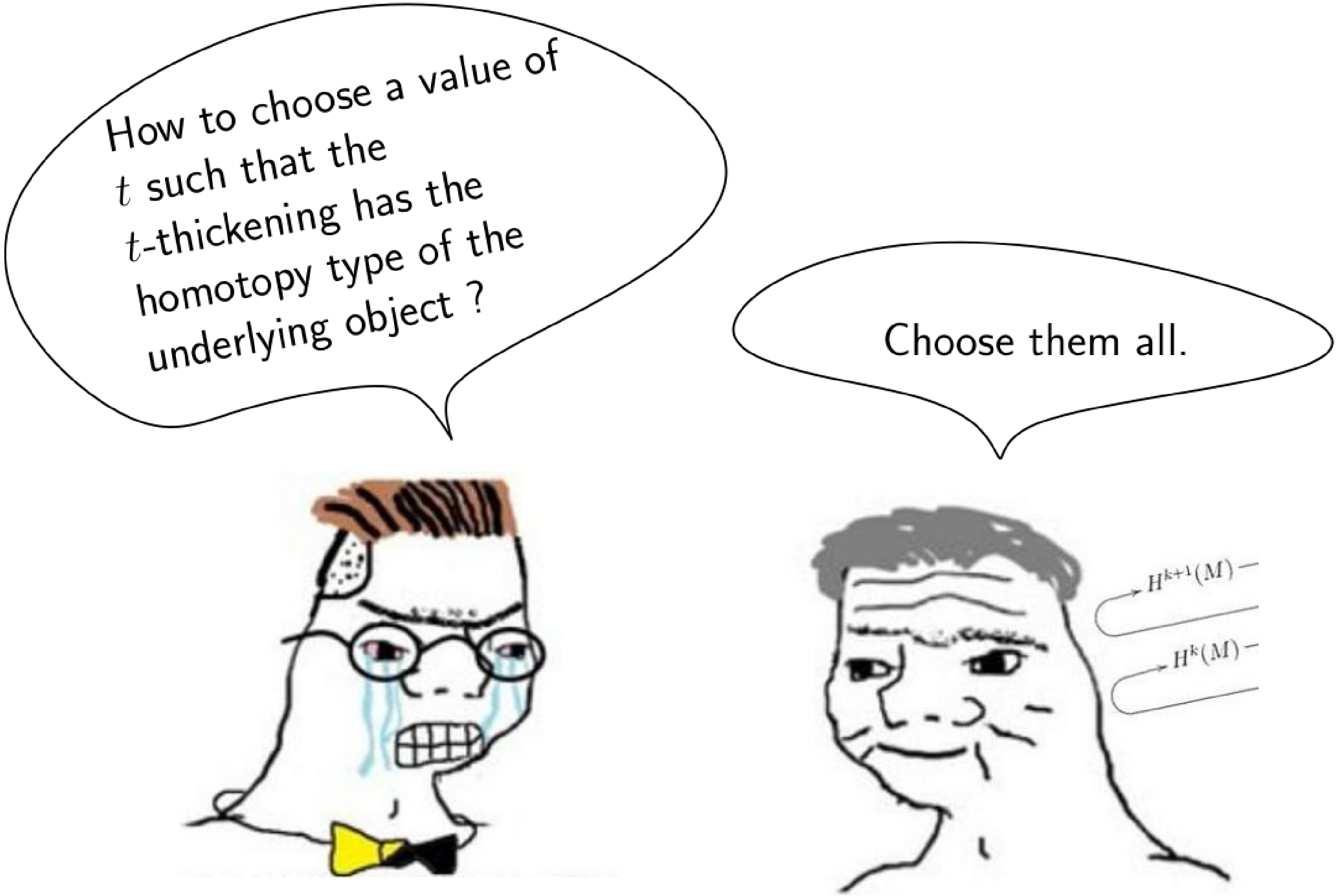


Let $X \subset \mathbb{R}^n$ finite. We want to estimate the homology groups of the 'underlying shape'.

Pipeline of homology inference:
- select a thickening X^t
- compute its homology via $\check{\mathrm{C}}\mathrm{ech}^t(X)$ or $\mathrm{Rips}^t(X)$



O problema da inferência homológica



Data analyst　　　　*Persistence theory*

I - Decomposition of persistence modules

1 - Simplicial filtrations

2 - Persistence modules

3 - Barcodes

II - Stability of persistence modules

1 - Bottleneck distance

2 - Interleaving distance

3 - Stability

III - Persistent homology in practice

1 - Data analysis

2 - Machine learning

3 - Variations on persistent homology

Funtorialidade singular

We have seen that homology transforms *topological spaces* into *vector spaces*

$$\begin{aligned} H_i \colon \mathrm{Top} &\longrightarrow \mathrm{Vect} \\ X &\longmapsto H_i(X) \end{aligned}$$

and transforms *continous maps* into *linear maps*

$$\big(f \colon X \to Y\big) \longmapsto \big(f_* \colon H_i(X) \to H_i(Y)\big)$$

We will adopt a simplicial point of view.

$$\begin{aligned} H_i \colon \mathrm{SimpComp} &\longrightarrow \mathrm{Vect} \\ K &\longmapsto H_i(K) \\ \big(f \colon K \to L\big) &\longmapsto \big(f_* \colon H_i(K) \to H_i(L)\big) \end{aligned}$$

Funtorialidade singular

We have seen that homology transforms *topological spaces* into *vector spaces*

$$\begin{aligned} H_i \colon \mathrm{Top} &\longrightarrow \mathrm{Vect} \\ X &\longmapsto H_i(X) \end{aligned}$$

and transforms *continous maps* into *linear maps*

$$\big(f \colon X \to Y\big) \longmapsto \big(f_* \colon H_i(X) \to H_i(Y)\big)$$

We will adopt a simplicial point of view.

$$\begin{aligned} H_i \colon \mathrm{SimpComp} &\longrightarrow \mathrm{Vect} \\ K &\longmapsto H_i(K) \\ \big(f \colon K \to L\big) &\longmapsto \big(f_* \colon H_i(K) \to H_i(L)\big) \end{aligned}$$

what is a map between simplicial complexes?

Mapas simpliciais

Definition: Let K and L be two simplicial complexes, and V_K, V_L their set of vertices. A **simplicial map** between K and L is a map $f\colon V_K \to V_L$ such that

$$\forall \sigma \in K, f(\sigma) \in L.$$

We may denote a simplicial map $f\colon K \to L$ instead of $f\colon V_K \to V_L$.

Mapas simpliciais

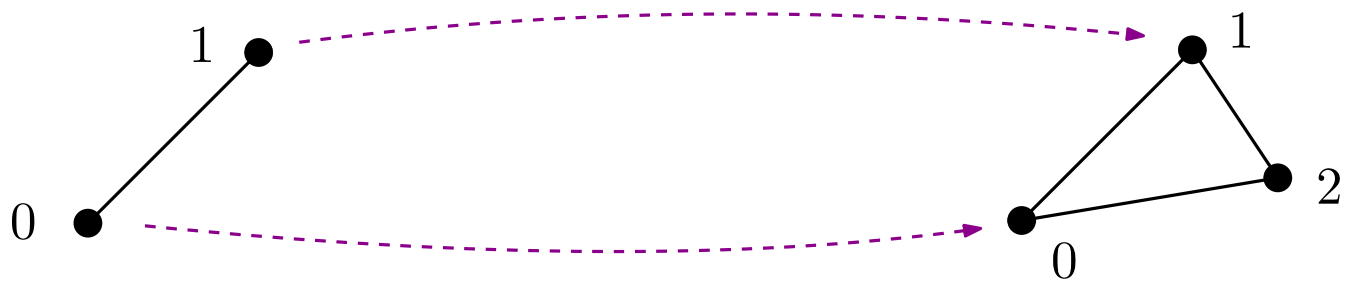
Definition: Let K and L be two simplicial complexes, and V_K, V_L their set of vertices.
A **simplicial map** between K and L is a map $f\colon V_K \to V_L$ such that

$$\forall \sigma \in K, f(\sigma) \in L.$$

We may denote a simplicial map $f\colon K \to L$ instead of $f\colon V_K \to V_L$.

Example: Let $K = \{[0],[1],[0,1]\}$, $L = \{[0],[1],[2],[0,1],[0,2],[1,2]\}$ and

$$f\colon \{0,1\} \to \{0,1,2\}$$
$$0 \mapsto 0$$
$$1 \mapsto 1$$



It is simplicial since $f([0,1]) = [0,1]$ is a simplex of L.

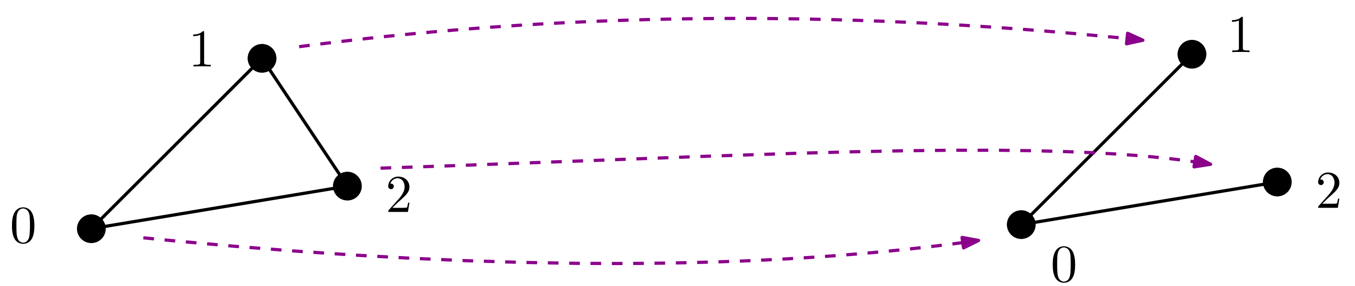
Definition: Let K and L be two simplicial complexes, and V_K, V_L their set of vertices. A **simplicial map** between K and L is a map $f\colon V_K \to V_L$ such that

$$\forall \sigma \in K, f(\sigma) \in L.$$

We may denote a simplicial map $f\colon K \to L$ instead of $f\colon V_K \to V_L$.

Example: Let $K = \{[0],[1],[2],[0,1],[0,2],[1,2]\}$, $L = \{[0],[1],[2],[0,1],[0,2]\}$ and

$$f\colon \{0,1\} \to \{0,1,2\}$$
$$0 \mapsto 0$$
$$1 \mapsto 1$$
$$2 \mapsto 2$$



It is not simplicial since $f([1,2]) = [1,2]$ is not a simplex of L.

Mapas simpliciais

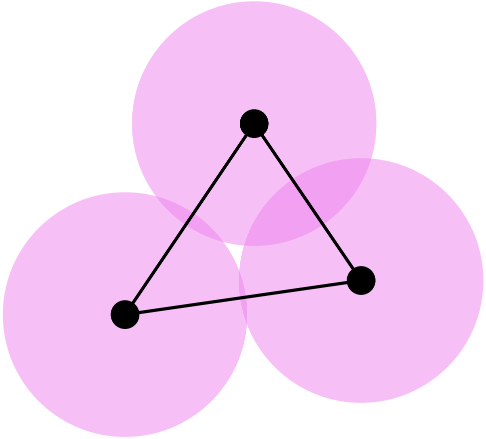
Definition: Let K and L be two simplicial complexes, and V_K, V_L their set of vertices. A **simplicial map** between K and L is a map $f\colon V_K \to V_L$ such that

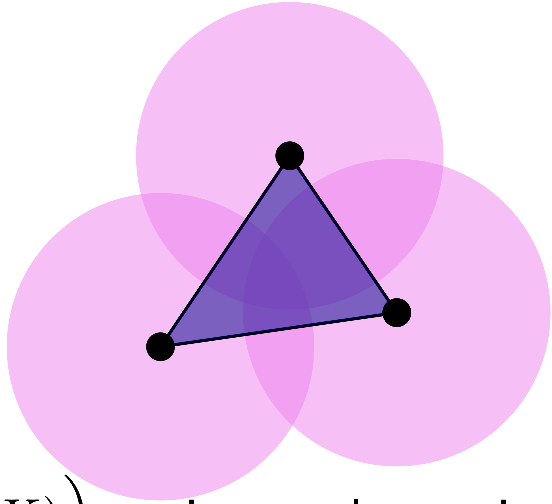
$$\forall \sigma \in K, f(\sigma) \in L.$$

We may denote a simplicial map $f\colon K \to L$ instead of $f\colon V_K \to V_L$.

Example: Let $X \subset \mathbb{R}^n$ and $s, t \geq 0$ such that $s \leq t$. Consider the Čech complexes $\check{\mathrm{C}}\mathrm{ech}^s(X)$ and $\check{\mathrm{C}}\mathrm{ech}^t(X)$.

The inclusion map $i\colon \check{\mathrm{C}}\mathrm{ech}^s(X) \to \check{\mathrm{C}}\mathrm{ech}^t(X)$ is a simplicial map.





Indeed, the sequence of simplicial complexes $\left(\check{\mathrm{C}}\mathrm{ech}^t(X)\right)_{t\geq 0}$ is non-decreasing.

Mapa linear induzido

Let $f\colon K \to L$ be a simplicial map. Let $n \geq 0$, and consider the groups of chains of K and L:

$$C_n(K) = \left\{ \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \mid \forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z} \right\}$$

$$C_n(L) = \left\{ \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \mid \forall \sigma \in L_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z} \right\}$$

We define a linear map as follows:

$$\begin{aligned} f_n\colon C_n(K) &\longrightarrow C_n(L) \\ \sigma &\longmapsto f(\sigma) \quad \text{if} \quad \dim(f(\sigma)) = n, \\ & \qquad 0 \quad \text{else.} \end{aligned}$$

Mapa linear induzido

Let $f\colon K \to L$ be a simplicial map. Let $n \geq 0$, and consider the groups of chains of K and L:

$$C_n(K) = \left\{ \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \mid \forall \sigma \in K_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z} \right\}$$

$$C_n(L) = \left\{ \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \mid \forall \sigma \in L_{(n)},\ \epsilon_\sigma \in \mathbb{Z}/2\mathbb{Z} \right\}$$

We define a linear map as follows:

$$\begin{aligned} f_n\colon C_n(K) &\longrightarrow C_n(L) \\ \sigma &\longmapsto f(\sigma) \quad \text{if} \quad \dim(f(\sigma)) = n, \\ & \qquad 0 \quad \text{else.} \end{aligned}$$

$$\begin{array}{ccccccccccc} \dashrightarrow & C_3(K) & \xrightarrow{\partial_3} & C_2(K) & \xrightarrow{\partial_2} & C_1(K) & \xrightarrow{\partial_1} & C_0(K) & \xrightarrow{\partial_0} & \{0\} \\ & \downarrow{\scriptstyle f_3} & & \downarrow{\scriptstyle f_2} & & \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle f_0} & & \downarrow{\scriptstyle f_{-1}} \\ \dashrightarrow & C_3(L) & \xrightarrow{\partial_3} & C_2(L) & \xrightarrow{\partial_2} & C_1(L) & \xrightarrow{\partial_1} & C_0(L) & \xrightarrow{\partial_0} & \{0\} \end{array}$$

Mapa linear induzido

$$
\begin{array}{ccccccccccc}
\dashrightarrow & C_3(K) & \xrightarrow{\partial_3} & C_2(K) & \xrightarrow{\partial_2} & C_1(K) & \xrightarrow{\partial_1} & C_0(K) & \xrightarrow{\partial_0} & \{0\} \\
& \downarrow f_3 & & \downarrow f_2 & & \downarrow f_1 & & \downarrow f_0 & & \downarrow f_{-1} \\
\dashrightarrow & C_3(L) & \xrightarrow{\partial_3} & C_2(L) & \xrightarrow{\partial_2} & C_1(L) & \xrightarrow{\partial_1} & C_0(L) & \xrightarrow{\partial_0} & \{0\}
\end{array}
$$

Lemma: For every $n \geq 0$, we have $\partial_n \circ f_n = f_{n-1} \circ \partial_n$.

Proof: Let $\sigma \in K_{(n)}$. We have the equalities

$$
\partial_n \circ f_n(\sigma) = \sum_{\substack{\mu \subset f(\sigma) \\ |\mu| = |\sigma| - 1}} \mu
$$

$$
f_{n-1} \circ \partial_n(\sigma) = \sum_{\substack{\tau \subset \sigma \\ |\tau| = |\sigma| - 1}} f_n(\tau)
$$

We should distinguish three cases:

- $|f(\sigma)| = |\sigma|$ (i.e. f is injective on σ),
- $|f(\sigma)| < |\sigma| - 1$,
- $|f(\sigma)| = |\sigma| - 1$.

Mapa linear induzido

$$
\begin{array}{ccccccccccc}
\dashrightarrow & C_3(K) & \xrightarrow{\partial_3} & C_2(K) & \xrightarrow{\partial_2} & C_1(K) & \xrightarrow{\partial_1} & C_0(K) & \xrightarrow{\partial_0} & \{0\} \\
 & \downarrow f_3 & & \downarrow f_2 & & \downarrow f_1 & & \downarrow f_0 & & \downarrow f_{-1} \\
\dashrightarrow & C_3(L) & \xrightarrow{\partial_3} & C_2(L) & \xrightarrow{\partial_2} & C_1(L) & \xrightarrow{\partial_1} & C_0(L) & \xrightarrow{\partial_0} & \{0\}
\end{array}
$$

Lemma: For every $n \geq 0$, we have $\partial_n \circ f_n = f_{n-1} \circ \partial_n$.

Proposition: For every $c \in Z_n(K)$, we have $f_n(c) \in Z_n(L)$.
For every $c \in B_n(K)$, we also have $f_n(c) \in B_n(L)$.

Proof: First, let $c \in Z_n(K)$. We have

$$\partial_n \circ f_n(c) = f_{n-1} \circ \partial_n(c) = f_{n-1}(0) = 0,$$

hence $f_n(c) \in Z_n(L)$.

Secondly, let $c \in B_n(K)$, and write $c = \partial_{n+1}(c')$ with $c' \in C_{n+1}(K)$. We get

$$f_n(c) = f_n \circ \partial_{n+1}(c') = \partial_{n+1} \circ f_{n+1}(c'),$$

hence $f_n(c) \in B_n(L)$.

Mapa linear induzido

$$
\begin{array}{ccccccccccc}
\dashrightarrow & C_3(K) & \xrightarrow{\partial_3} & C_2(K) & \xrightarrow{\partial_2} & C_1(K) & \xrightarrow{\partial_1} & C_0(K) & \xrightarrow{\partial_0} & \{0\} \\
 & \downarrow f_3 & & \downarrow f_2 & & \downarrow f_1 & & \downarrow f_0 & & \downarrow f_{-1} \\
\dashrightarrow & C_3(L) & \xrightarrow{\partial_3} & C_2(L) & \xrightarrow{\partial_2} & C_1(L) & \xrightarrow{\partial_1} & C_0(L) & \xrightarrow{\partial_0} & \{0\}
\end{array}
$$

Lemma: For every $n \geq 0$, we have $\partial_n \circ f_n = f_{n-1} \circ \partial_n$.

Proposition: For every $c \in Z_n(K)$, we have $f_n(c) \in Z_n(L)$.
For every $c \in B_n(K)$, we also have $f_n(c) \in B_n(L)$.

We have $B_n(K) \subset Z_n(K)$, $B_n(L) \subset Z_n(L)$, $f(Z_n(K)) \subset f(Z_n(K))$ and $f(B_n(K)) \subset f(B_n(K))$.
Hence we can define a linear map between quotient vector spaces:

$$(f_n)_* \colon Z_n(K)/B_n(K) \longrightarrow Z_n(L)/B_n(L).$$

By definition of the homology groups, we have defined a map

$$(f_n)_* \colon H_n(K) \longrightarrow H_n(L).$$

It is called the **induced map in homology**.

Mapa linear induzido

$$\begin{array}{ccccccccccc}
\dashrightarrow & C_3(K) & \xrightarrow{\partial_3} & C_2(K) & \xrightarrow{\partial_2} & C_1(K) & \xrightarrow{\partial_1} & C_0(K) & \xrightarrow{\partial_0} & \{0\} \\
 & \downarrow f_3 & & \downarrow f_2 & & \downarrow f_1 & & \downarrow f_0 & & \downarrow f_{-1} \\
\dashrightarrow & C_3(L) & \xrightarrow{\partial_3} & C_2(L) & \xrightarrow{\partial_2} & C_1(L) & \xrightarrow{\partial_1} & C_0(L) & \xrightarrow{\partial_0} & \{0\}
\end{array}$$

Lemma: For every $n \geq 0$, we have $\partial_n \circ f_n = f_{n-1} \circ \partial_n$.

Proposition: For every $c \in Z_n(K)$, we have $f_n(c) \in Z_n(L)$.
For every $c \in B_n(K)$, we also have $f_n(c) \in B_n(L)$.

$$\begin{array}{ccccc}
\cdots & H_3(K) & H_2(K) & H_1(K) & H_0(K) \\
 & \downarrow (f_3)_* & \downarrow (f_2)_* & \downarrow (f_1)_* & \downarrow (f_0)_* \\
\cdots & H_3(L) & H_2(L) & H_1(L) & H_0(L)
\end{array}$$

$(f_n)_*$ can be defined as

$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \longmapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma)$$

Mapa linear induzido

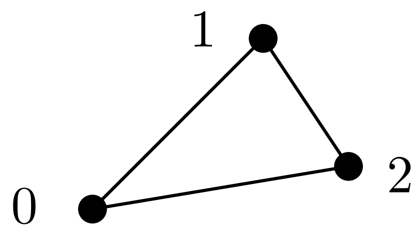
Example: Consider the simplicial complexes $K = L = \{[0], [1], [2], [0,1], [0,2], [1,2]\}$.

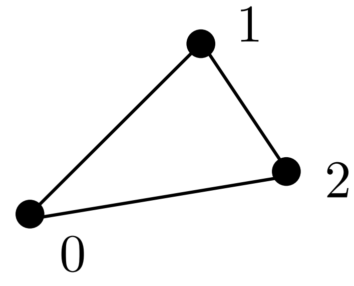
The inclusion $i\colon K \to L$ induces the identity in H_0:

$$(i_0)_*\colon H_0(K) \simeq \mathbb{Z}/2\mathbb{Z} \longrightarrow H_0(L) \simeq \mathbb{Z}/2\mathbb{Z}$$
$$1 \longmapsto 1$$

The inclusion $i\colon K \to L$ induces the identity in H_1:

$$(i_1)_*\colon H_1(K) \simeq \mathbb{Z}/2\mathbb{Z} \longrightarrow H_1(L) \simeq \mathbb{Z}/2\mathbb{Z}$$
$$1 \longmapsto 1$$





$(f_n)_*$ can be defined as

$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \longmapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma)$$

Mapa linear induzido

Example: Consider the simplicial complexes $K = \{[0],[1],[2],[0,1],[0,2],[1,2]\}$ and $L = \{[0],[1],[2],[0,1],[0,2],[1,2],[0,1,2]\}$.

The inclusion $i\colon K \to L$ induces the zero map in H^1:

$$(i_1)_*\colon H_1(K) \simeq \mathbb{Z}/2\mathbb{Z} \longrightarrow H_1(L) \simeq \{0\}$$
$$1 \longmapsto 0$$

$(f_n)_*$ can be defined as

$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \longmapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma)$$

Mapa linear induzido

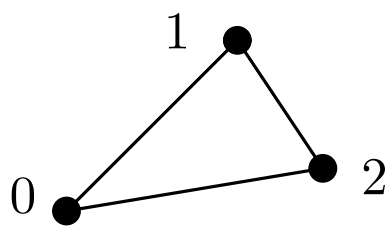
Example: Consider the simplicial complexes $K = \{[0],[1],[2],[0,1],[0,2],[1,2]\}$ and $L = \{[0],[1],[2],[3],[0,1],[0,2],[1,2],[1,3],[2,3]\}$.

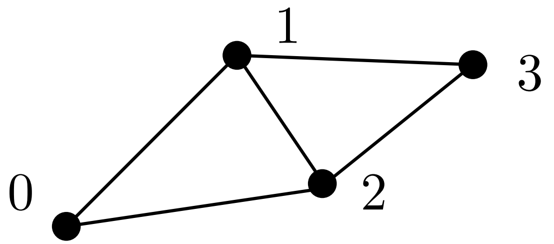
The homology group $H_1(L)$ is isomorphic to the vector space $(\mathbb{Z}/2\mathbb{Z})^2$ by identifying $[0,1]+[0,2]+[1,2] \mapsto (1,0)$ and $[1,2]+[2,3]+[1,3] \mapsto (0,1)$.

The inclusion $i\colon K \to L$ induces the following map between 1^{st} homology groups:

$$(i_1)_*\colon H_1(K) \simeq \mathbb{Z}/2\mathbb{Z} \longrightarrow H_1(L) \simeq (\mathbb{Z}/2\mathbb{Z})^2$$
$$1 \longmapsto (1,0)$$

It can be represented as the matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.





$(f_n)_*$ can be defined as

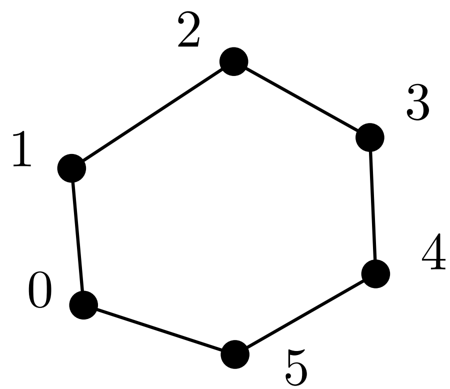
$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \longmapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma)$$

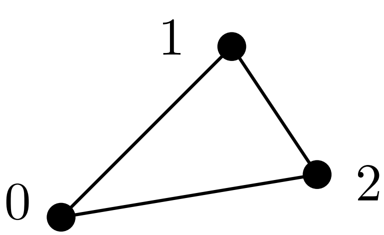
Mapa linear induzido

Example: Let $K = \{[0],[1],[2],[3],[4],[5],[0,1],[1,2],[2,3],[3,4],[4,5],[5,0]\}$ and $L = \{[0],[1],[2],[0,1],[1,2],[2,0]\}$.

Consider the simplical map $f\colon i \mapsto i$ (modulo 3).

The induced map $(f_1)_*$ is zero.





$(f_n)_*$ can be defined as

$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \longmapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma)$$

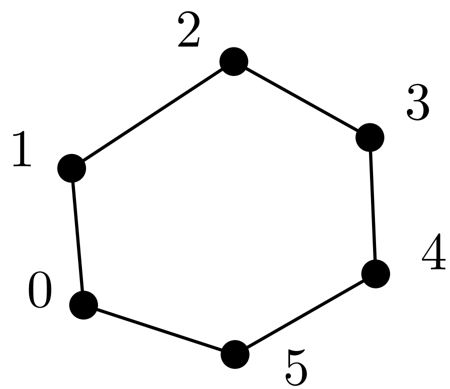
Mapa linear induzido

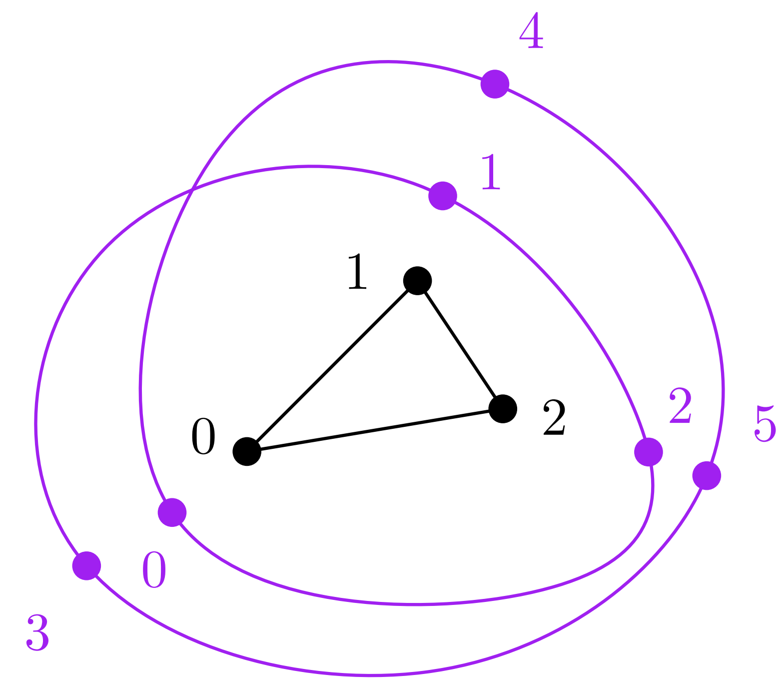
Example: Let $K = \{[0],[1],[2],[3],[4],[5],[0,1],[1,2],[2,3],[3,4],[4,5],[5,0]\}$ and $L = \{[0],[1],[2],[0,1],[1,2],[2,0]\}$.

Consider the simplical map $f\colon i \mapsto i$ (modulo 3).

The induced map $(f_1)_*$ is zero.

In $\mathbb{Z}/2\mathbb{Z}$, we have $2 = 0$





$(f_n)_*$ can be defined as

$$c = \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot \sigma \longmapsto \sum_{\sigma \in K_{(n)}} \epsilon_\sigma \cdot f_n(\sigma)$$

Funtorialidade simplicial

Proposition: Let K, L, M be three simplicial complexes, and consider two simplicial maps $f\colon K \to L$ and $g\colon L \to M$.

For any $n \geq 0$, the induced map $((g \circ f)_n)_*\colon H_n(K) \to H_n(M)$ and $(g_n)_* \circ (f_n)_*\colon H_n(K) \to H_n(M)$ are equal.

$$K \xrightarrow[f]{} L \xrightarrow[g]{} M, \qquad K \xrightarrow{g \circ f} M$$

$$H_n(K) \xrightarrow[f_*]{} H_n(L) \xrightarrow[g_*]{} H_n(M). \qquad H_n(K) \xrightarrow{(g \circ f)_*} H_n(M)$$

Funtorialidade simplicial

Proposition: Let K, L, M be three simplicial complexes, and consider two simplicial maps $f\colon K \to L$ and $g\colon L \to M$.

For any $n \geq 0$, the induced map $((g \circ f)_n)_*\colon H_n(K) \to H_n(M)$ and $(g_n)_* \circ (f_n)_*\colon H_n(K) \to H_n(M)$ are equal.

$$K \xrightarrow[f]{} L \xrightarrow[g]{} M, \qquad (g\circ f\colon K \to M)$$

$$H_n(K) \xrightarrow[f_*]{} H_n(L) \xrightarrow[g_*]{} H_n(M). \qquad ((g\circ f)_*\colon H_n(K) \to H_n(M))$$

Proof: Let $\sigma \in K_{(n)}$. The image $(g \circ f)_n(\sigma)$ is

- $(g \circ f)(\sigma)$ if $g \circ f$ is injective on σ,
- 0 else.

If $g \circ f$ is injective on σ, then f is injective on σ **and** g is injective on $f(\sigma)$, hence $g_n \circ f_n(\sigma) = g \circ f(\sigma)$, and we deduce the result.

If $g \circ f$ is not injective on σ, then f is not injective on σ **or** g is not injective on $f(\sigma)$, hence $g_n \circ f_n(\sigma) = 0$, and we deduce the result.

Acompanhar os ciclos ao longo do tempo

Let $X \subset \mathbb{R}^n$. The collection its thickenings is an non-decreasing **sequence of subsets**

$$\ldots \subset X^{t_1} \subset X^{t_2} \subset X^{t_3} \subset \ldots$$

By considering the corresponding Čech complexes, we obtain an non-decreasing **sequence of simplicial complexes**

$$\ldots \subset \check{\mathrm{C}}\mathrm{ech}^{t_1}(X) \subset \check{\mathrm{C}}\mathrm{ech}^{t_2}(X) \subset \check{\mathrm{C}}\mathrm{ech}^{t_3}(X) \subset \ldots$$

Let us denote i_s^t the inclusion map corresponding to $\check{\mathrm{C}}\mathrm{ech}^s(X) \subset \check{\mathrm{C}}\mathrm{ech}^t(X)$. We can write

$$\dashrightarrow \check{\mathrm{C}}\mathrm{ech}^{t_1}(X) \xrightarrow{i_{t_1}^{t_2}} \check{\mathrm{C}}\mathrm{ech}^{t_2}(X) \xrightarrow{i_{t_2}^{t_3}} \check{\mathrm{C}}\mathrm{ech}^{t_3}(X) \dashrightarrow$$

Applying the i^{th} homology functor yields a **diagram of vector spaces**

$$\dashrightarrow H_i(\check{\mathrm{C}}\mathrm{ech}^{t_1}(X)) \xrightarrow{(i_{t_1}^{t_2})_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_2}(X)) \xrightarrow{(i_{t_2}^{t_3})_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_3}(X)) \dashrightarrow$$

where the maps $(i_s^t)_*$ are those induced in homology by the inclusions i_s^t.

Acompanhar os ciclos ao longo do tempo

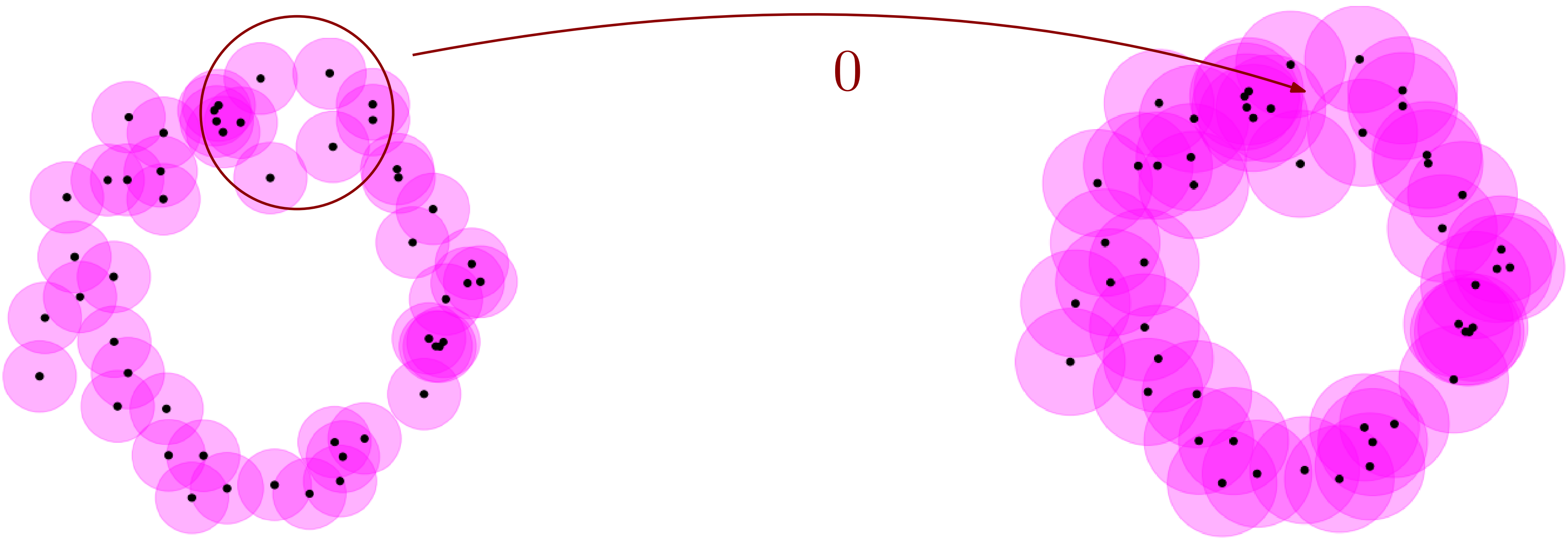
$$\dashrightarrow H_i(\check{\mathrm{C}}\mathrm{ech}^{t_1}(X)) \xrightarrow{\left(i_{t_1}^{t_2}\right)_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_2}(X)) \xrightarrow{\left(i_{t_2}^{t_3}\right)_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_3}(X)) \dashrightarrow$$

Let $i \geq 0$, $t_0 \geq 0$ and consider a cycle $c \in H_i(\check{\mathrm{C}}\mathrm{ech}^{t_0}(X))$.

Its **death time** is: $\sup\left\{t \geq t_0 \mid \left(i_{t_0}^{t}\right)(c) \neq 0\right\}$,

its **birth time** is: $\inf\left\{t \geq t_0 \mid \left(i_t^{t_0}\right)^{-1}(\{c\}) \neq \emptyset\right\}$,

its **persistence** is the difference.



Acompanhar os ciclos ao longo do tempo

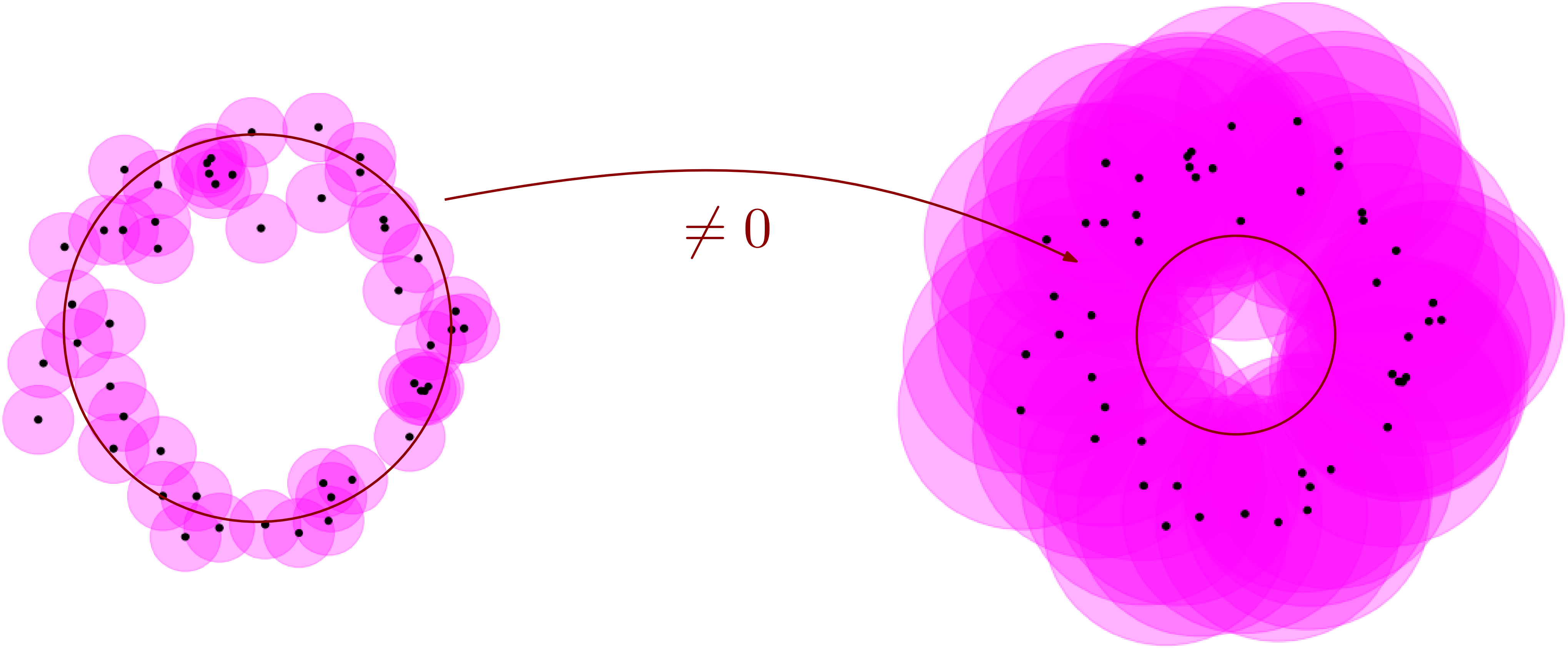
$$\dashrightarrow H_i(\check{\mathrm{C}}\mathrm{ech}^{t_1}(X)) \xrightarrow{\left(i_{t_1}^{t_2}\right)_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_2}(X)) \xrightarrow{\left(i_{t_2}^{t_3}\right)_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_3}(X)) \dashrightarrow$$

Let $i \geq 0$, $t_0 \geq 0$ and consider a cycle $c \in H_i(\check{\mathrm{C}}\mathrm{ech}^{t_0}(X))$.

Its **death time** is: $\sup\left\{t \geq t_0 \mid \left(i_{t_0}^{t}\right)(c) \neq 0\right\}$,

its **birth time** is: $\inf\left\{t \geq t_0 \mid \left(i_t^{t_0}\right)^{-1}(\{c\}) \neq \emptyset\right\}$,

its **persistence** is the difference.



$$\dashrightarrow H_i(\check{\mathrm{C}}\mathrm{ech}^{t_1}(X)) \xrightarrow{\left(i_{t_1}^{t_2}\right)_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_2}(X)) \xrightarrow{\left(i_{t_2}^{t_3}\right)_*} H_i(\check{\mathrm{C}}\mathrm{ech}^{t_3}(X)) \dashrightarrow$$

Let $i \geq 0$, $t_0 \geq 0$ and consider a cycle $c \in H_i(\check{\mathrm{C}}\mathrm{ech}^{t_0}(X))$.

Its **death time** is: $\sup \left\{t \geq t_0 \mid \left(i_{t_0}^{t}\right)(c) \neq 0\right\}$,

its **birth time** is: $\inf \left\{t \geq t_0 \mid \left(i_{t}^{t_0}\right)^{-1}(\{c\}) \neq \emptyset\right\}$,

its **persistence** is the difference.

As a rule of thumb:

- cycles with **large persistence** correspond to important topological features of the dataset,
- cycles with **short persistence** corresponds to topological noise.

Módulos de persistência

Definition: A **persistence module** $\mathbb{V}$ over $\mathbb{R}^+$ with coefficients in $\mathbb{Z}/2\mathbb{Z}$ is a pair $(\mathbb{V}, v)$ where $\mathbb{V} = (V^t)_{t\in\mathbb{R}^+}$ is a family of $\mathbb{Z}/2\mathbb{Z}$-vector spaces, and $v = (v_s^t \colon V^s \to V^t)_{s\leq t\in\mathbb{R}^+}$ a family of linear maps such that:

- for every $t \in \mathbb{R}^+$, $v_t^t \colon V^t \to V^t$ is the identity map,
- for every $r, s, t \in \mathbb{R}^+$ such that $r \leq s \leq t$, we have $v_s^t \circ v_r^s = v_r^t$.

Módulos de persistência

Definition: A **persistence module** $\mathbb{V}$ over $\mathbb{R}^+$ with coefficients in $\mathbb{Z}/2\mathbb{Z}$ is a pair $(\mathbb{V}, v)$ where $\mathbb{V} = (V^t)_{t\in\mathbb{R}^+}$ is a family of $\mathbb{Z}/2\mathbb{Z}$-vector spaces, and $v = (v_s^t \colon V^s \to V^t)_{s\leq t\in\mathbb{R}^+}$ a family of linear maps such that:

- for every $t \in \mathbb{R}^+$, $v_t^t \colon V^t \to V^t$ is the identity map,
- for every $r, s, t \in \mathbb{R}^+$ such that $r \leq s \leq t$, we have $v_s^t \circ v_r^s = v_r^t$.

In practice, one builds persistence modules from **filtrations**, that is, non-decreasing families of simplicial complexes $\mathbb{S} = (S^t)_{t\in\mathbb{R}^+}$.

By applying the i^{th} homology functor to a filtration, we obtain a persistence module $\mathbb{V}[\mathbb{S}] = (H_i(S^t))_{t\in\mathbb{R}^+}$, with maps $((i_s^t)_* \colon H_i(S^s) \to H_i(S^t))_{s\leq t}$ induced by the inclusions.

$$\dashrightarrow S^{t_1} \xhookrightarrow{i_{t_1}^{t_2}} S^{t_2} \xhookrightarrow{i_{t_2}^{t_3}} S^{t_3} \xhookrightarrow{i_{t_3}^{t_4}} S^{t_4} \hookleftarrow\dashleftarrow$$

$$\dashrightarrow H_i(S^{t_1}) \xrightarrow{\left(i_{t_1}^{t_2}\right)_*} H_i(S^{t_2}) \xrightarrow{\left(i_{t_2}^{t_3}\right)_*} H_i(S^{t_3}) \xrightarrow{\left(i_{t_3}^{t_4}\right)_*} H_i(S^{t_4}) \dashrightarrow$$

Isomorfismos de módulos

Definition: An **isomorphism** between two persistence modules $(\mathbb{V}, v)$ and $(\mathbb{W}, w)$ is a family of isomorphisms of vector spaces $\phi = (\phi_t \colon V^t \to W^t)_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^s & \xrightarrow{w_s^t} & W^t \end{array}$$

Decomponibilidade

Definition: Let $(\mathbb{V}, v)$ and $(\mathbb{W}, w)$ be two persistence modules.

Their **sum** is the persistence module $\mathbb{V} \oplus \mathbb{W}$ defined with the vector spaces $(V \oplus W)^t = V^t \oplus W^t$ and the linear maps

$$(v \oplus w)_s^t \colon (x, y) \in (V \oplus W)^s \longmapsto (v_s^t(x), w_s^t(y)) \in (V \oplus W)^t.$$

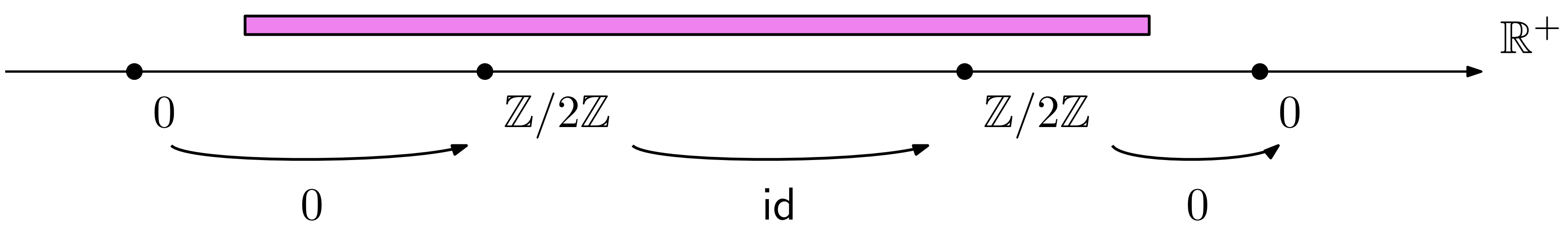
A persistence module $\mathbb{U}$ is **indecomposable** if for every pair of persistence modules $\mathbb{V}$ and $\mathbb{W}$ such that $\mathbb{U}$ is isomorphic to the sum $\mathbb{V} \oplus \mathbb{W}$, then one of the summands has to be a trivial persistence module, that is, equal to zero for every $t \in \mathbb{R}^+$.
Otherwise, $\mathbb{U}$ is said **decomposable**.

Módulos intervalo

Definition: Let $I \subset \mathbb{R}^+$ be an interval: $[a,b]$, $(a,b]$, $[a,b)$ or (a,b), with $a,b \in \mathbb{R}^+$ such that $a \le b$, and potentially $a = -\infty$ or $b = +\infty$.

The **interval module** associated to I is the persistence module $\mathbb{B}[I]$ with vector spaces $\mathbb{B}^t[I]$ and linear maps $v_s^t \colon \mathbb{B}^s[I] \to \mathbb{B}^t[I]$ defined as

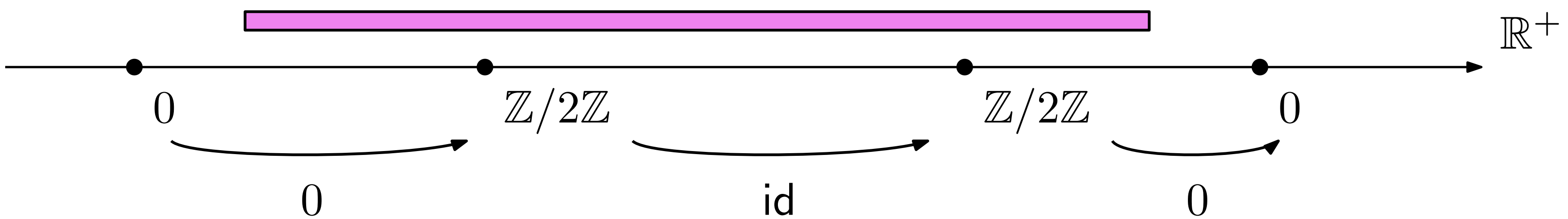
$$\mathbb{B}^t[I] = \begin{cases} \mathbb{Z}/2\mathbb{Z} & \text{if } t \in I, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad v_s^t = \begin{cases} \text{id} & \text{if } s,t \in I, \\ 0 & \text{otherwise.} \end{cases}$$



Definition: Let $I \subset \mathbb{R}^+$ be an interval: $[a,b]$, $(a,b]$, $[a,b)$ or (a,b), with $a,b \in \mathbb{R}^+$ such that $a \leq b$, and potentially $a = -\infty$ or $b = +\infty$.

The **interval module** associated to I is the persistence module $\mathbb{B}[I]$ with vector spaces $\mathbb{B}^t[I]$ and linear maps $v_s^t \colon \mathbb{B}^s[I] \to \mathbb{B}^t[I]$ defined as

$$\mathbb{B}^t[I] = \begin{cases} \mathbb{Z}/2\mathbb{Z} & \text{if } t \in I, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad v_s^t = \begin{cases} \text{id} & \text{if } s,t \in I, \\ 0 & \text{otherwise.} \end{cases}$$



Lemma: Interval modules are indecomposable.

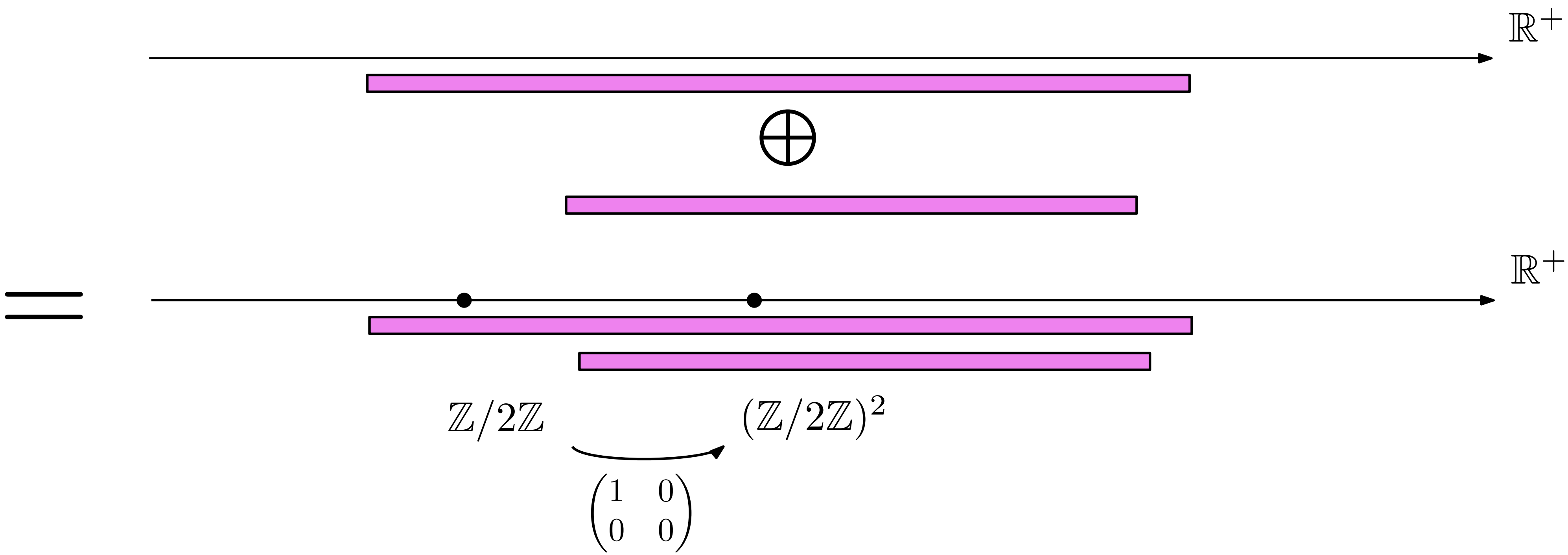
Módulos intervalo

Definition: Let $I \subset \mathbb{R}^+$ be an interval: $[a,b]$, $(a,b]$, $[a,b)$ or (a,b), with $a,b \in \mathbb{R}^+$ such that $a \leq b$, and potentially $a = -\infty$ or $b = +\infty$.

The **interval module** associated to I is the persistence module $\mathbb{B}[I]$ with vector spaces $\mathbb{B}^t[I]$ and linear maps $v_s^t \colon \mathbb{B}^s[I] \to \mathbb{B}^t[I]$ defined as

$$\mathbb{B}^t[I] = \begin{cases} \mathbb{Z}/2\mathbb{Z} & \text{if } t \in I, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad v_s^t = \begin{cases} \mathrm{id} & \text{if } s,t \in I, \\ 0 & \text{otherwise.} \end{cases}$$

We can sum interval modules:

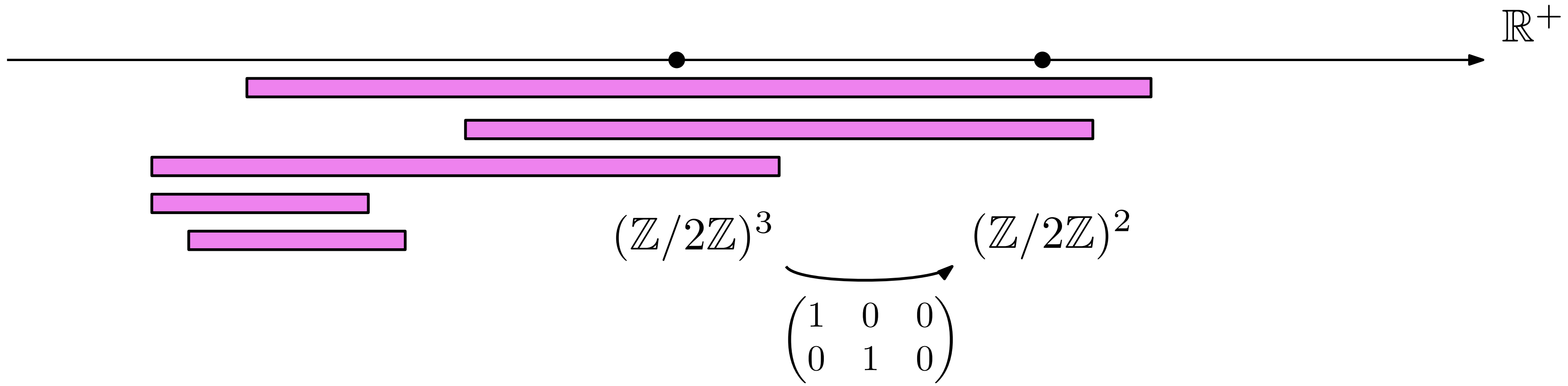


Módulos intervalo

Definition: Let $I \subset \mathbb{R}^+$ be an interval: $[a,b]$, $(a,b]$, $[a,b)$ or (a,b), with $a,b \in \mathbb{R}^+$ such that $a \leq b$, and potentially $a = -\infty$ or $b = +\infty$.

The **interval module** associated to I is the persistence module $\mathbb{B}[I]$ with vector spaces $\mathbb{B}^t[I]$ and linear maps $v_s^t \colon \mathbb{B}^s[I] \to \mathbb{B}^t[I]$ defined as

$$\mathbb{B}^t[I] = \begin{cases} \mathbb{Z}/2\mathbb{Z} & \text{if } t \in I, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad v_s^t = \begin{cases} \text{id} & \text{if } s,t \in I, \\ 0 & \text{otherwise.} \end{cases}$$



Barcodes

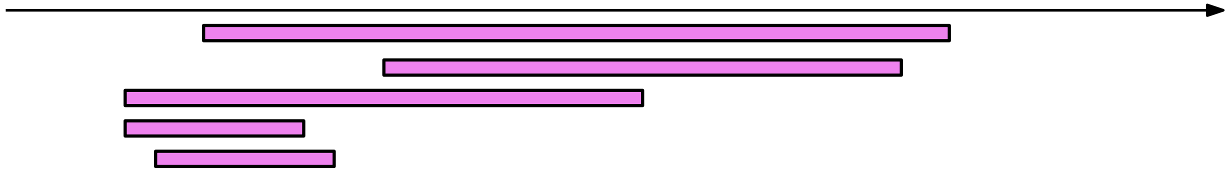
A persistence module $\mathbb{V}$ **decomposes into interval module** if there exists a multiset $\mathcal{I}$ of intervals of T such that

$$\mathbb{V} \simeq \bigoplus_{I \in \mathcal{I}} \mathbb{B}[I].$$

Multiset means that $\mathcal{I}$ may contain several copies of the same interval I.

Theorem (consequence of Krull–Remak–Schmidt–Azumaya): If a persistence module decomposes into interval modules, then the multiset $\mathcal{I}$ of intervals is unique.

In this case, $\mathcal{I}$ is called the **persistence barcode** of $\mathbb{V}$. It is written $\mathrm{Barcode}\,(\mathbb{V})$.



Barcodes

A persistence module $\mathbb{V}$ **decomposes into interval module** if there exists a multiset $\mathcal{I}$ of intervals of T such that

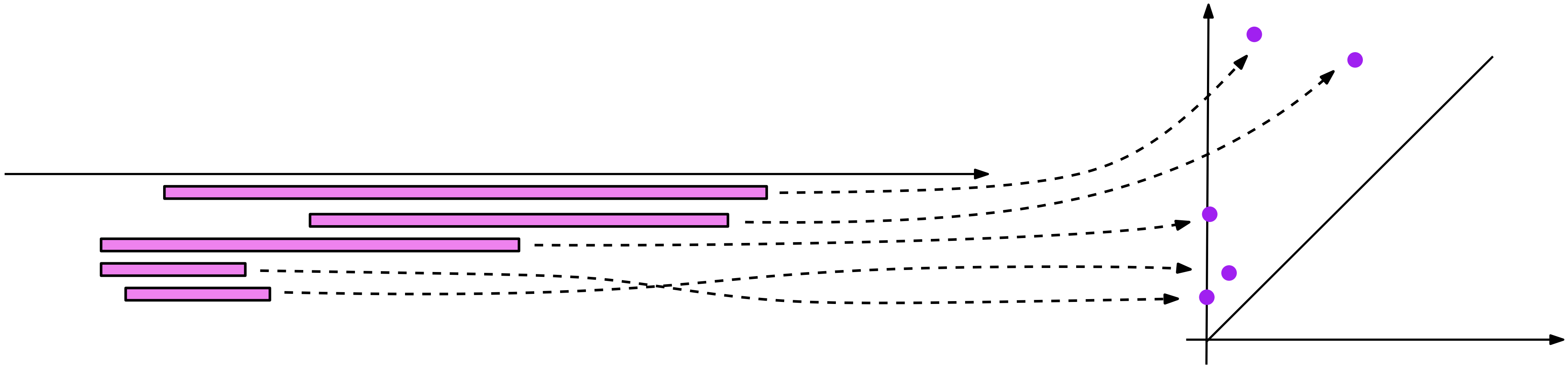
$$\mathbb{V} \simeq \bigoplus_{I \in \mathcal{I}} \mathbb{B}[I].$$

Multiset means that $\mathcal{I}$ may contain several copies of the same interval I.

Theorem (consequence of Krull–Remak–Schmidt–Azumaya): If a persistence module decomposes into interval modules, then the multiset $\mathcal{I}$ of intervals is unique.

In this case, $\mathcal{I}$ is called the **persistence barcode** of $\mathbb{V}$. It is written $\mathrm{Barcode}\,(\mathbb{V})$.

For every $[a, b]$, $(a, b]$, $[a, b)$ or (a, b) in $\mathrm{Barcode}\,(\mathbb{V})$, consider the point (a, b) of $\mathbb{R}^2$. The collection of all such points is the **persistence diagram** of $\mathbb{V}$.



Barcodes

A persistence module $\mathbb{V}$ **decomposes into interval module** if there exists a multiset $\mathcal{I}$ of intervals of T such that

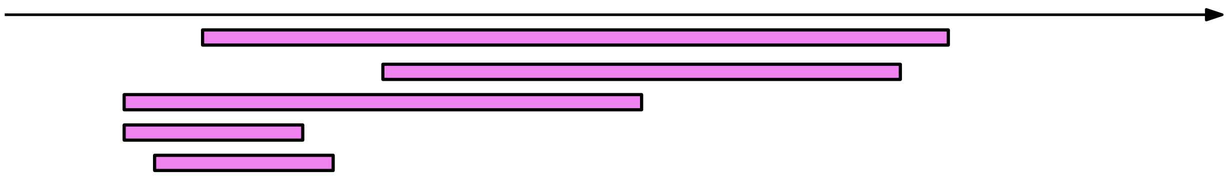
$$\mathbb{V} \simeq \bigoplus_{I \in \mathcal{I}} \mathbb{B}[I].$$

Multiset means that $\mathcal{I}$ may contain several copies of the same interval I.

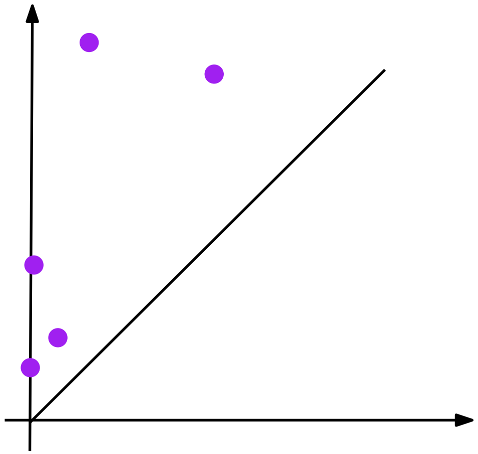
Theorem (consequence of Krull–Remak–Schmidt–Azumaya): If a persistence module decomposes into interval modules, then the multiset $\mathcal{I}$ of intervals is unique.

In this case, $\mathcal{I}$ is called the **persistence barcode** of $\mathbb{V}$. It is written $\mathrm{Barcode}(\mathbb{V})$.

For every $[a, b]$, $(a, b]$, $[a, b)$ or (a, b) in $\mathrm{Barcode}(\mathbb{V})$, consider the point (a, b) of $\mathbb{R}^2$. The collection of all such points is the **persistence diagram** of $\mathbb{V}$.



$\mathrm{Barcode}(\mathbb{V})$



$\mathrm{Diagram}(\mathbb{V})$

Barcodes

A persistence module $\mathbb{V}$ is said **pointwise finite dimensional** if $\dim V^t < +\infty$ for all t.

Theorem (Crawley-Boevey, 2015): Every pointwise finite-dimensional persistence module decomposes into interval modules.

Barcodes

A persistence module $\mathbb{V}$ is said **pointwise finite dimensional** if $\dim V^t < +\infty$ for all t.

Theorem (Crawley-Boevey, 2015): Every pointwise finite-dimensional persistence module decomposes into interval modules.

Proof (Zomorodian, Carlsson, 2005): **Simpler case:** the persistence module is finite-dimensional *and* has finitely many terms.
We can write our persistence module as

$$V^1 \xrightarrow{v_1^2} V^2 \xrightarrow{v_2^3} V^3 \xrightarrow{v_3^4} V^4 \dashrightarrow \dots \dashrightarrow V^n$$

Consider the vector space $\mathcal{V} = \bigotimes_{1\leq i\leq n} V^i = V^1 \times \cdots \times V^n$.
Let $\mathbb{Z}/2\mathbb{Z}[x]$ denote the space of polynomials with coefficients in $\mathbb{Z}/2\mathbb{Z}$. We give $\mathcal{V}$ an action of $\mathbb{Z}/2\mathbb{Z}[x]$ via

$$x \cdot (a^1, a^2, ..., a^n) = (0, v_1^2(a^1), v_2^3(a^2), ..., v_{n-1}^n(a^{n-1})).$$

Hence $\mathcal{V}$ can be seen as a finitely generated module over the principal ideal domain $\mathbb{Z}/2\mathbb{Z}[x]$. By classification, $\mathcal{V}$ is isomorphic to a sum

$$\mathcal{V} \simeq \bigoplus_{i\in I} \mathbb{Z}/2\mathbb{Z}[x]/x^i \cdot \mathbb{Z}/2\mathbb{Z}[x].$$

We identify the components $\mathbb{Z}/2\mathbb{Z}[x]/x^i \cdot \mathbb{Z}/2\mathbb{Z}[x]$ with bars of the barcode of length i.

Barcodes

Barcodes

On a barcode we can read homology **at each step**, and see how it **evolves**.

Algoritmo

The Čech or the Rips filtration define an increasing sequence of simplices

$$\ldots \subset \check{\mathrm{C}}\mathrm{ech}^{t_1}(X) \subset \check{\mathrm{C}}\mathrm{ech}^{t_2}(X) \subset \check{\mathrm{C}}\mathrm{ech}^{t_3}(X) \subset \ldots$$

We can turn it consistently into an ordering of the simplices, by inserting the simplices by order of apparition in the filtration.

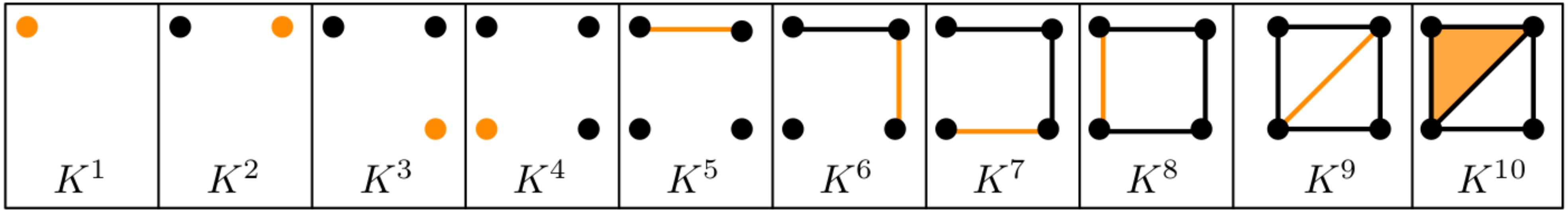
$$\sigma^1 < \sigma^2 < \ldots < \sigma^n$$

Denote $t(\sigma)$ the time of apparition of the simplex σ in the filtration. The total order on the simplices satisfies

$$t(\sigma^i) < t(\sigma^j) \text{ for all } i < j.$$

In practice several simplices may appear at the same time. If this occurs, choose an order of the simplices.

⟶ Consider the boundary matrix, and compute a Gauss reduction.



Algoritmo

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	1	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	σ^8	σ^9	σ^{10}
σ^1	0	0	0	0	1	0	0	1	0	0
σ^2	0	0	0	0	1	1	0	0	1	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	1	1	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

For any $j \in [\![1, n]\!]$,

$$\delta(j) = \max\{i \in [\![1, n]\!] \mid \Delta_{i,j} \neq 0\},$$

and $\Delta_{i,j} = 0$ for all j, then $\delta(j)$ is *undefined*.

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Algoritmo

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.

$$\begin{array}{c|cccccccccc}
 & \sigma^1 & \sigma^2 & \sigma^3 & \sigma^4 & \sigma^5 & \sigma^6 & \sigma^7 & \sigma^5+\sigma^6+\sigma^7+\sigma^8 & \sigma^9+\sigma^6+\sigma^7 & \sigma^{10} \\
\hline
\sigma^1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\sigma^2 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
\sigma^3 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
\sigma^4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\sigma^5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\sigma^6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\sigma^7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\sigma^8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\sigma^9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\sigma^{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Algoritmo

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1 ($+\infty$)	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Algoritmo

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.

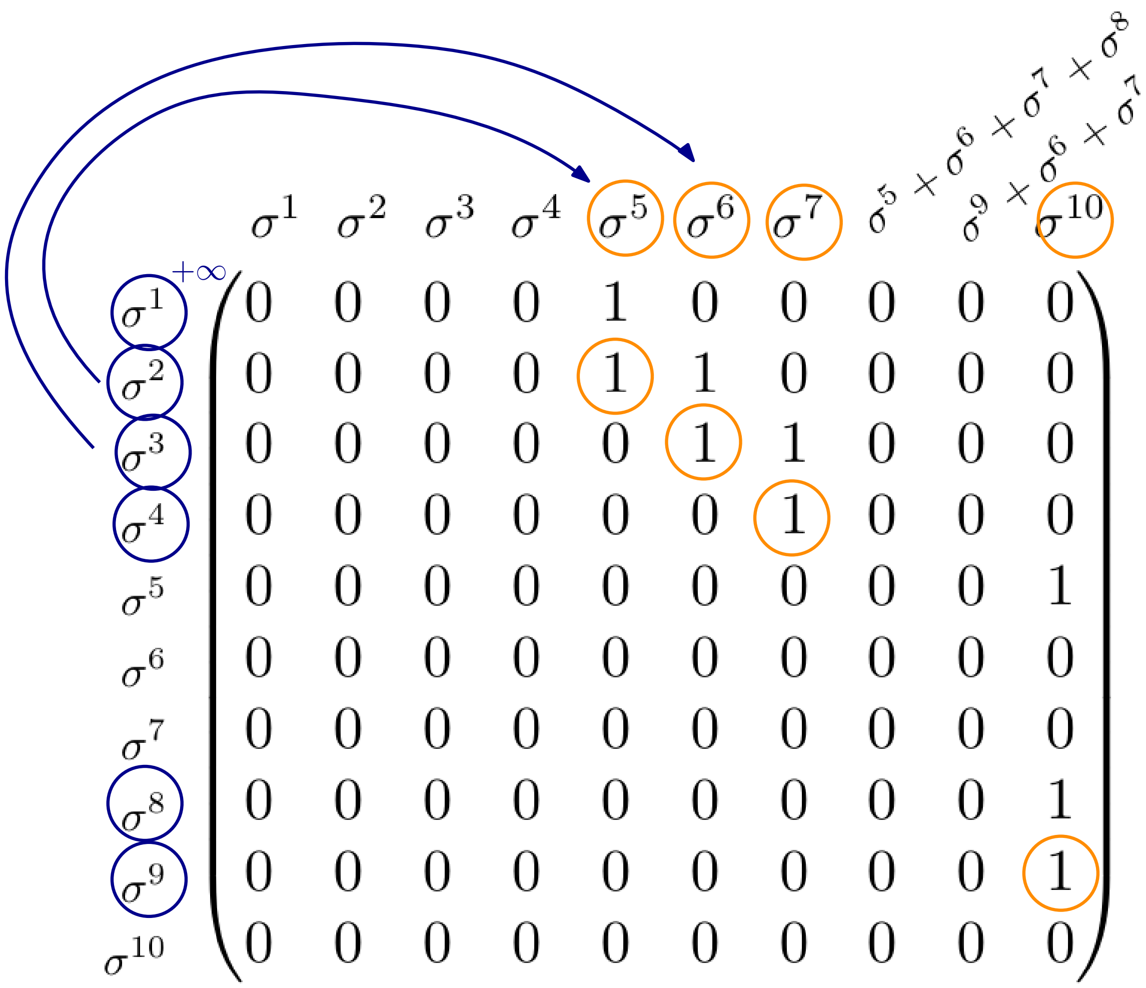
	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1 ($+\infty$)	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.



Algoritmo

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.

	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1 ($+\infty$)	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Algoritmo

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.

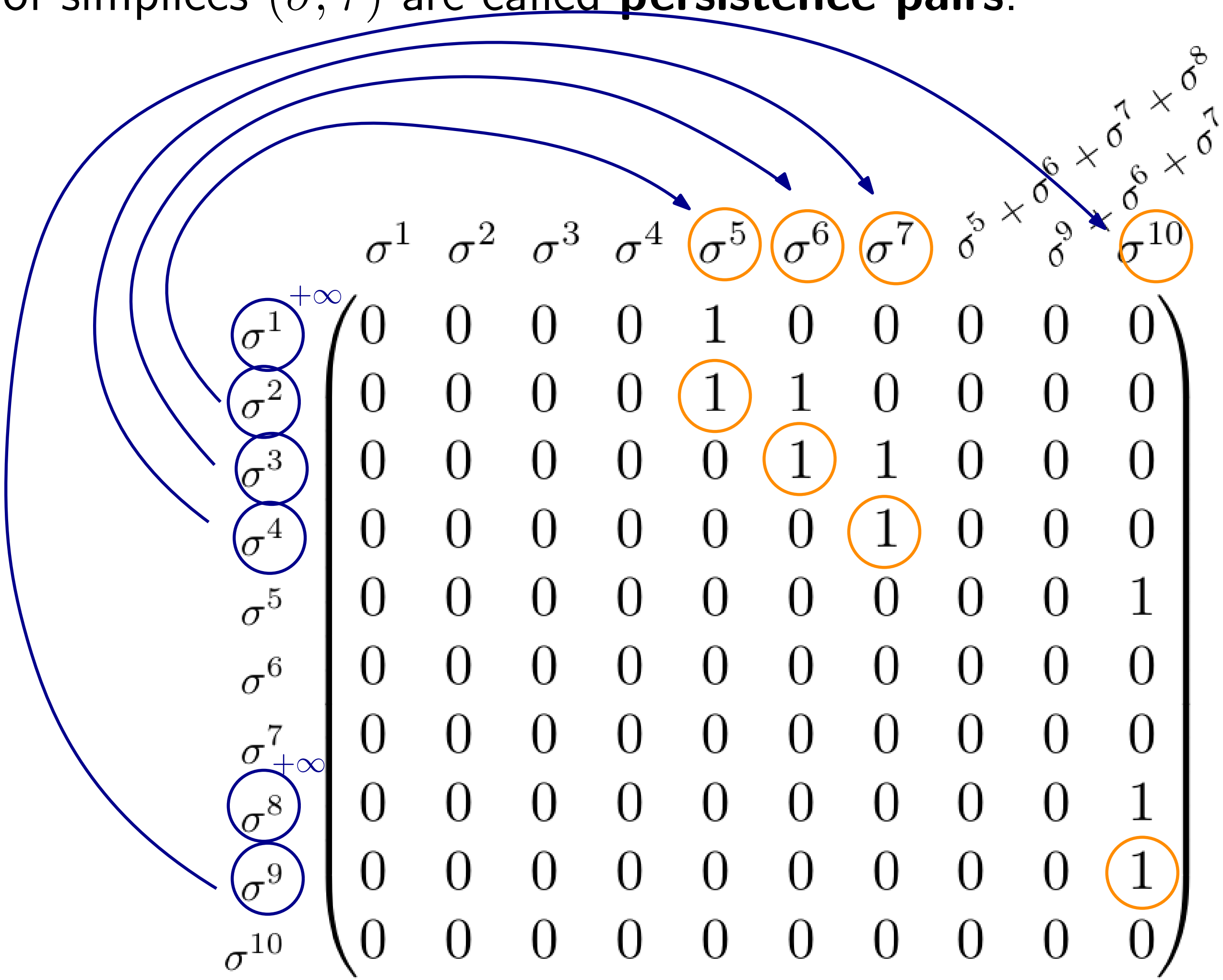
	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1 $+\infty$	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8 $+\infty$	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.



$$\begin{pmatrix}
0&0&0&0&1&0&0&0&0&0\\
0&0&0&0&1&1&0&0&0&0\\
0&0&0&0&0&1&1&0&0&0\\
0&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0
\end{pmatrix}$$

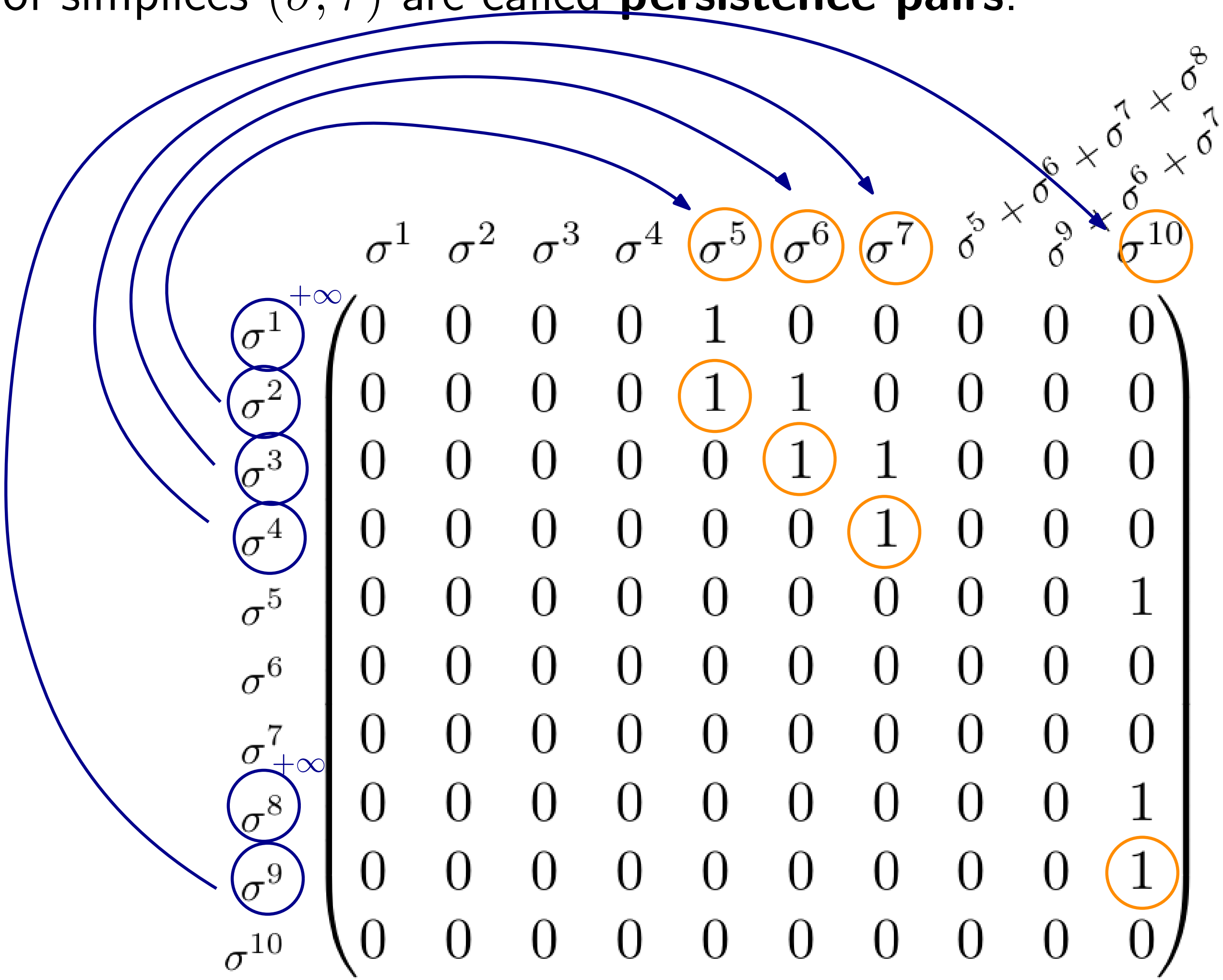
Algoritmo

Now, for all j such that $\delta(j)$ is defined, consider the pair of simplices

$$\left(\sigma^{\delta(j)}, \sigma^{j}\right).$$

Also, for the values $i \notin \mathrm{Im}(\delta)$, we set: $\left(\sigma^{i}, +\infty\right)$.

The pairs of simplices (σ, τ) are called **persistence pairs**.



	σ^1	σ^2	σ^3	σ^4	σ^5	σ^6	σ^7	$\sigma^5+\sigma^6+\sigma^7+\sigma^8$	$\sigma^9+\sigma^6+\sigma^7$	σ^{10}
σ^1	0	0	0	0	1	0	0	0	0	0
σ^2	0	0	0	0	1	1	0	0	0	0
σ^3	0	0	0	0	0	1	1	0	0	0
σ^4	0	0	0	0	0	0	1	0	0	0
σ^5	0	0	0	0	0	0	0	0	0	1
σ^6	0	0	0	0	0	0	0	0	0	0
σ^7	0	0	0	0	0	0	0	0	0	0
σ^8	0	0	0	0	0	0	0	0	0	1
σ^9	0	0	0	0	0	0	0	0	0	1
σ^{10}	0	0	0	0	0	0	0	0	0	0

$(\sigma^1, +\infty)$

(σ^2, σ^5)

(σ^3, σ^6)

(σ^4, σ^7)

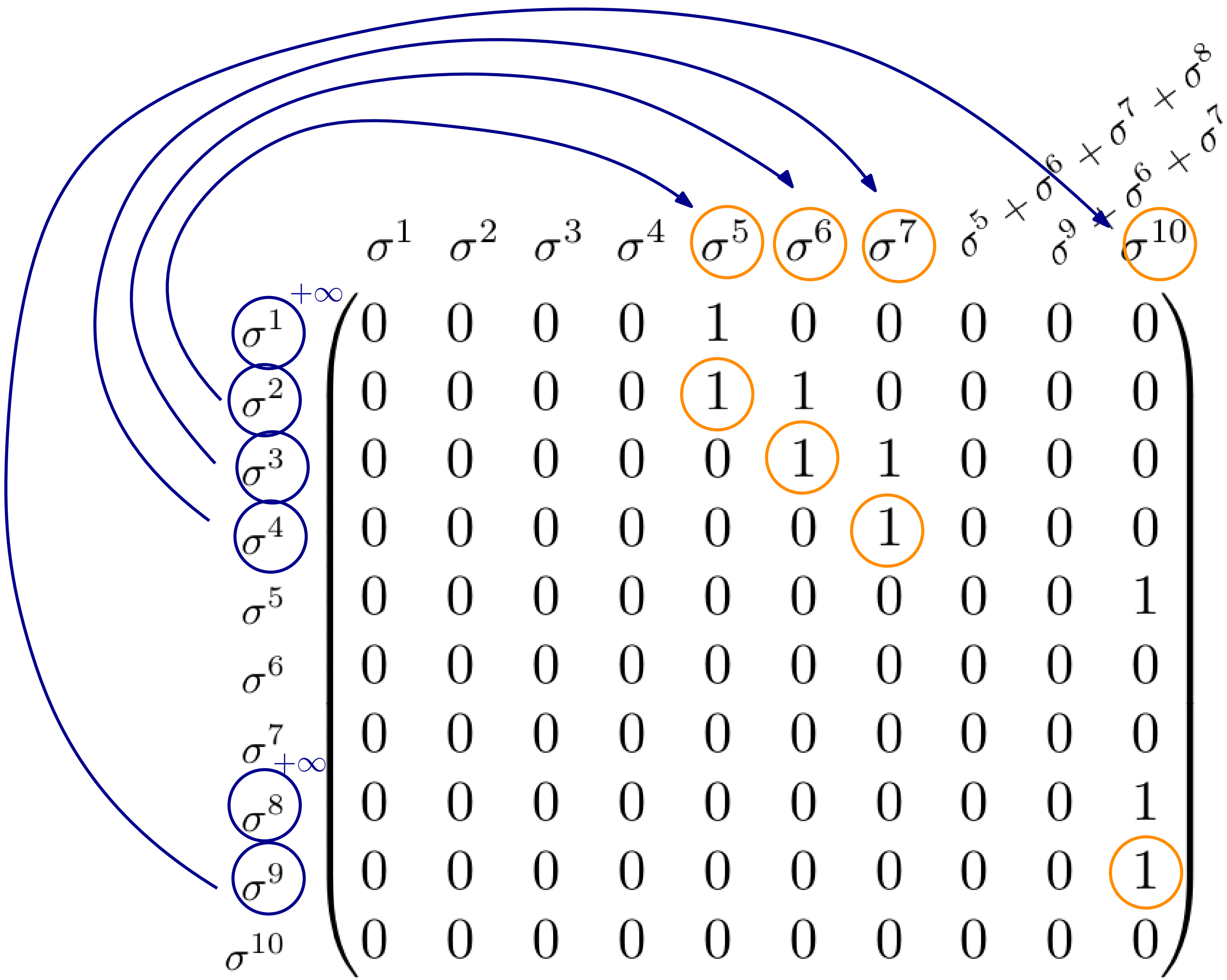
$(\sigma^8, +\infty)$

(σ^9, σ^{10})

Algoritmo

Proposition: The barcodes of the filtration consists in the intervals

$$\mathcal{I} = \left\{ (t(\sigma), t(\tau)) \text{ for all persistence pair } (\sigma, \tau) \text{ such that } t(\sigma) \neq t(\tau) \right\}.$$



$(\sigma^1, +\infty)$

(σ^2, σ^5)

(σ^3, σ^6)

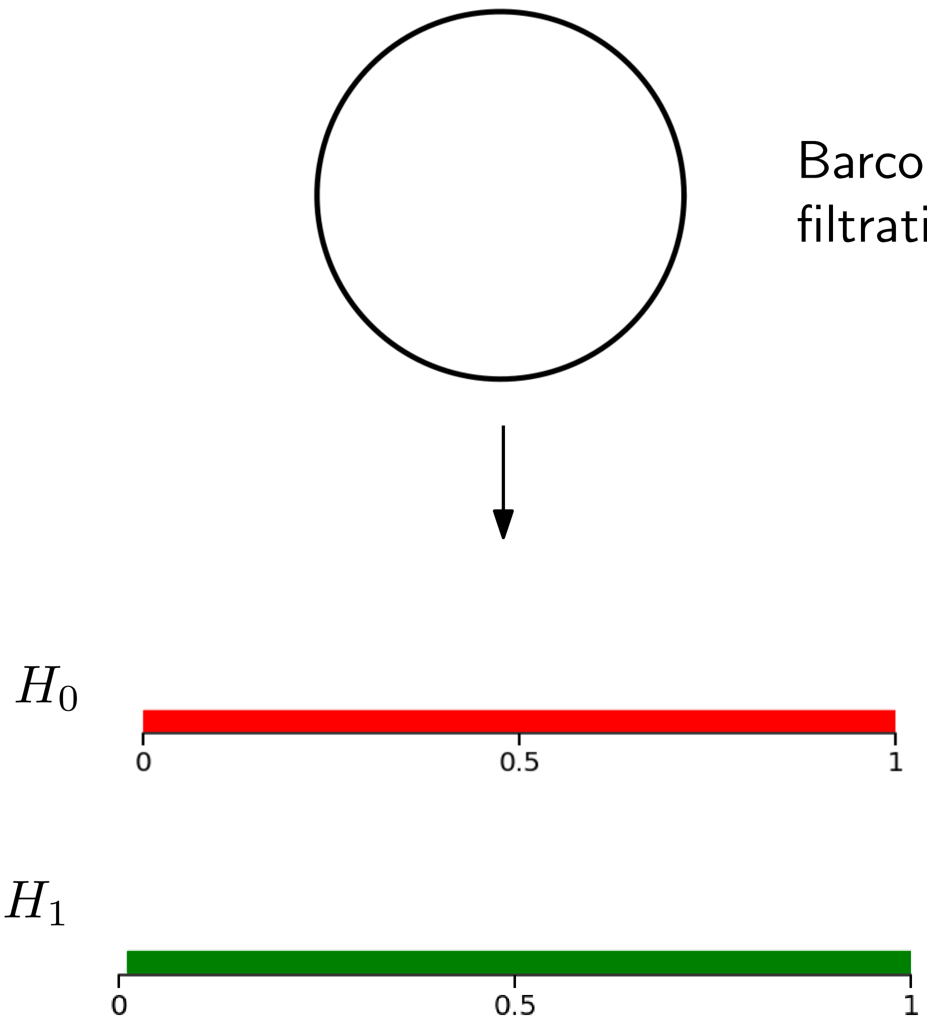
(σ^4, σ^7)

$(\sigma^8, +\infty)$

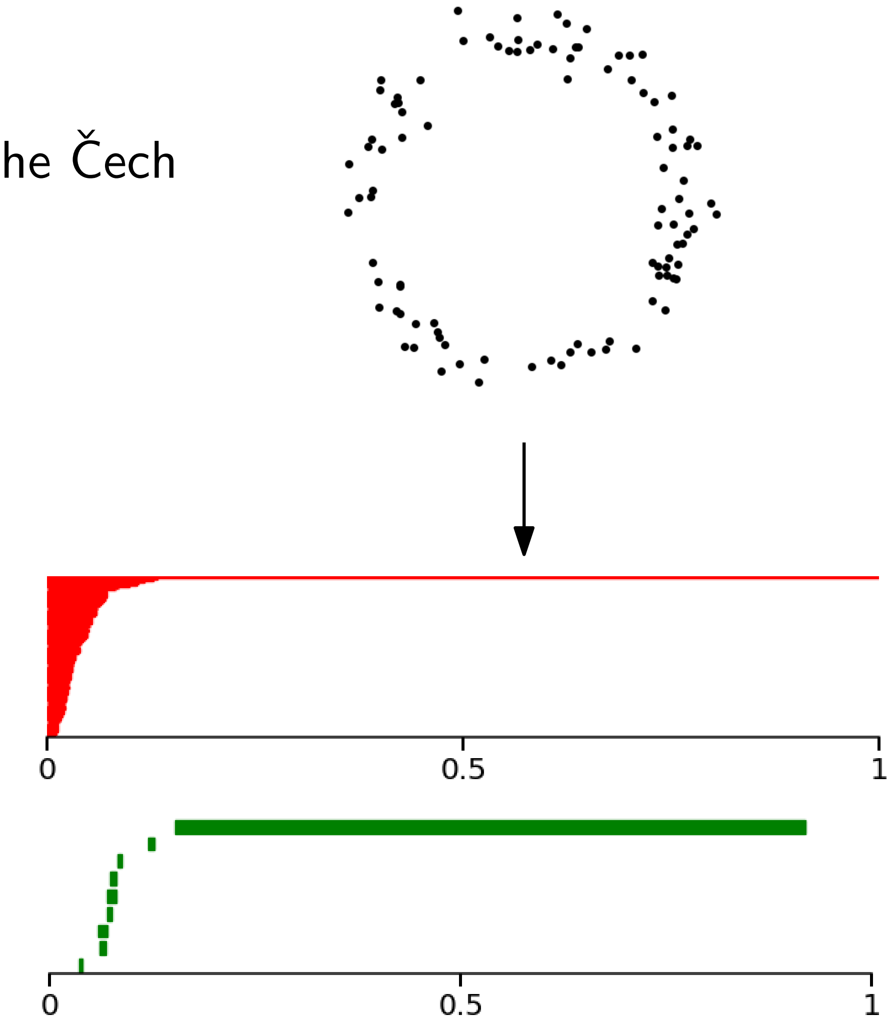
(σ^9, σ^{10})

O problema da estabilidade

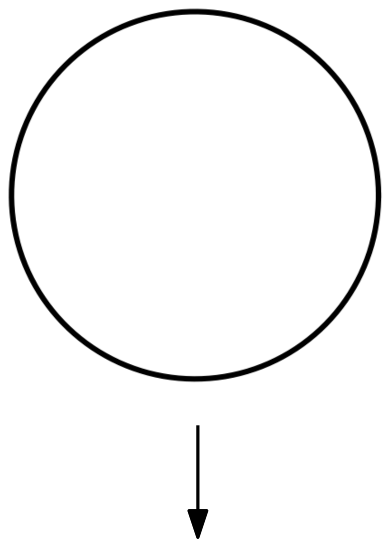
Let $X \subset \mathbb{R}^n$ finite, seen as a sample of $\mathcal{M}$.



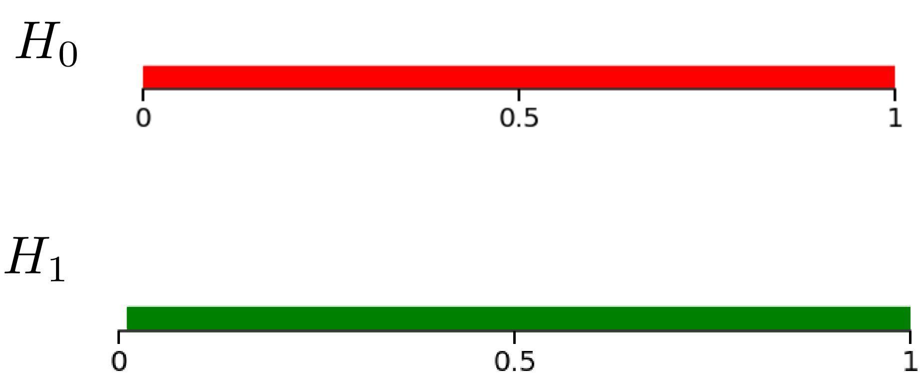
Barcodes of the Čech filtration

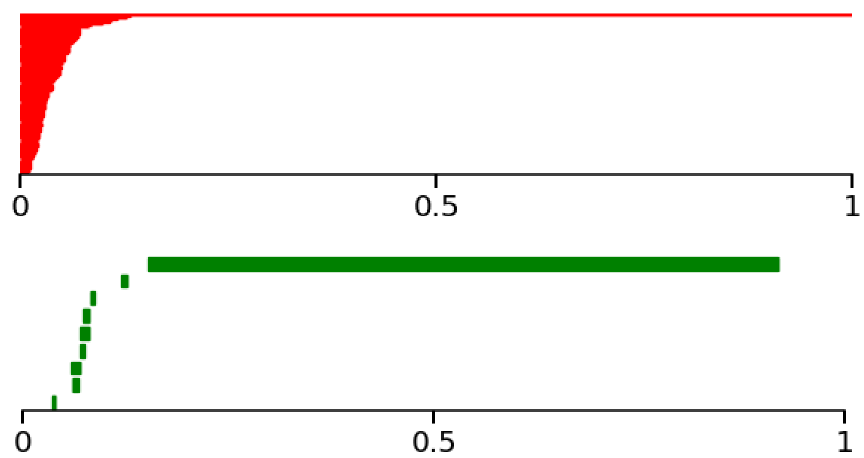


Let $X \subset \mathbb{R}^n$ finite, seen as a sample of $\mathcal{M}$.



Barcodes of the Čech filtration

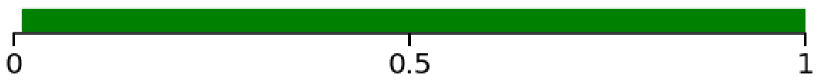


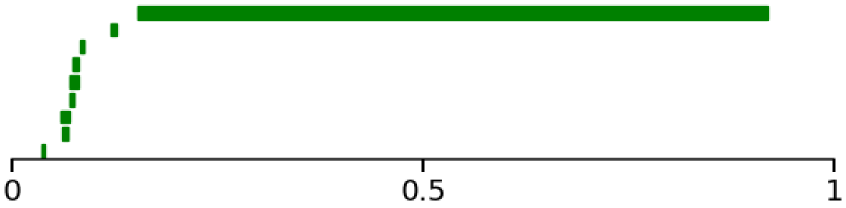


stability

Distância bottleneck

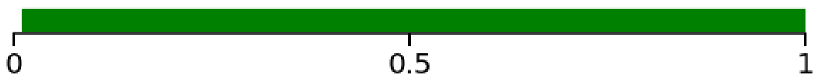
Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.

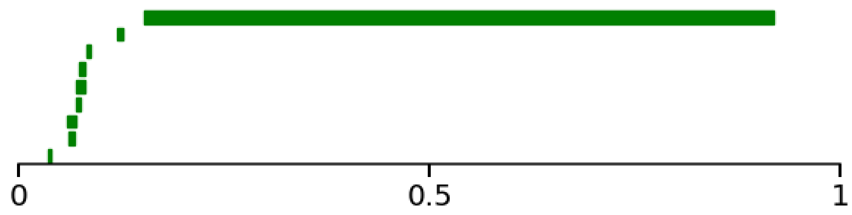




Distância bottleneck

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.



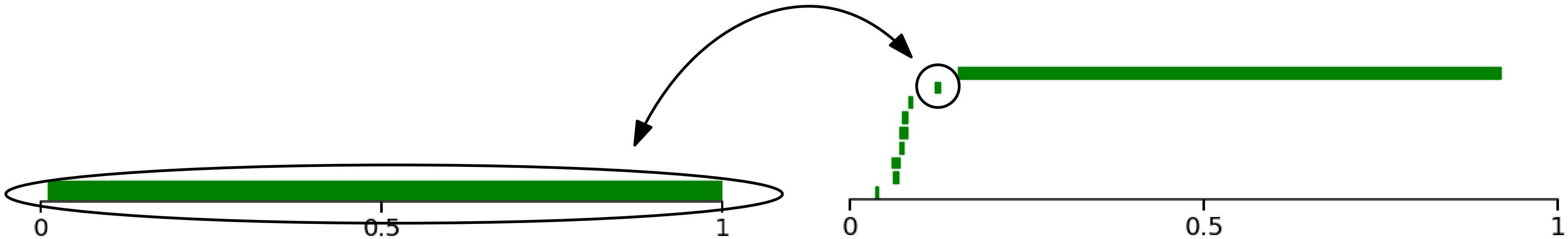


A **partial matching** between the barcodes is a subset $M \subset P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The bars $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said **matched** by M.

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.

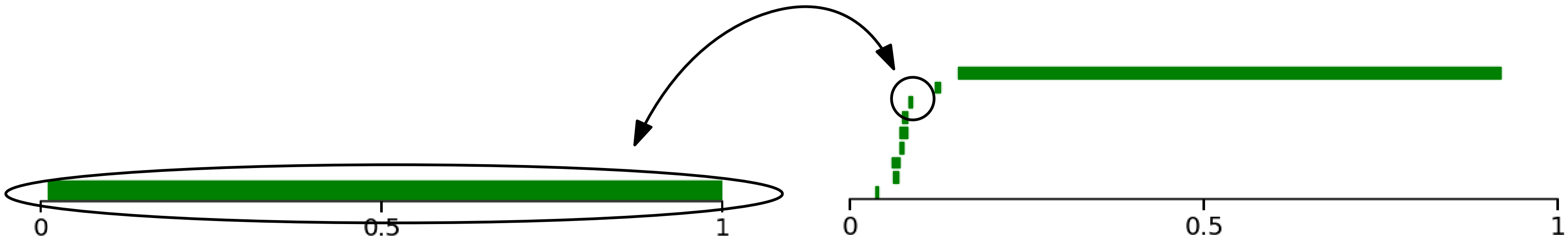


A **partial matching** between the barcodes is a subset $M \subset P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The bars $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said **matched** by M.

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.



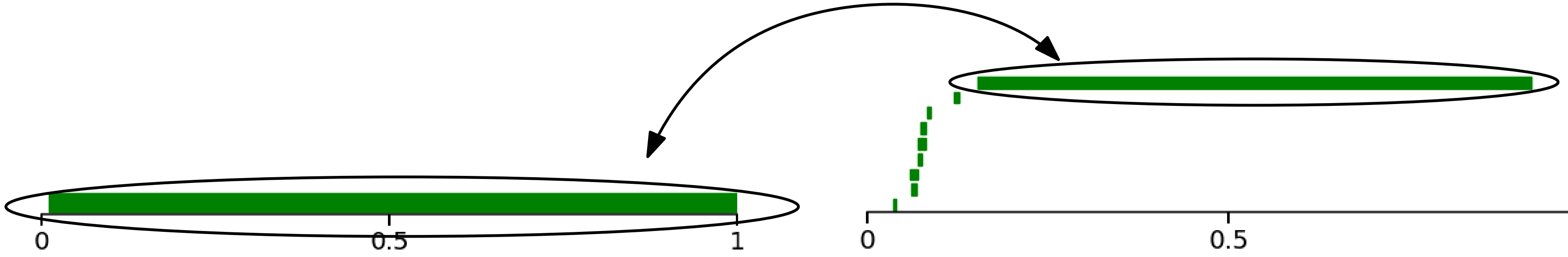
A **partial matching** between the barcodes is a subset $M \subset P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The bars $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said **matched** by M.

Distância bottleneck

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.

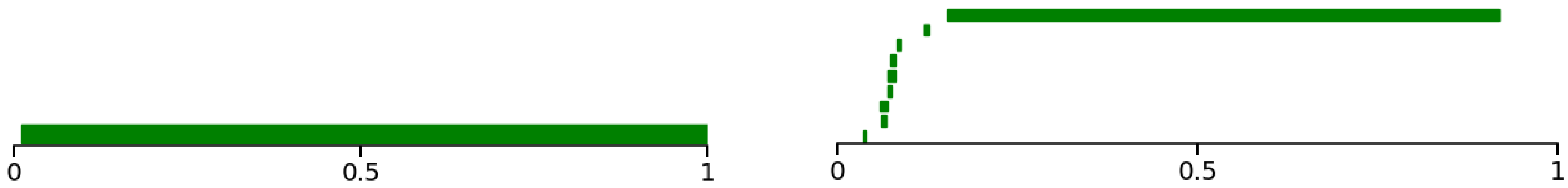


A **partial matching** between the barcodes is a subset $M \subset P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The bars $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said **matched** by M.

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.



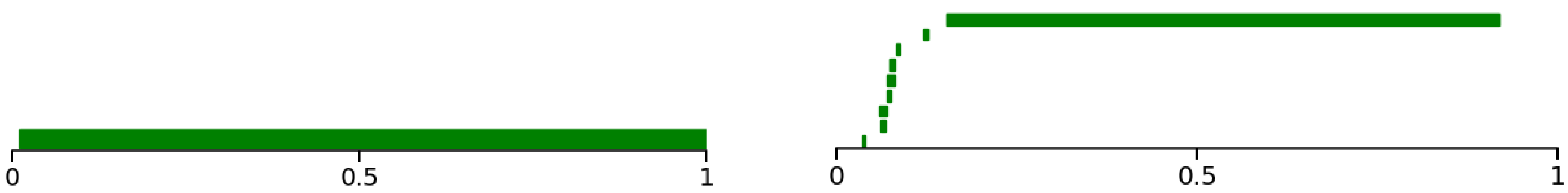
A **partial matching** between the barcodes is a subset $M \subset P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The bars $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said **matched** by M.

If a bar $p \in P$ (resp. $q \in Q$) is not matched by M, we consider that it is matched with the singleton $\overline{p} = \left[\frac{p_1+p_2}{2}, \frac{p_1+p_2}{2}\right]$ (resp. $\overline{q} = \left[\frac{q_1+q_2}{2}, \frac{q_1+q_2}{2}\right]$).

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.



A **partial matching** between the barcodes is a subset $M \subset P \times Q$ such that

- for every $p \in P$, there exists at most one $q \in Q$ such that $(p, q) \in M$,
- for every $q \in Q$, there exists at most one $p \in P$ such that $(p, q) \in M$.

The bars $p \in P$ (resp. $q \in Q$) such that there exists $q \in Q$ (resp. $p \in P$) with $(p, q) \in M$ are said **matched** by M.

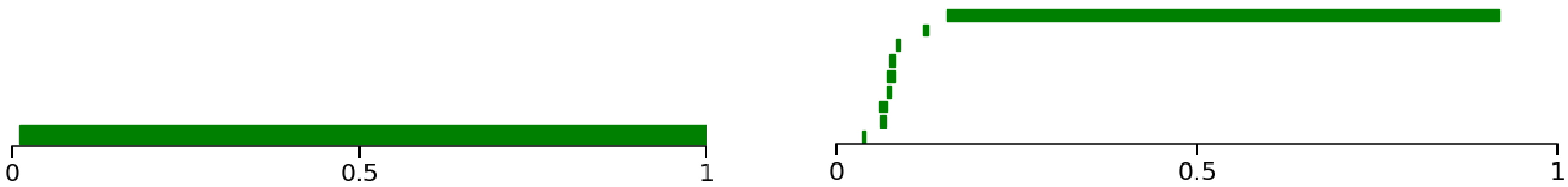
If a bar $p \in P$ (resp. $q \in Q$) is not matched by M, we consider that it is matched with the singleton $\overline{p} = \left[\frac{p_1+p_2}{2}, \frac{p_1+p_2}{2}\right]$ (resp. $\overline{q} = \left[\frac{q_1+q_2}{2}, \frac{q_1+q_2}{2}\right]$).

The **cost** of a matched pair (p, q) (resp. $(p, \overline{p})$, resp. $(\overline{q}, q)$) is the sup norm $\|p - q\|_\infty = \sup\{|p_1 - q_1|, |p_2 - q_2|\}$ (resp. $\|p - \overline{p}\|_\infty$, resp. $\|\overline{q} - q\|_\infty$).

The **cost** of the partial matching M, denoted $\mathrm{cost}(M)$, is the supremum of all costs.

Distância bottleneck

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.

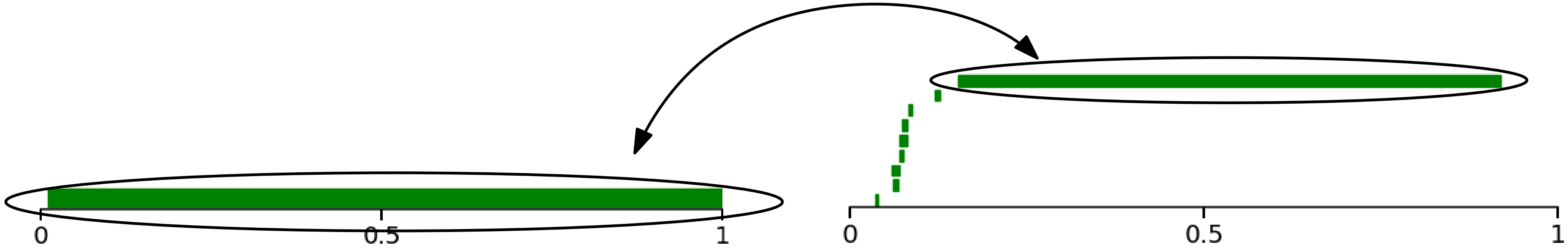


Definition: The **bottleck distance** between P and Q is defined as the infimum of costs over all the partial matchings:

$$\mathrm{d_b}(P, Q) = \inf\{\mathrm{cost}(M) \mid M \text{ is a partial matching between } P \text{ and } Q\}.$$

Distância bottleneck

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.

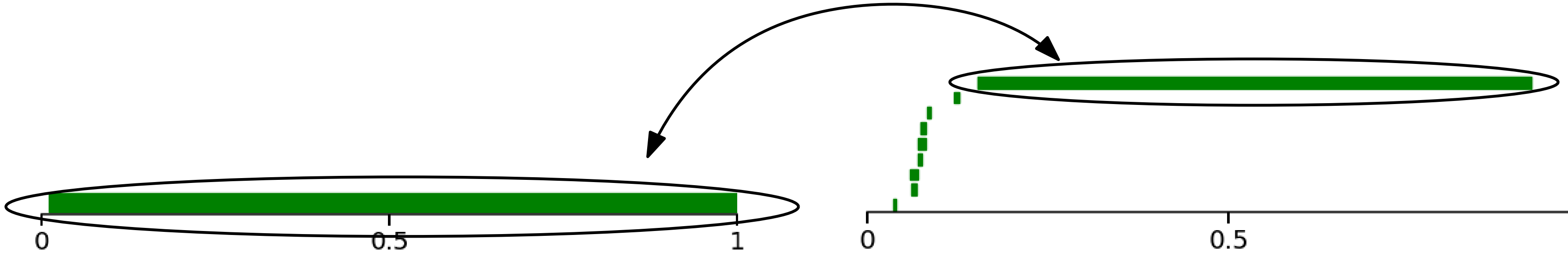


Definition: The **bottleck distance** between P and Q is defined as the infimum of costs over all the partial matchings:

$$\mathrm{d_b}(P, Q) = \inf\{\mathrm{cost}(M) \mid M \text{ is a partial matching between } P \text{ and } Q\}.$$

Distância bottleneck

Consider two barcodes P and Q, that is, multisets of intervals $\{(a_i, b_i), i \in \mathcal{I}\}$ of $(\overline{\mathbb{R}^+})^2$ such that $a_i \leq b_i$ for all $i \in \mathcal{I}$.



Definition: The **bottleck distance** between P and Q is defined as the infimum of costs over all the partial matchings:

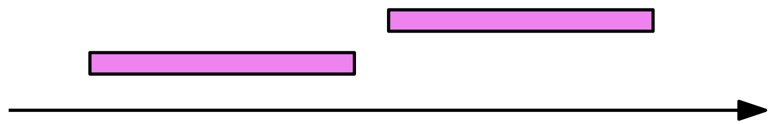
$$\mathrm{d_b}(P, Q) = \inf\{\mathrm{cost}(M) \mid M \text{ is a partial matching between } P \text{ and } Q\}.$$

If $\mathbb{U}$ and $\mathbb{V}$ are two decomposable persistence modules, we define their **bottleneck distance** as

$$\mathrm{d_b}(\mathbb{U}, \mathbb{V}) = \mathrm{d_b}(\mathrm{Barcode}(\mathbb{U}), \mathrm{Barcode}(\mathbb{V})).$$

Distância bottleneck

Example: Consider $a \leq b$, $a' \leq b'$, and the barcodes $P = \{[a,b]\}$ and $Q = \{[a',b']\}$.



First maching: the empty matching $M = \emptyset$. The intervals are matched to their midpoint, and the cost is

$$\left|(a,b) - \left(\frac{a+b}{2}, \frac{a+b}{2}\right)\right|_\infty = \frac{b-a}{2}, \qquad \left|(a',b') - \left(\frac{a'+b'}{2}, \frac{a'+b'}{2}\right)\right|_\infty = \frac{b'-a'}{2}$$

The total cost is $\operatorname{cost}(M) = \max\left\{\frac{b-a}{2}, \frac{b'-a'}{2}\right\}$.

Second maching: $M' = \left\{\left((a,b),(a',b')\right)\right\}$. The intervals are matched together, and the cost of the pair is

$$|(a,b) - (a',b')|_\infty = \max\{|a-a'|, |b-b'|\}.$$

which is also $\operatorname{cost}(M')$.

These are the only two partial matchings, and we deduce the bottleneck distance

$$\mathrm{d_b}(P,Q) = \min\left\{\max\left\{\frac{b-a}{2}, \frac{b'-a'}{2}\right\}, \max\{|a-a'|, |b-b'|\}\right\}.$$

Example: Consider $a \leq b$, $a' \leq b'$, and the barcodes $P = \{[a,b]\}$ and $Q = \{[a',b']\}$.

We have

$$\mathrm{d_b}\,(P,Q) = \min\left\{\max\left\{\frac{b-a}{2}, \frac{b'-a'}{2}\right\}, \max\{|a-a'|, |b-b'|\}\right\}.$$

Example: Consider the interval-modules $\mathbb{B}[a,b]$ and $\mathbb{B}[a',b']$.

Their barcodes are the sets P and Q of the previous example, from which we deduce

$$\mathrm{d_b}\,(\mathbb{B}[a,b], \mathbb{B}[a',b']) = \min\left\{\max\left\{\frac{b-a}{2}, \frac{b'-a'}{2}\right\}, \max\{|a-a'|, |b-b'|\}\right\}.$$

Distância de entrelaçamento

Consider two persistence modules $\mathbb{V}$ and $\mathbb{W}$:

$$\dashrightarrow V^{t_1} \xrightarrow{v_{t_1}^{t_2}} V^{t_2} \xrightarrow{v_{t_2}^{t_3}} V^{t_3} \xrightarrow{v_{t_3}^{t_4}} V^{t_4} \text{ - - - - - - - - - -}$$

$$\dashrightarrow W^{t_1} \xrightarrow{w_{t_1}^{t_2}} W^{t_2} \xrightarrow{w_{t_2}^{t_3}} W^{t_3} \xrightarrow{w_{t_3}^{t_4}} W^{t_4} \text{ - - - - - - - - - -}$$

Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

Distância de entrelaçamento

Consider two persistence modules $\mathbb{V}$ and $\mathbb{W}$:

$$\dashrightarrow V^{t_1} \xrightarrow{v_{t_1}^{t_2}} V^{t_2} \xrightarrow{v_{t_2}^{t_3}} V^{t_3} \xrightarrow{v_{t_3}^{t_4}} V^{t_4} \text{ - - - - - - - -}$$

$$\dashrightarrow W^{t_1} \xrightarrow{w_{t_1}^{t_2}} W^{t_2} \xrightarrow{w_{t_2}^{t_3}} W^{t_3} \xrightarrow{w_{t_3}^{t_4}} W^{t_4} \text{ - - - - - - - -}$$

Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow \phi_s & & \downarrow \phi_t \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

An ϵ-**interleaving** between $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ and $(\psi_t \colon W^t \to V^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagrams commute for every $t \in \mathbb{R}^+$:

$$\begin{array}{ccccc} V^t & & \xrightarrow{v_t^{t+2\epsilon}} & & V^{t+2\epsilon} \\ & \searrow \phi_t & & \nearrow \psi_{t+\epsilon} & \\ & & W^{t+\epsilon} & & \end{array} \qquad \begin{array}{ccccc} & & V^{t+\epsilon} & & \\ & \nearrow \psi_t & & \searrow \phi_{t+\epsilon} & \\ W^t & & \xrightarrow{w_t^{t+2\epsilon}} & & W^{t+2\epsilon} \end{array}$$

Distância de entrelaçamento

Consider two persistence modules $\mathbb{V}$ and $\mathbb{W}$:

$$\dashrightarrow V^{t_1} \xrightarrow{v_{t_1}^{t_2}} V^{t_2} \xrightarrow{v_{t_2}^{t_3}} V^{t_3} \xrightarrow{v_{t_3}^{t_4}} V^{t_4} \dashrightarrow$$

$$\dashrightarrow W^{t_1} \xrightarrow{w_{t_1}^{t_2}} W^{t_2} \xrightarrow{w_{t_2}^{t_3}} W^{t_3} \xrightarrow{w_{t_3}^{t_4}} W^{t_4} \dashrightarrow$$

Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t\in\mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

An ϵ-**interleaving** between $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t \colon V^t \to W^{t+\epsilon})_{t\in\mathbb{R}^+}$ and $(\psi_t \colon W^t \to V^{t+\epsilon})_{t\in\mathbb{R}^+}$ such that the following diagrams commute for every $t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^t & \xrightarrow{v_t^{t+2\epsilon}} & V^{t+2\epsilon} \\ & {\scriptstyle \phi_t} \searrow \quad \nearrow {\scriptstyle \psi_{t+\epsilon}} & \\ & W^{t+\epsilon} & \end{array} \qquad \begin{array}{ccc} & V^{t+\epsilon} & \\ & {\scriptstyle \psi_t} \nearrow \quad \searrow {\scriptstyle \phi_{t+\epsilon}} & \\ W^t & \xrightarrow{w_t^{t+2\epsilon}} & W^{t+2\epsilon} \end{array}$$

The **interleaving distance** is: $\mathrm{d_i}(\mathbb{V},\mathbb{W}) = \inf\{\epsilon \geq 0 \mid \mathbb{V} \text{ and } \mathbb{W} \text{ are } \epsilon\text{-interleaved}\}$.

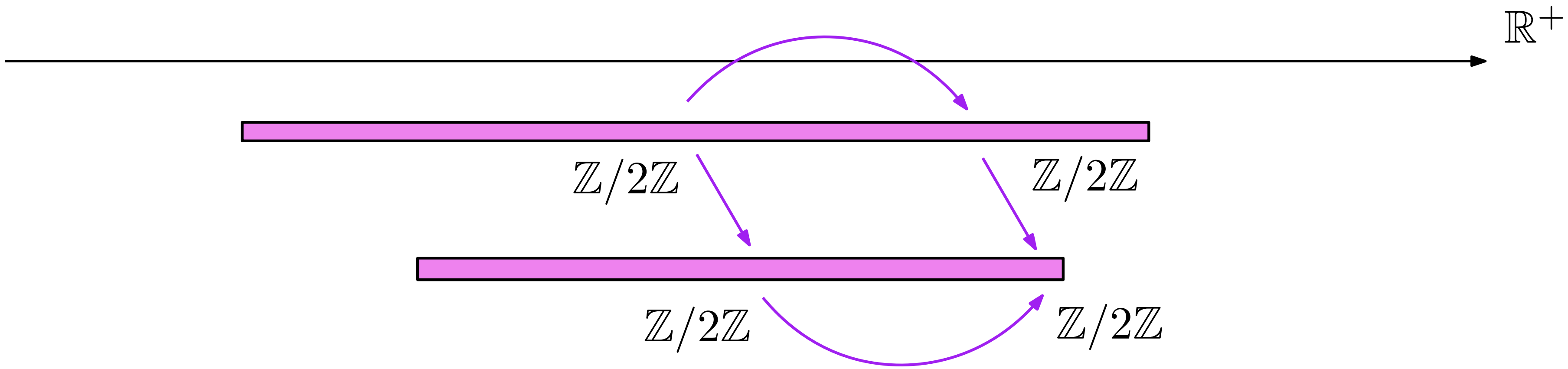
Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a,b]$ and $\mathbb{B}[a',b']$.

Let us find an ϵ-interleaving.

Distância de entrelaçamento

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a, b]$ and $\mathbb{B}[a', b']$.

Let us find an ϵ-interleaving.

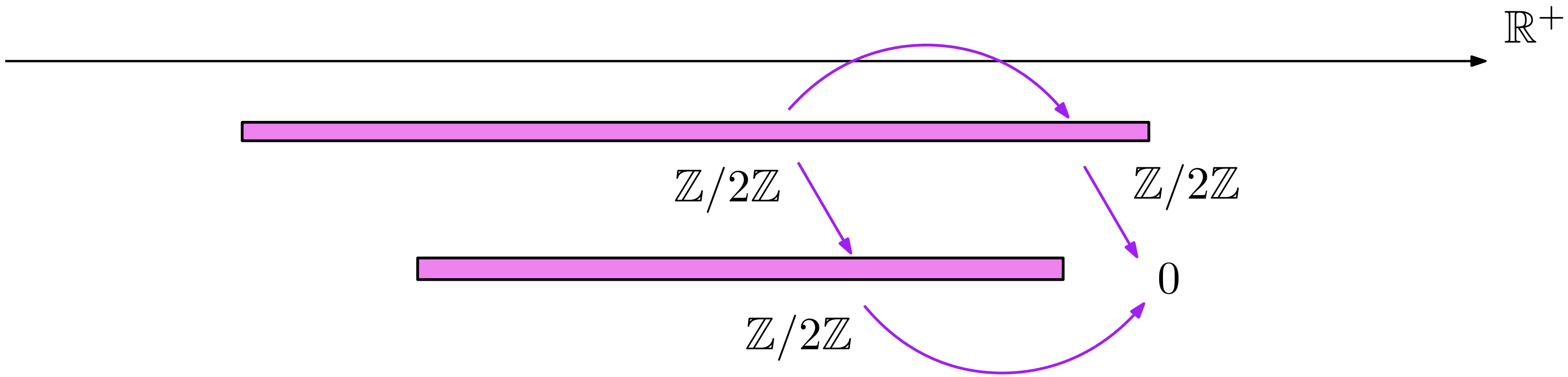


Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a,b]$ and $\mathbb{B}[a',b']$.

Let us find an ϵ-interleaving.



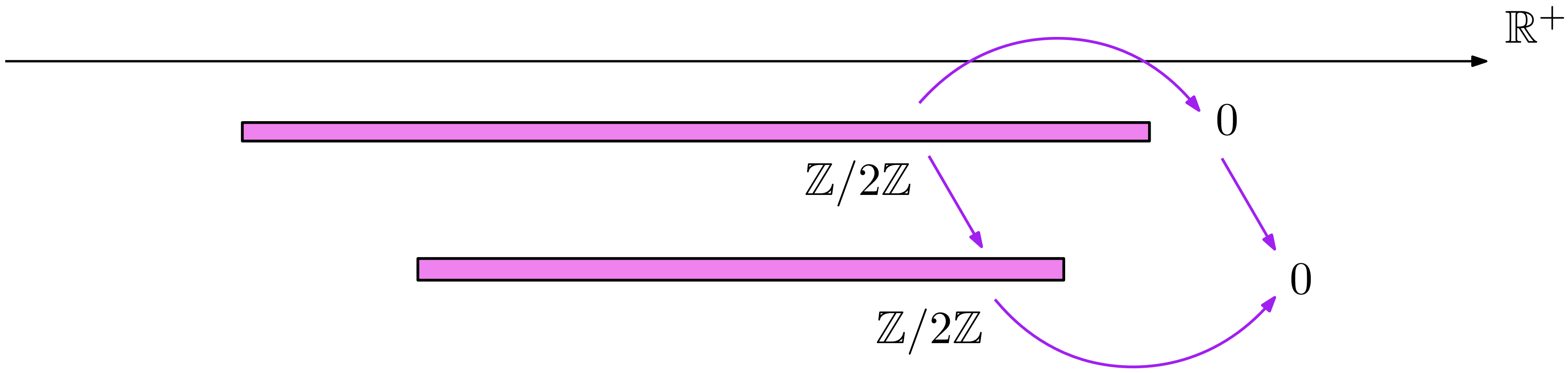
Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t\in\mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

Distância de entrelaçamento

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a, b]$ and $\mathbb{B}[a', b']$.

Let us find an ϵ-interleaving.



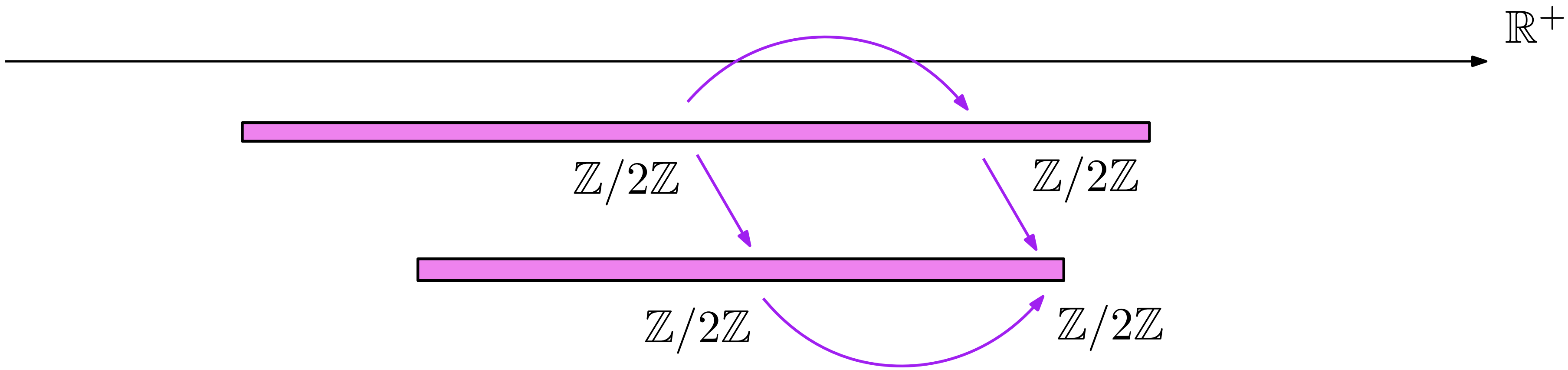
Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

Distância de entrelaçamento

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a, b]$ and $\mathbb{B}[a', b']$.

Let us find an ϵ-interleaving.



Given $\epsilon \geq 0$, an ϵ-**morphism** between $\mathbb{V}$ and $\mathbb{W}$ is a family of linear maps $\phi = (\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagram commutes for every $s \leq t \in \mathbb{R}^+$:

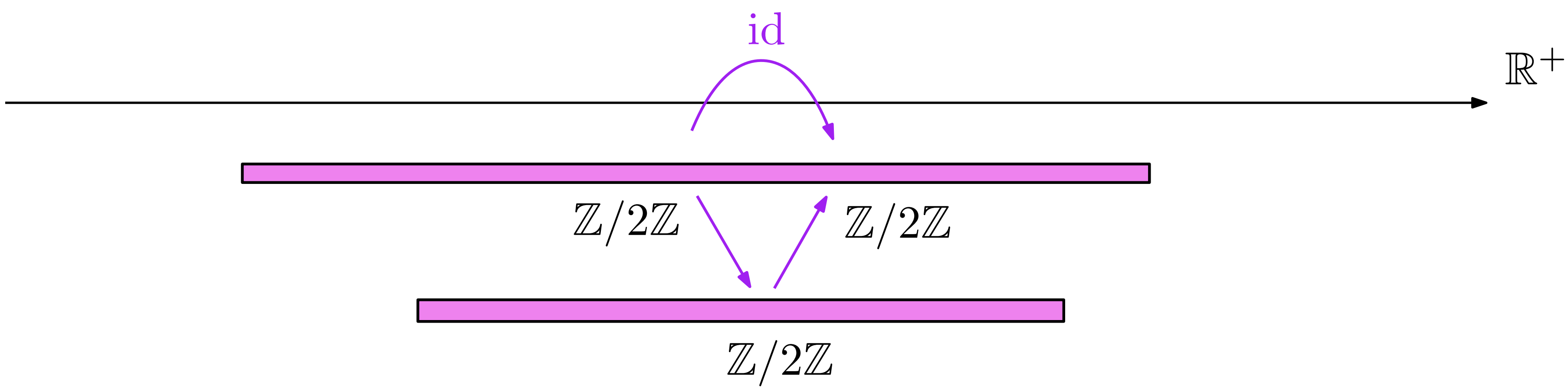
$$\begin{array}{ccc} V^s & \xrightarrow{v_s^t} & V^t \\ \downarrow{\scriptstyle \phi_s} & & \downarrow{\scriptstyle \phi_t} \\ W^{s+\epsilon} & \xrightarrow{w_{s+\epsilon}^{t+\epsilon}} & W^{t+\epsilon} \end{array}$$

$\longrightarrow$ Only two possibilities for ϕ:

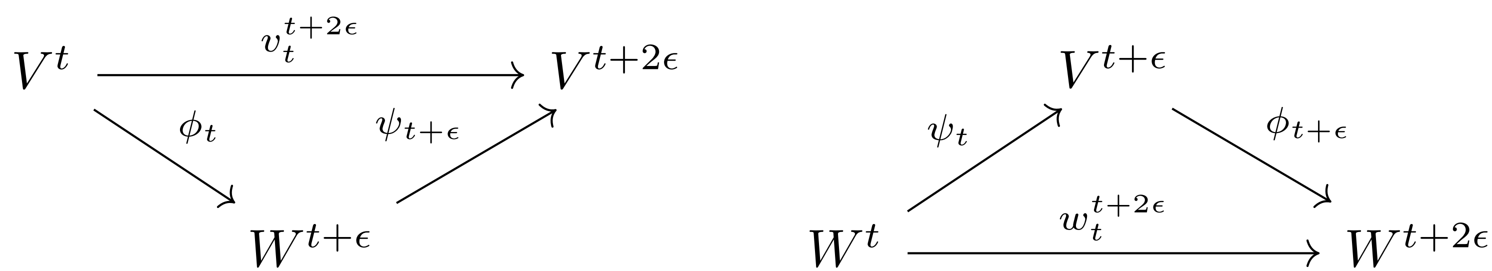
- always the zero map
- always nonzero when V^t and $W^{t+\epsilon}$ are nonzero

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a, b]$ and $\mathbb{B}[a', b']$.

Let us find an ϵ-interleaving.



An ϵ-**interleaving** between $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ and $(\psi_t \colon W^t \to V^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagrams commute for every $t \in \mathbb{R}^+$:

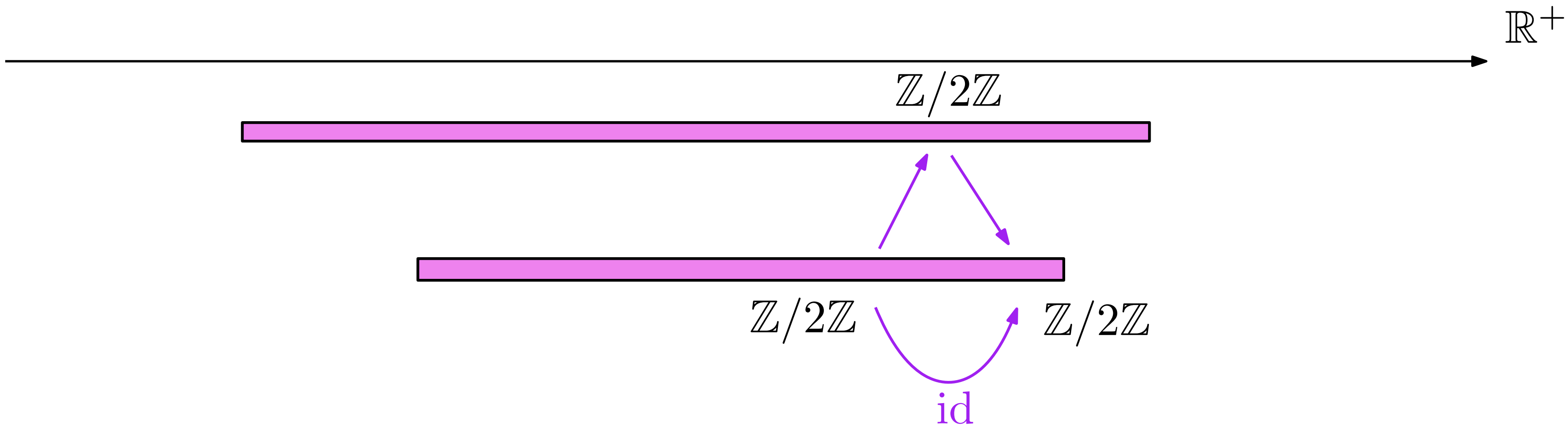


→ $\psi_{t+\epsilon} \circ \phi_t$ must be nonzero when $[t, t+\epsilon] \subset [a, b]$

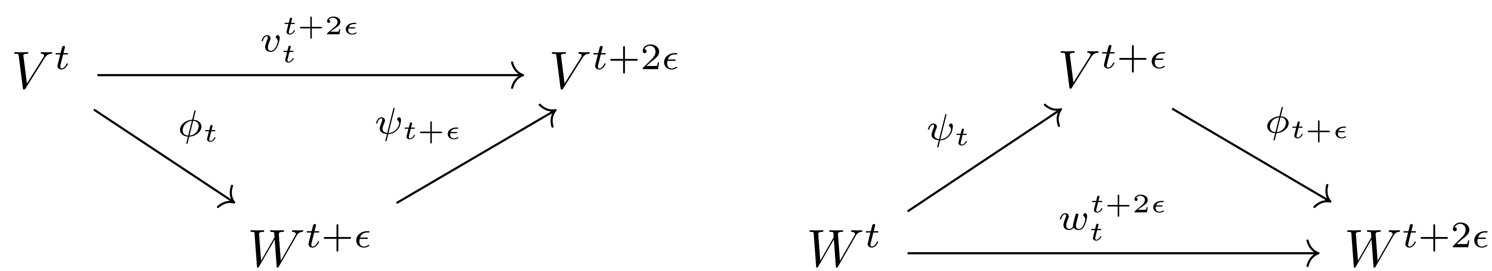
Distância de entrelaçamento

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a, b]$ and $\mathbb{B}[a', b']$.

Let us find an ϵ-interleaving.



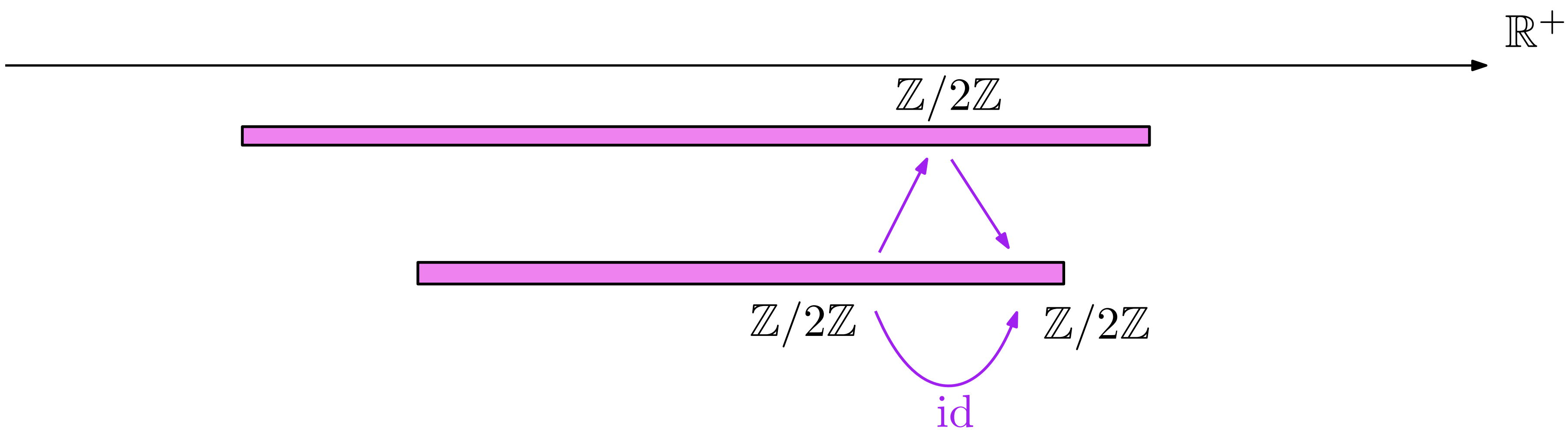
An ϵ-**interleaving** between $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t : V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ and $(\psi_t : W^t \to V^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagrams commute for every $t \in \mathbb{R}^+$:



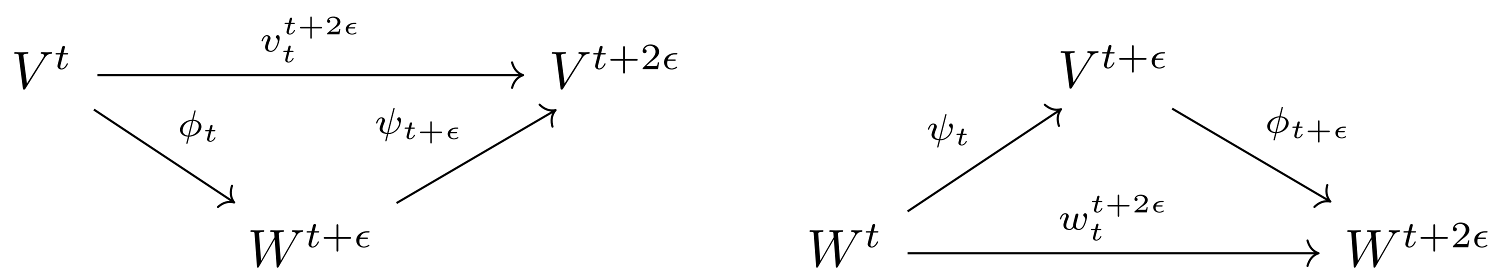
→ $\psi_{t+\epsilon} \circ \phi_t$ must be nonzero when $[t, t+\epsilon] \subset [a, b]$

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a,b]$ and $\mathbb{B}[a',b']$.

Let us find an ϵ-interleaving.



An ϵ-**interleaving** between $\mathbb{V}$ and $\mathbb{W}$ is a pair of ϵ-morphisms $(\phi_t \colon V^t \to W^{t+\epsilon})_{t \in \mathbb{R}^+}$ and $(\psi_t \colon W^t \to V^{t+\epsilon})_{t \in \mathbb{R}^+}$ such that the following diagrams commute for every $t \in \mathbb{R}^+$:



→ $\psi_{t+\epsilon} \circ \phi_t$ must be nonzero when $[t, t+\epsilon] \subset [a,b]$
$\phi_{t+\epsilon} \circ \psi_t$ must be nonzero when $[t, t+\epsilon] \subset [a',b']$

Distância de entrelaçamento

Example: Let $a, a', b, b' \in \mathbb{R}^+$ such that $a \leq b$ and $a' \leq b'$. Consider the interval-modules $\mathbb{B}[a, b]$ and $\mathbb{B}[a', b']$.

Let us find an ϵ-interleaving.

→ Only two possibilities for ϕ:
- always the zero map
- always nonzero when V^t and $W^{t+\epsilon}$ are nonzero

→ $\psi_{t+\epsilon} \circ \phi_t$ must be nonzero when $[t, t+\epsilon] \subset [a, b]$
$\phi_{t+\epsilon} \circ \psi_t$ must be nonzero when $[t, t+\epsilon] \subset [a', b']$

We deduce that either
- $|a - b| \leq 2\epsilon$ and $|a' - b'| \leq 2\epsilon$, or
- $|a - a'| \leq \epsilon$ and $|b - b'| \leq \epsilon$

Conclusion: $$\mathrm{d_i}\left(\mathbb{B}[a, b], \mathbb{B}[a', b']\right) = \min\left\{\max\left\{\frac{b-a}{2}, \frac{b'-a'}{2}\right\}, \max\{|a-a'|, |b-b'|\}\right\}$$

Teorema de isometria

Theorem (Chazal, de Silva, Glisse, Oudot, 2009): If the persistence modules $\mathbb{U}$ and $\mathbb{V}$ are interval-decomposable, then $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) = \mathrm{d_b}(\mathbb{U}, \mathbb{V})$.

⟶ **Stability**: $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \geq \mathrm{d_b}(\mathbb{U}, \mathbb{V})$

⟶ **Converse stability**: $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \leq \mathrm{d_b}(\mathbb{U}, \mathbb{V})$

Teorema de isometria

Theorem (Chazal, de Silva, Glisse, Oudot, 2009): If the persistence modules $\mathbb{U}$ and $\mathbb{V}$ are interval-decomposable, then $\mathrm{d_i}\left(\mathbb{U},\mathbb{V}\right) = \mathrm{d_b}\left(\mathbb{U},\mathbb{V}\right)$.

Proof: Let us write the decomposition of the persistence modules in intervals:

$$\mathbb{V} \simeq \bigoplus_{I\in\mathcal{I}} \mathbb{B}[I] \qquad\qquad \mathbb{W} \simeq \bigoplus_{J\in\mathcal{J}} \mathbb{B}[J]$$

Suppose that we have a ϵ-partial matching $M \subset \mathcal{I}\times\mathcal{J}$. This gives a matching of some intervals (I, J), where $I = (a, b)$ and $J = (a', b')$, such that $|a - a'| \leq \epsilon$ and $|b - b'| \leq \epsilon$.

We can build an ϵ-interleaving between $\mathbb{B}[I]$ and $\mathbb{B}[J]$, that we denote $(\phi_{(I,J)}, \psi_{(I,J)})$.

Some intervals I (resp. J) are not matched, in which case their length is not greater than 2ϵ, and we can build an ϵ-interleaving with the zero persistence module. We denote this interleaving $(\phi_{(I,0)}, \psi_{(I,0)})$ (resp. $(\phi_{(0,J)}, \psi_{(0,J)})$).

Now, let us consider the sums of all these linear maps:

$$\overline{\phi} = \bigoplus_{(I,J)\ \text{matched}} \phi_{(I,J)} \bigoplus_{I\ \text{not matched}} \phi_{(I,0)}, \qquad \overline{\psi} = \bigoplus_{(I,J)\ \text{matched}} \psi_{(I,J)} \bigoplus_{J\ \text{not matched}} \phi_{(0,J)}$$

⟶ $(\overline{\phi}, \overline{\psi})$ is an ϵ-interleaving ⟶ $\mathrm{d_i}\left(\mathbb{U},\mathbb{V}\right) \leq \mathrm{d_b}\left(\mathbb{U},\mathbb{V}\right)$

Teorema de isometria

Theorem (Chazal, de Silva, Glisse, Oudot, 2009): If the persistence modules $\mathbb{U}$ and $\mathbb{V}$ are interval-decomposable, then $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) = \mathrm{d_b}(\mathbb{U}, \mathbb{V})$.

→ **Stability**: $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \geq \mathrm{d_b}(\mathbb{U}, \mathbb{V})$

→ **Converse stability**: $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \leq \mathrm{d_b}(\mathbb{U}, \mathbb{V})$

The stability part is more difficult.

A first strategy uses the interpolation lemma, and concludes with the box lemma.

Interpolation lemma: If $\mathbb{U}$ and $\mathbb{V}$ are δ-interleaved, then there exists a family of persistence modules $(\mathbb{U}_t)_{t\in[0,\delta]}$ such that $\mathbb{U}_0 = \mathbb{U}$, $\mathbb{U}_\delta = \mathbb{V}$ and $\mathrm{d_i}(\mathbb{U}_s, \mathbb{U}_t) \leq |s-t|$ for every $s, t \in [0, \delta]$.

Another proof builds an explicit partial matching from an interleaving (Bauer, Lesnick, 2013).

I - Decomposition of persistence modules

1 - Simplicial filtrations

2 - Persistence modules

3 - Barcodes

II - Stability of persistence modules

1 - Bottleneck distance

2 - Interleaving distance

3 - Stability

III - Persistent homology in practice

1 - Data analysis

2 - Machine learning

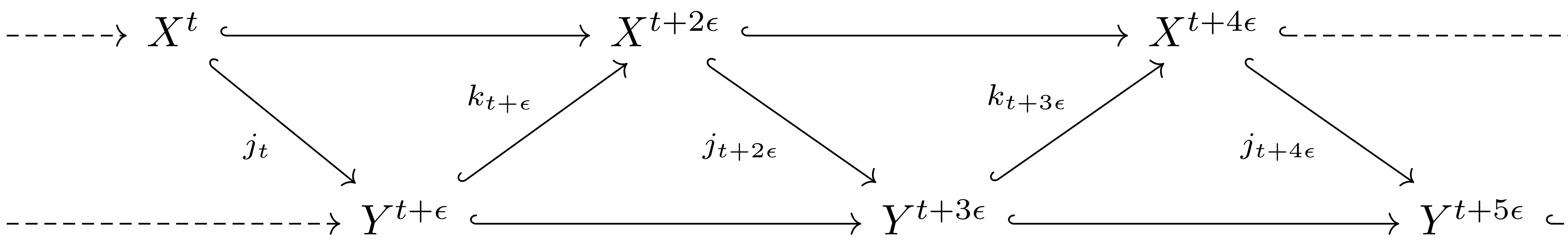
3 - Variations on persistent homology

De volta aos espessamentos

Let X and Y be two subets of $\mathbb{R}^n$. Define $\epsilon = \mathrm{d_H}(X, Y)$ (Hausdorff distance).

We have seen that $X \subset Y^\epsilon$ and $Y \subset X^\epsilon$. We even have that $X^t \subset Y^{t+\epsilon}$ and $Y^t \subset X^{t+\epsilon}$ for all $t \geq 0$.

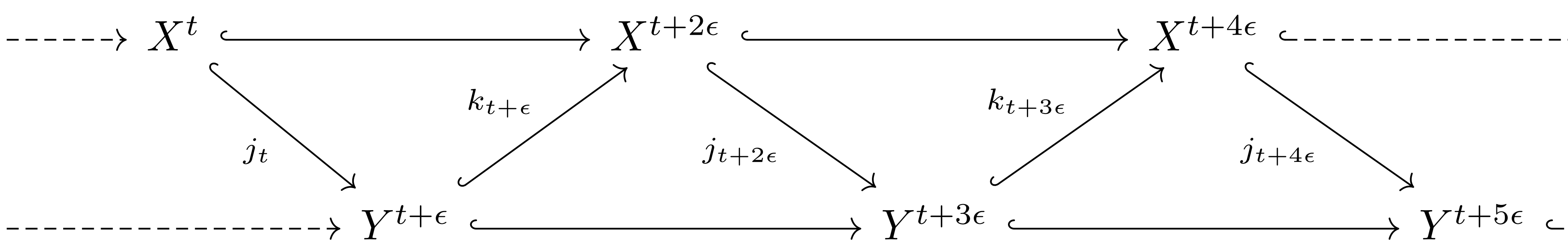
By denoting j and k these inclusions, we have a commutative diagram



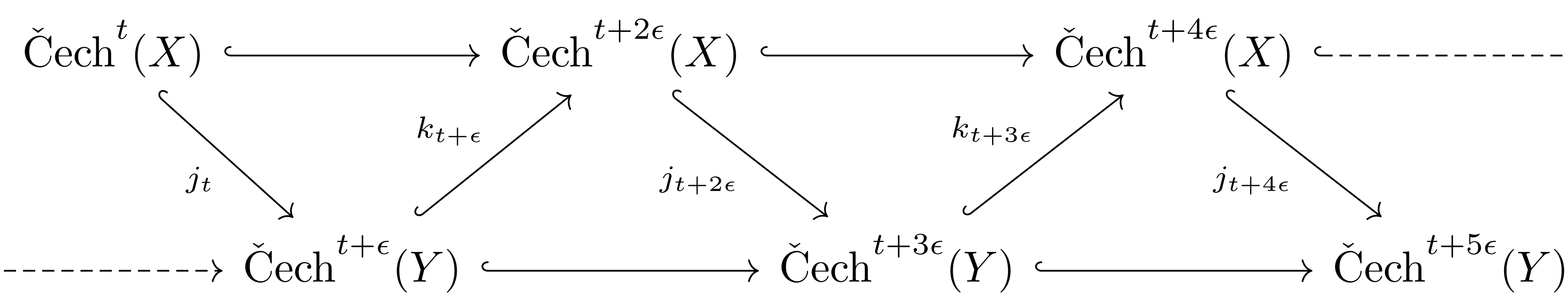
Let X and Y be two subets of $\mathbb{R}^n$. Define $\epsilon = \mathrm{d_H}(X, Y)$ (Hausdorff distance).

We have seen that $X \subset Y^\epsilon$ and $Y \subset X^\epsilon$. We even have that $X^t \subset Y^{t+\epsilon}$ and $Y^t \subset X^{t+\epsilon}$ for all $t \geq 0$.

By denoting j and k these inclusions, we have a commutative diagram



This also gives inclusions between Čech complexes:



De volta aos espessamentos

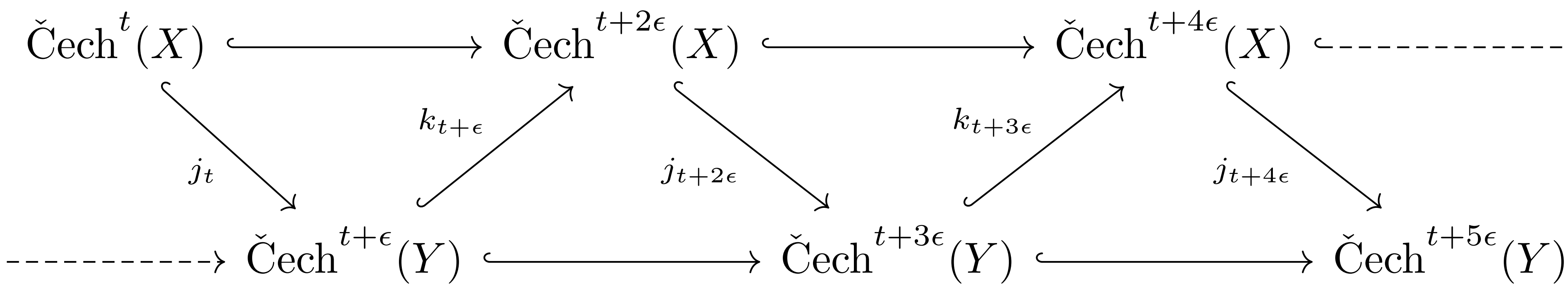
[...] This also gives inclusions between Čech complexes:

$$
\begin{array}{ccccccccccc}
\check{C}\mathrm{ech}^{t}(X) & & \hookrightarrow & & \check{C}\mathrm{ech}^{t+2\epsilon}(X) & & \hookrightarrow & & \check{C}\mathrm{ech}^{t+4\epsilon}(X) & & \hookrightarrow \cdots \\
& \searrow^{j_t} & & \nearrow^{k_{t+\epsilon}} & & \searrow^{j_{t+2\epsilon}} & & \nearrow^{k_{t+3\epsilon}} & & \searrow^{j_{t+4\epsilon}} & \\
\cdots \to & & \check{C}\mathrm{ech}^{t+\epsilon}(Y) & & \hookrightarrow & & \check{C}\mathrm{ech}^{t+3\epsilon}(Y) & & \hookrightarrow & & \check{C}\mathrm{ech}^{t+5\epsilon}(Y)
\end{array}
$$

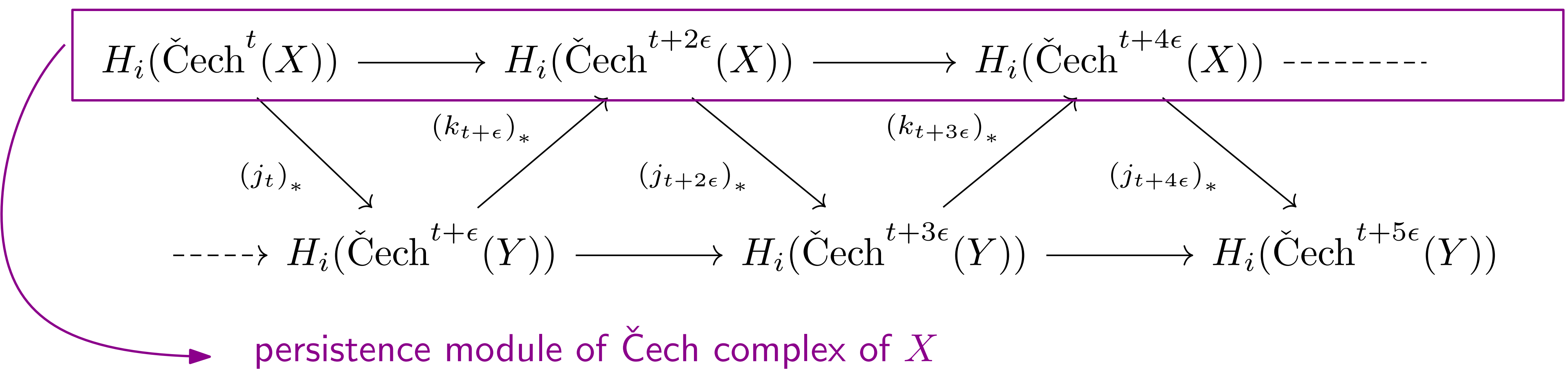
Now, we apply the i^{th} homology functor.

$$
\begin{array}{ccccccccccc}
H_i(\check{C}\mathrm{ech}^{t}(X)) & & \longrightarrow & & H_i(\check{C}\mathrm{ech}^{t+2\epsilon}(X)) & & \longrightarrow & & H_i(\check{C}\mathrm{ech}^{t+4\epsilon}(X)) & & \cdots \\
& \searrow^{(j_t)_*} & & \nearrow^{(k_{t+\epsilon})_*} & & \searrow^{(j_{t+2\epsilon})_*} & & \nearrow^{(k_{t+3\epsilon})_*} & & \searrow^{(j_{t+4\epsilon})_*} & \\
\cdots \to & & H_i(\check{C}\mathrm{ech}^{t+\epsilon}(Y)) & & \longrightarrow & & H_i(\check{C}\mathrm{ech}^{t+3\epsilon}(Y)) & & \longrightarrow & & H_i(\check{C}\mathrm{ech}^{t+5\epsilon}(Y))
\end{array}
$$

[...] This also gives inclusions between Čech complexes:

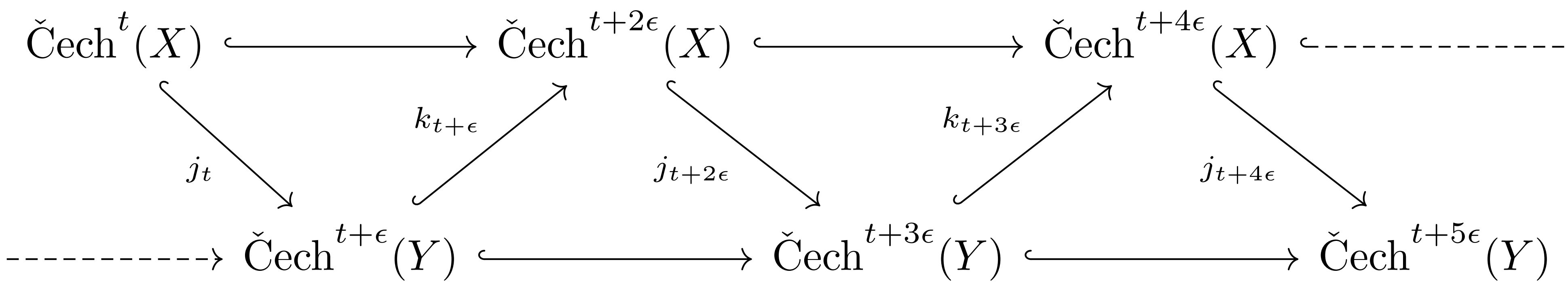


Now, we apply the i^{th} homology functor.

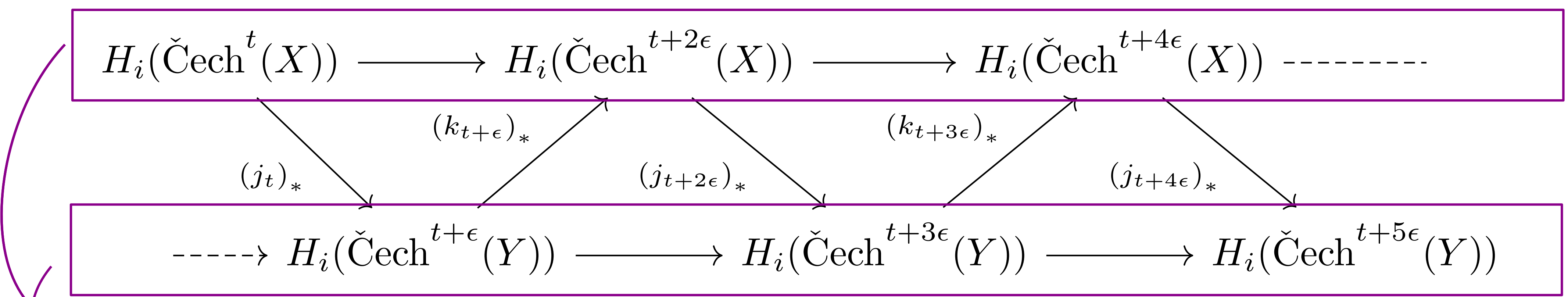


De volta aos espessamentos

[...] This also gives inclusions between Čech complexes:



Now, we apply the i^{th} homology functor.

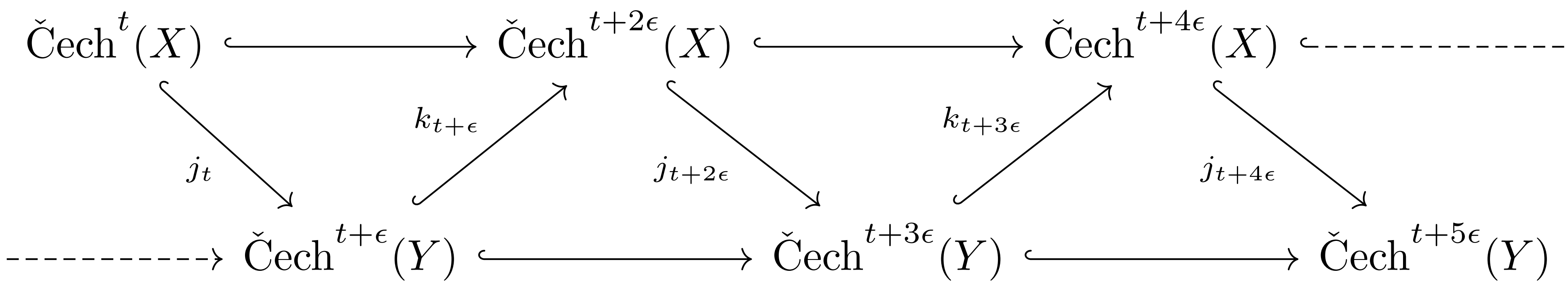


persistence module of Čech complex of X

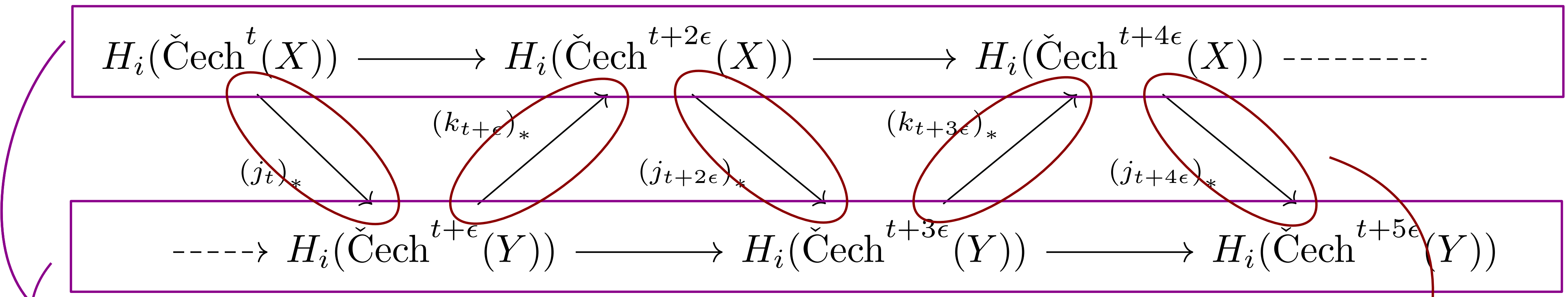
persistence module of Čech complex of Y

De volta aos espessamentos

[...] This also gives inclusions between Čech complexes:



Now, we apply the i^{th} homology functor.



persistence module of Čech complex of X

persistence module of Čech complex of Y

ϵ-interleaving between the persistence modules

De volta aos espessamentos

[...]

Hence the persistence modules $\left(H_i(\check{\mathrm{C}}\mathrm{ech}^t(X))\right)_{t\geq 0}$ and $\left(H_i(\check{\mathrm{C}}\mathrm{ech}^t(Y))\right)_{t\geq 0}$ are ϵ-interleaved.

$\mathbb{U}$ $\mathbb{V}$

Hence $\mathrm{d_i}\left(\mathbb{U}, \mathbb{V}\right) \leq \epsilon$.

De volta aos espessamentos

[...]

Hence the persistence modules $\left(H_i(\check{\mathrm{C}}\mathrm{ech}^t(X))\right)_{t\geq 0}$ and $\left(H_i(\check{\mathrm{C}}\mathrm{ech}^t(Y))\right)_{t\geq 0}$ are ϵ-interleaved.

$\mathbb{U}$ $\qquad$ $\mathbb{V}$

Hence $\mathrm{d_i}\left(\mathbb{U},\mathbb{V}\right)\leq\epsilon$.

We use the isometry theorem: $\mathrm{d_b}\left(\mathbb{U},\mathbb{V}\right)=\mathrm{d_i}\left(\mathbb{U},\mathbb{V}\right)\leq\epsilon$.

De volta aos espessamentos

[...]

Hence the persistence modules $\left(H_i(\check{\mathrm{C}}\mathrm{ech}^t(X))\right)_{t\geq 0}$ and $\left(H_i(\check{\mathrm{C}}\mathrm{ech}^t(Y))\right)_{t\geq 0}$ are ϵ-interleaved.

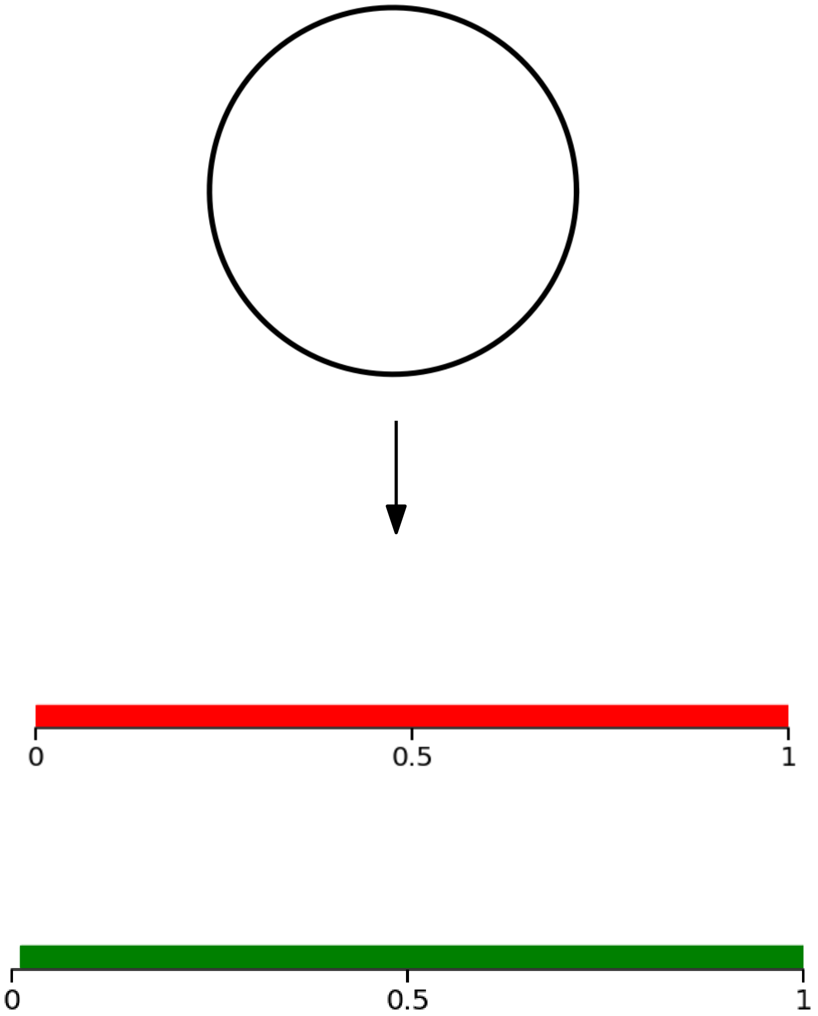
$\mathbb{U}$ $\qquad$ $\mathbb{V}$

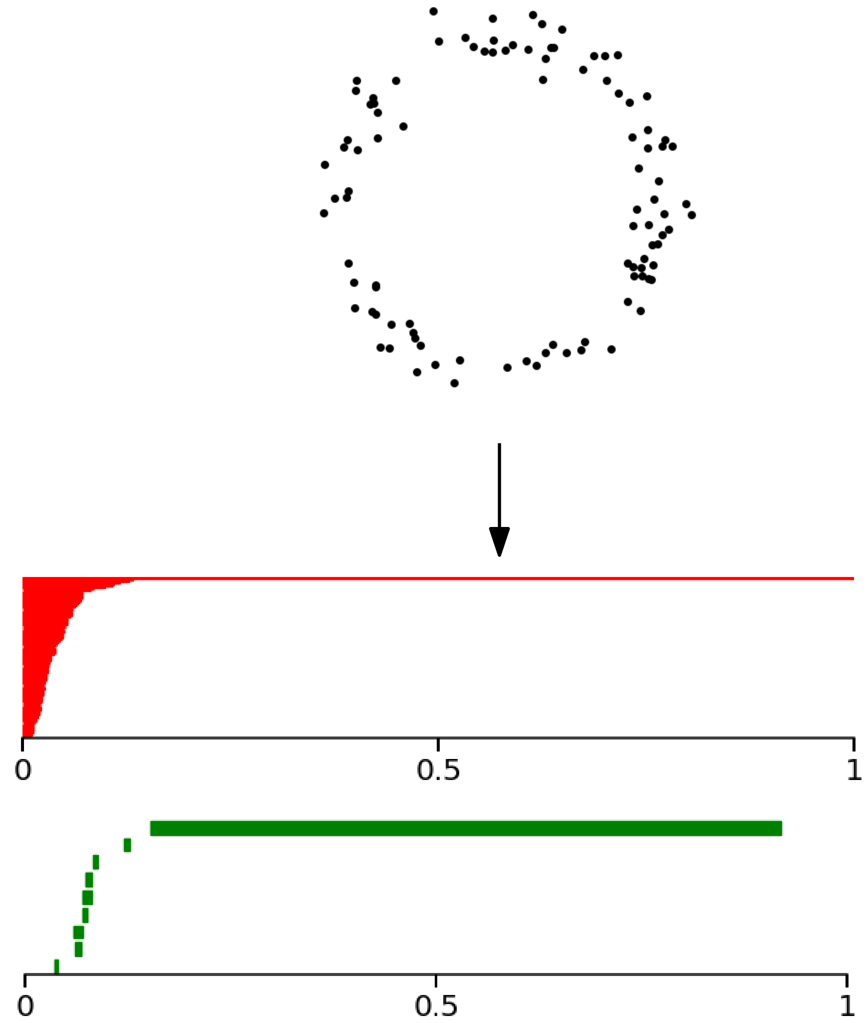
Hence $\mathrm{d_i}(\mathbb{U}, \mathbb{V}) \leq \epsilon$.

We use the isometry theorem: $\mathrm{d_b}(\mathbb{U}, \mathbb{V}) = \mathrm{d_i}(\mathbb{U}, \mathbb{V}) \leq \epsilon$.

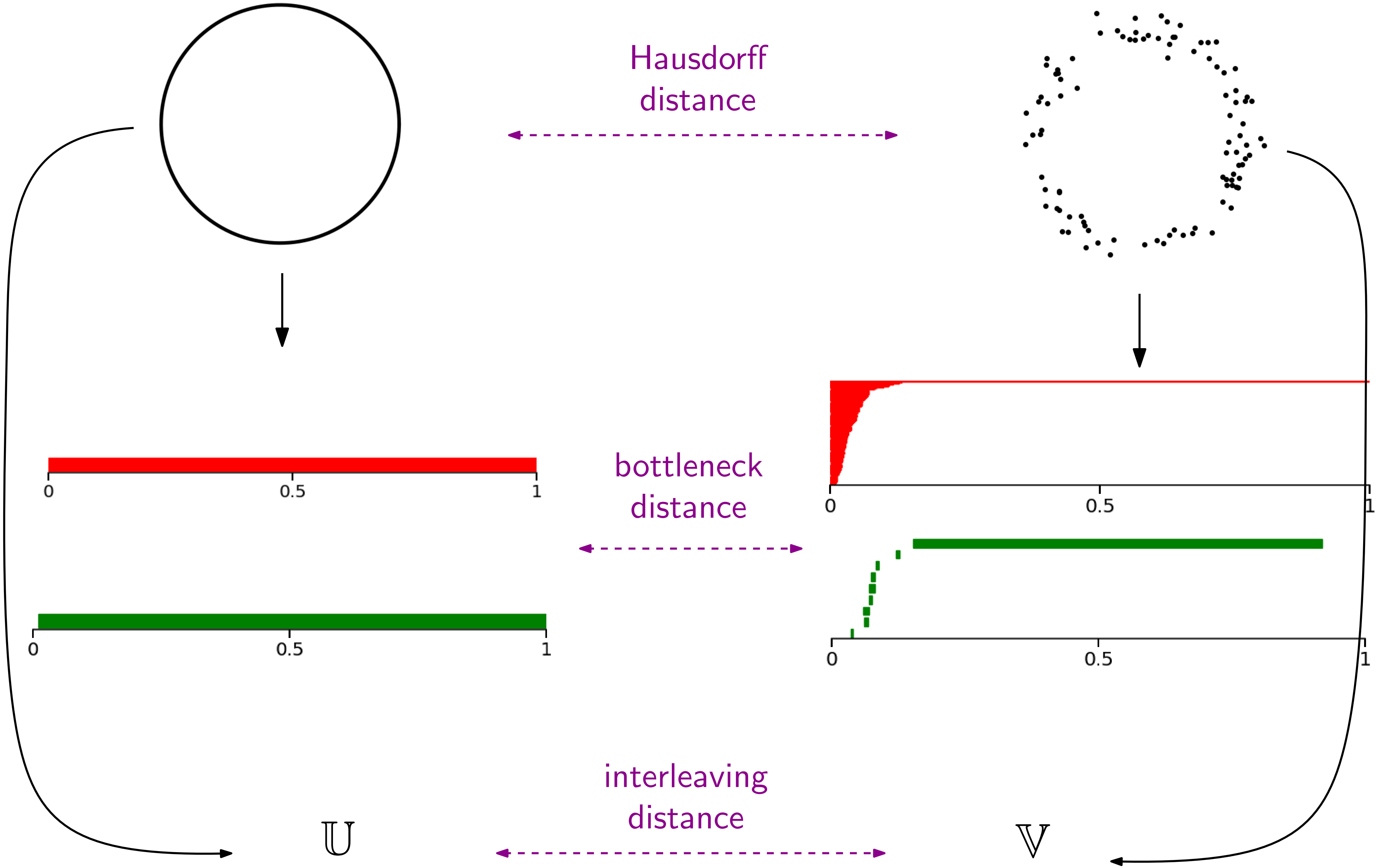
Theorem (Cohen-Steiner, Edelsbrunner, Harer, 2005): Let X and Y be two subsets of $\mathbb{R}^n$. Consider their Čech (resp. Rips) filtrations, and the corresponding i^{th} homology persistence modules, $\mathbb{U}$ and $\mathbb{V}$. Suppose that they are interval-decomposables. Then $\mathrm{d_b}(\mathbb{U}, \mathbb{V}) \leq \mathrm{d_H}(X, Y)$.

Resumo





Resumo


Hausdorff distance
0
0.5
1
0
0.5
1
bottleneck distance
0
0.5
1
0
0.5
1
interleaving distance
$\mathbb{U}$
$\mathbb{V}$

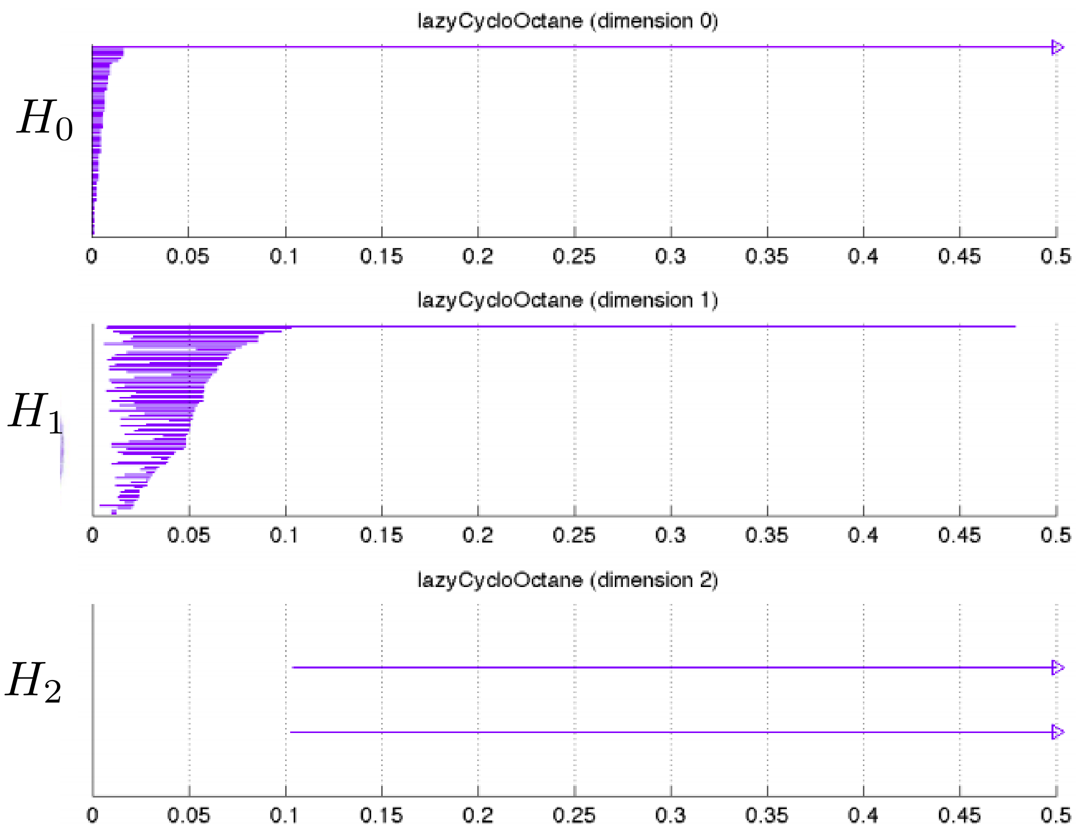
Topological inference I

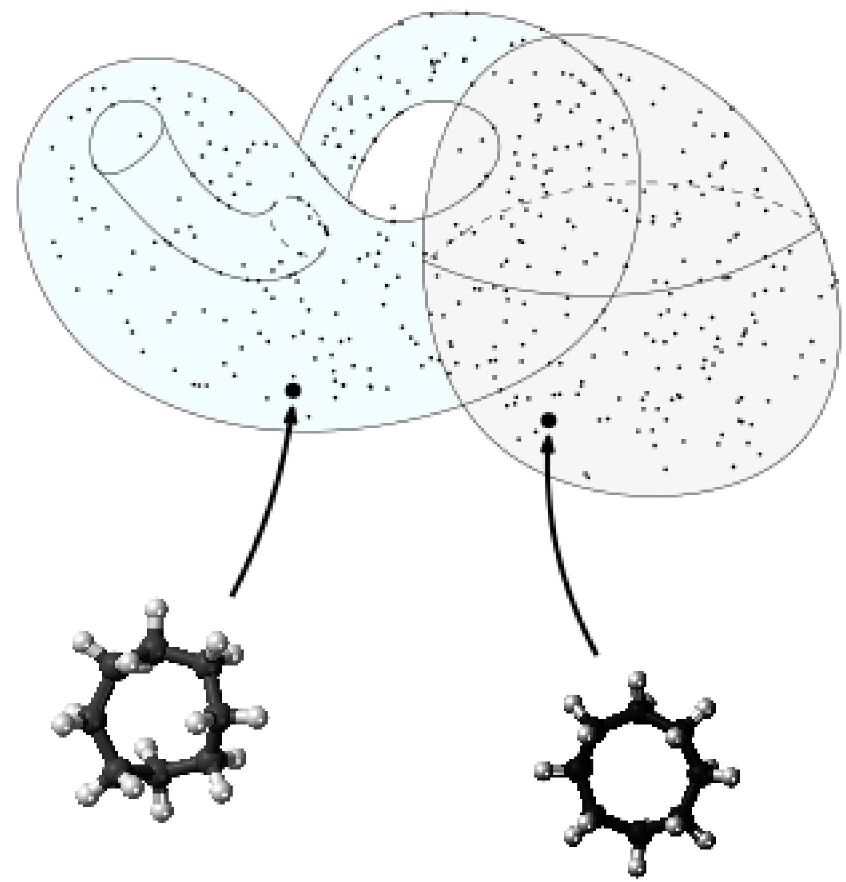
S. Martin, A. Thompson, E. A. Coutsias, and J-P. Watson, Topology of cyclo-octane energy landscape , 2010

https://www.researchgate.net/publication/44697030_Topology_of_Cyclooctane_Energy_Landscape

The cyclo-octane molecule C_8H_{16} contains 24 atoms.
By generating many of these molecules, we obtain a point cloud in $\mathbb{R}^{72}$ $(3 \times 24 = 72)$.

We obtain the barcodes:



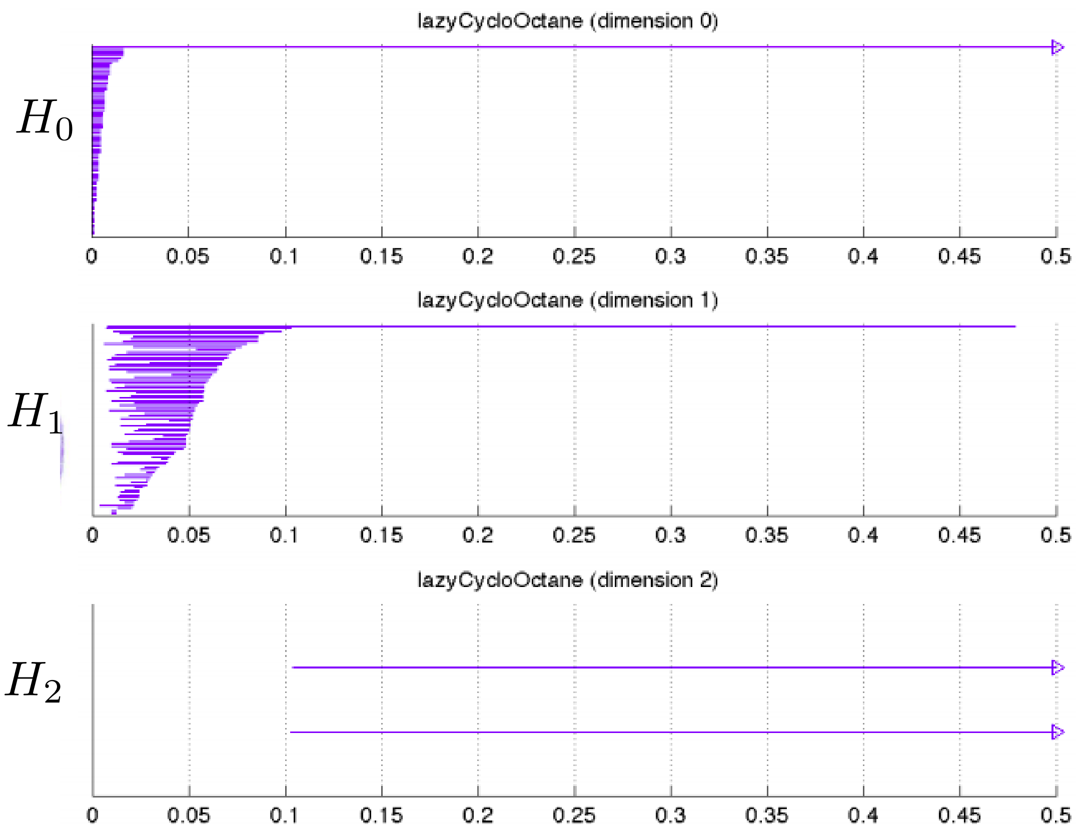


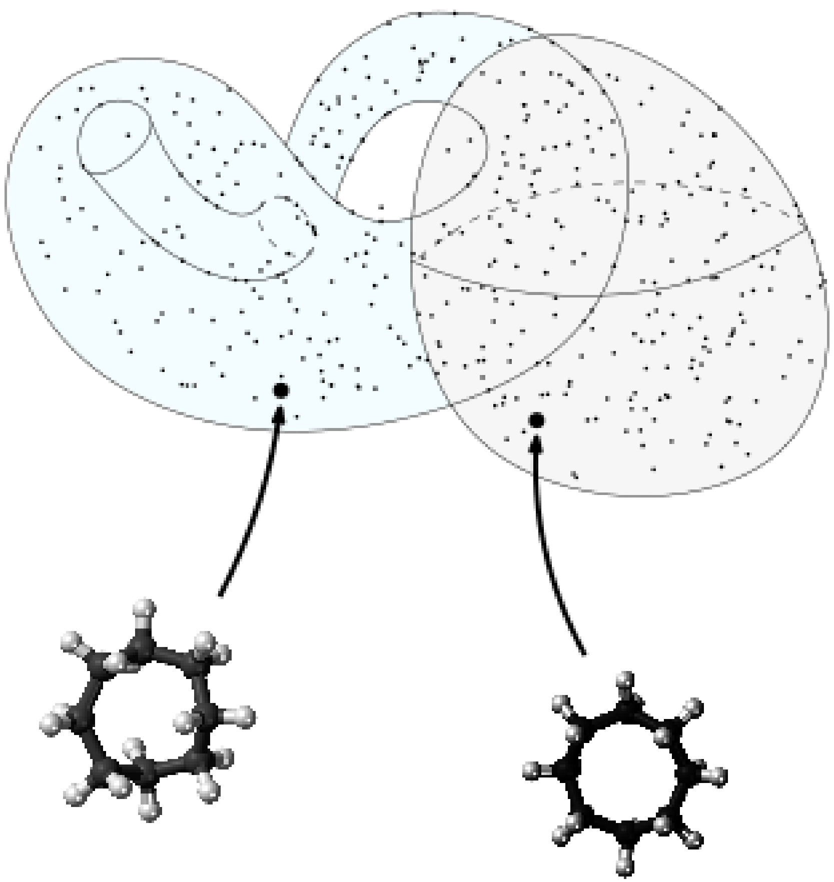
S. Martin, A. Thompson, E. A. Coutsias, and J-P. Watson, Topology of cyclo-octane energy landscape , 2010

https://www.researchgate.net/publication/44697030_Topology_of_Cyclooctane_Energy_Landscape

The cyclo-octane molecule C_8H_{16} contains 24 atoms.
By generating many of these molecules, we obtain a point cloud in $\mathbb{R}^{72}$ $(3 \times 24 = 72)$.

We obtain the barcodes:





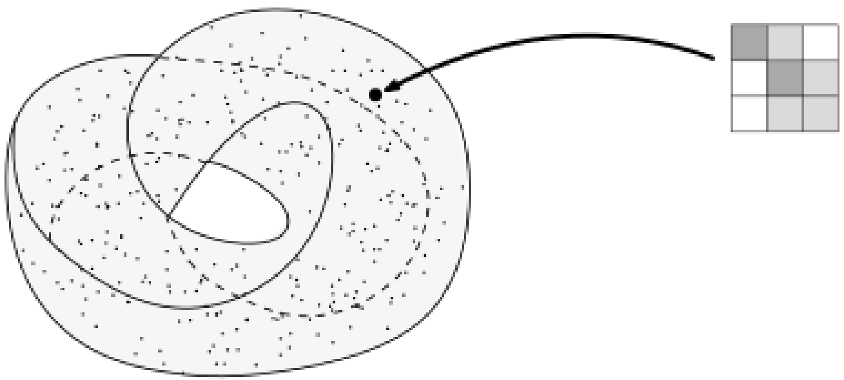
We deduce : $H_0 = \mathbb{Z}/2\mathbb{Z}, \quad H_1 = \mathbb{Z}/2\mathbb{Z}, \quad H_2 = (\mathbb{Z}/2\mathbb{Z})^2$

Topological inference II

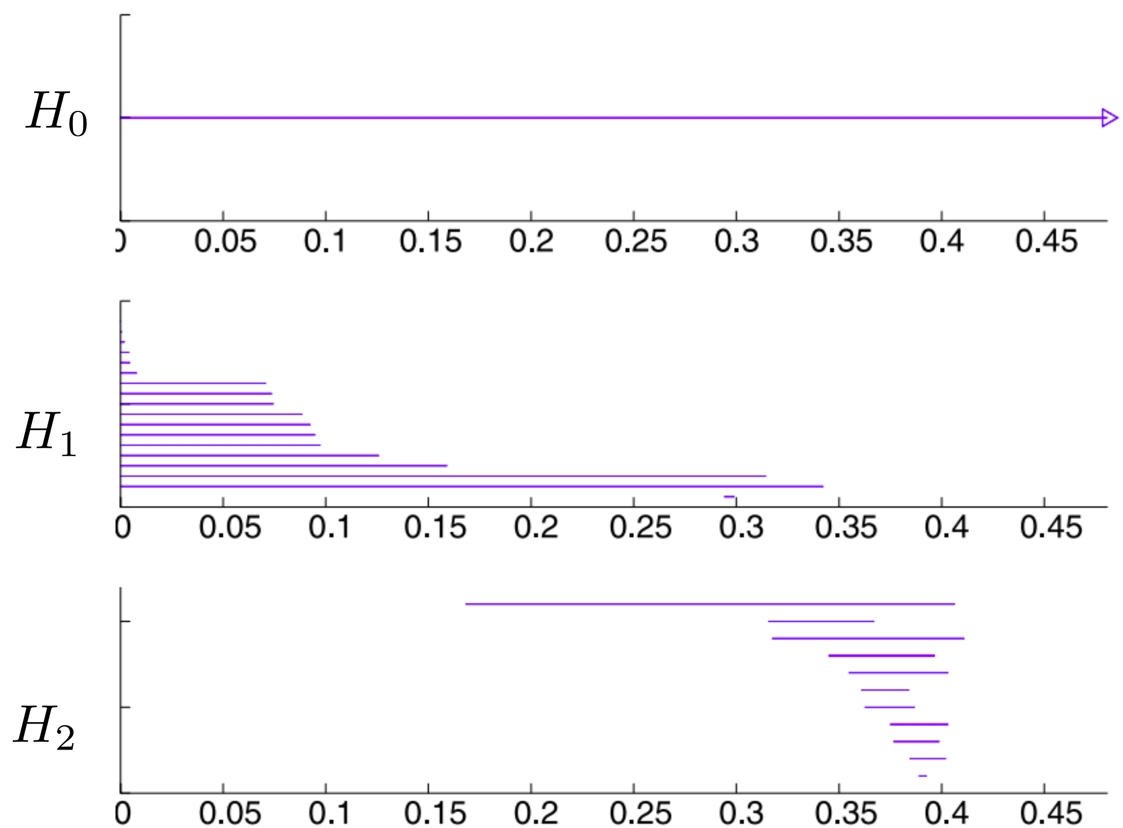
G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, On the Local Behavior of Spaces of Natural Images, 2008

https://link.springer.com/article/10.1007/s11263-007-0056-x

From a large collection of natural images, the authors extract 3×3 patches. Since it consists of 9 pixels, each of these patches can be seen as a 9-dimensional vector, and the whole set as a point cloud in $\mathbb{R}^9$.



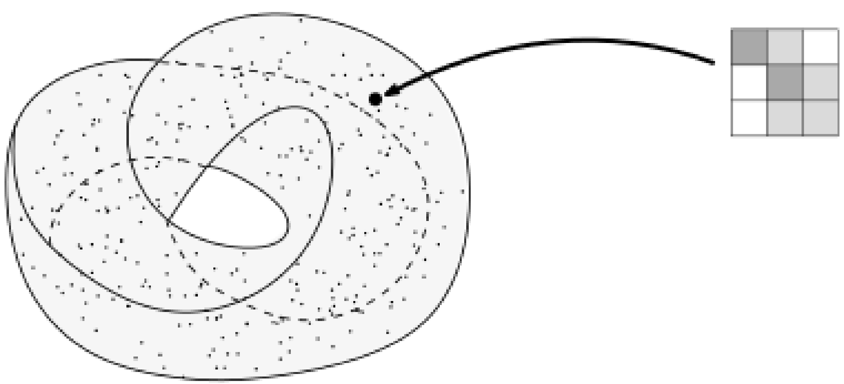
We get the barcodes:



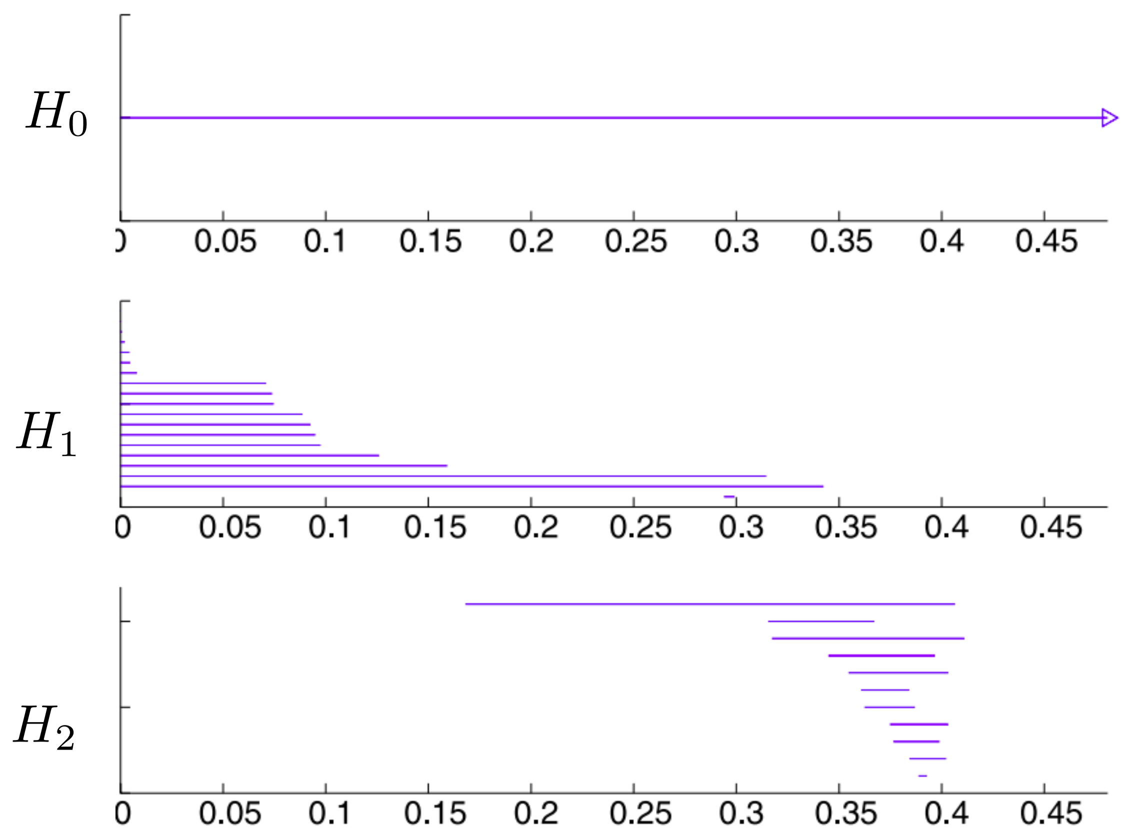
G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, On the Local Behavior of Spaces of Natural Images, 2008

https://link.springer.com/article/10.1007/s11263-007-0056-x

From a large collection of natural images, the authors extract 3×3 patches. Since it consists of 9 pixels, each of these patches can be seen as a 9-dimensional vector, and the whole set as a point cloud in $\mathbb{R}^9$.



We get the barcodes:

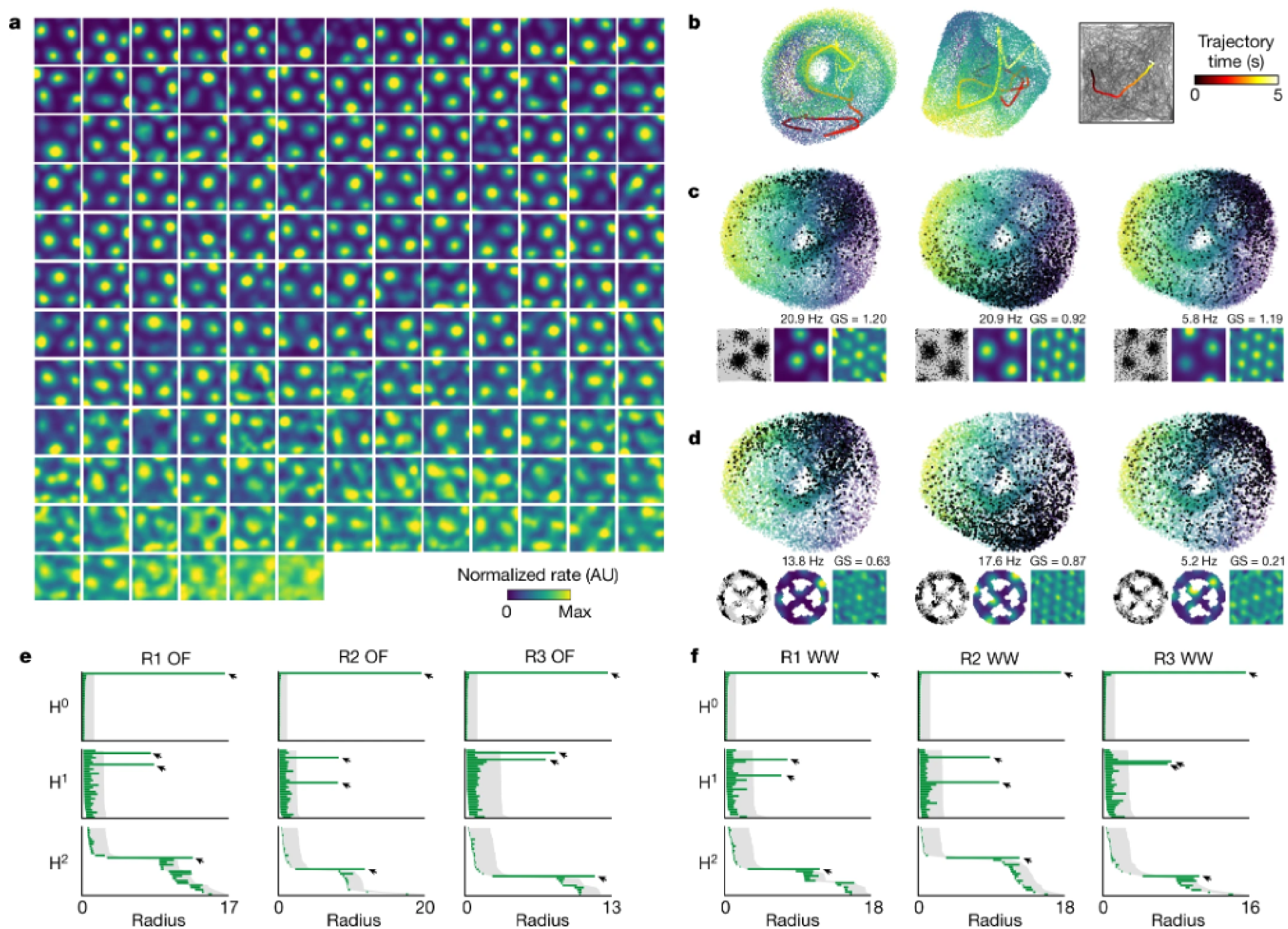


We deduce:

$$H_0 = \mathbb{Z}/2\mathbb{Z},$$
$$H_1 = (\mathbb{Z}/2\mathbb{Z})^2,$$
$$H_2 = \mathbb{Z}/2\mathbb{Z}.$$

[Richard J. Gardner et al, Toroidal topology of population activity in grid cells , 2022]

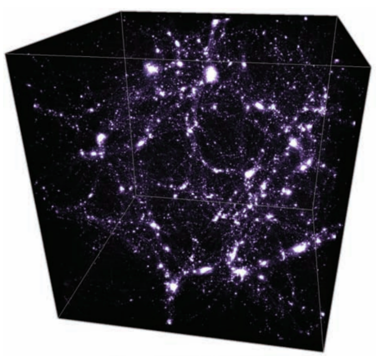
Os autores registraram spikes de grid-cells de ratos, e aplicaram redução de dimensionalidade na matriz de firing. Ao aplicar a homologia persistente a esta nuvem de pontos, observamos a homologia de um toro.



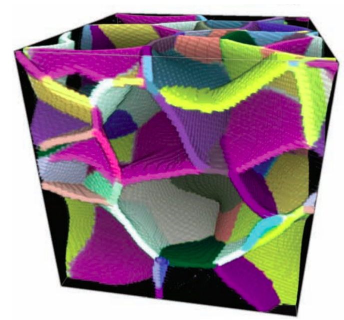
Multiscale analysis I

T. Sousbie, The persistent cosmic web and its filamentary structure, 2011

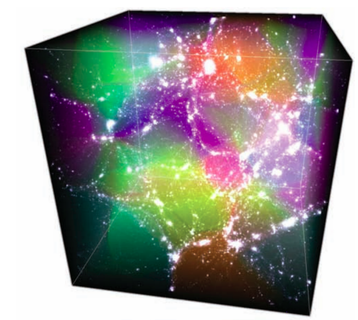
https://www.giss.nasa.gov/staff/mway/cluster/sousbie2011mnras.pdf



seen as an object of dimension 1



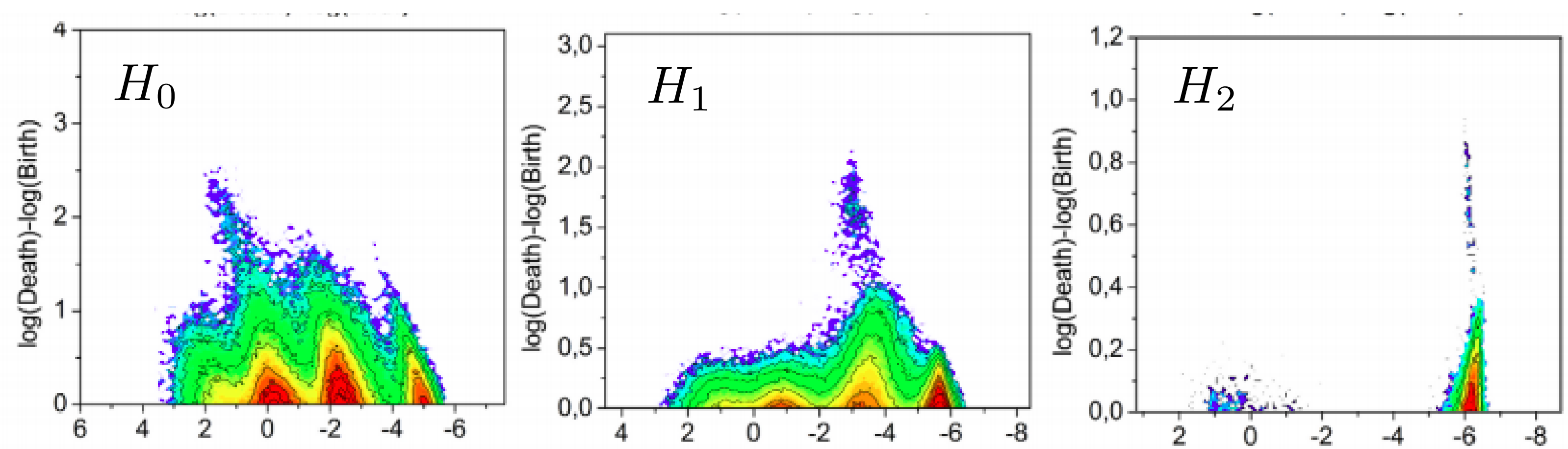
of dimension 2



of dimension 3

P. Pranav, H. Edelsbrunner, R. de Weygaert, G. Vegter, M. Kerber, B. Jones and M. Wintraecken, The topology of the cosmic web in terms of persistent Betti numbers, 2016

https://arxiv.org/pdf/1608.04519.pdff

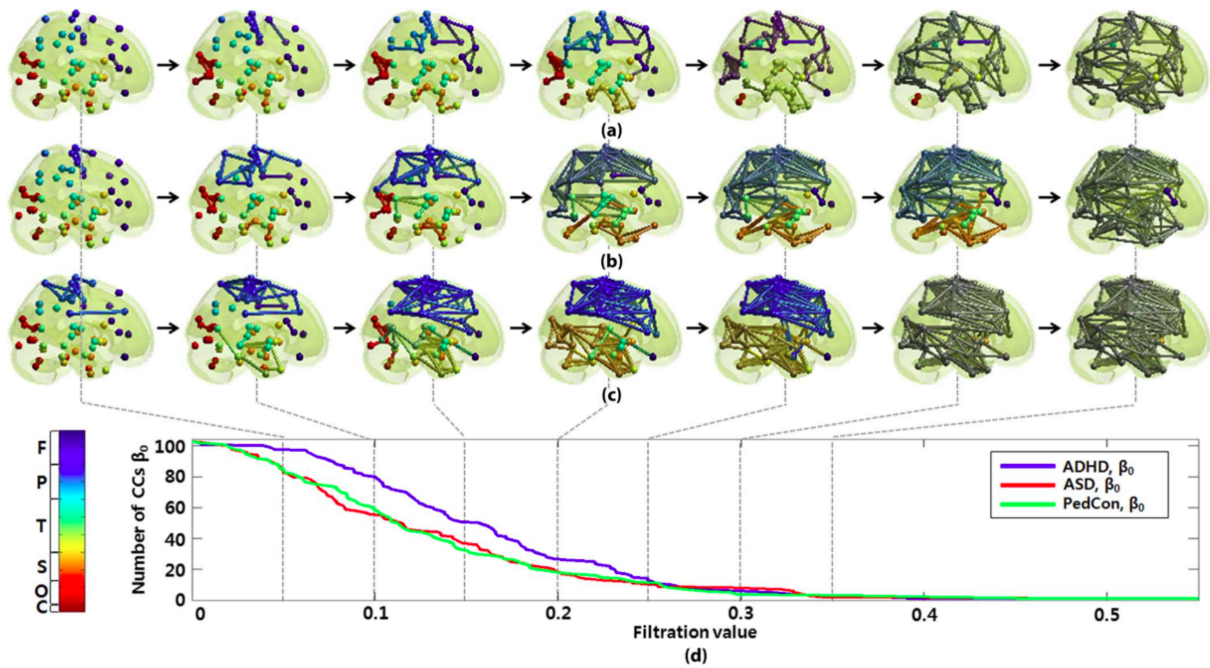


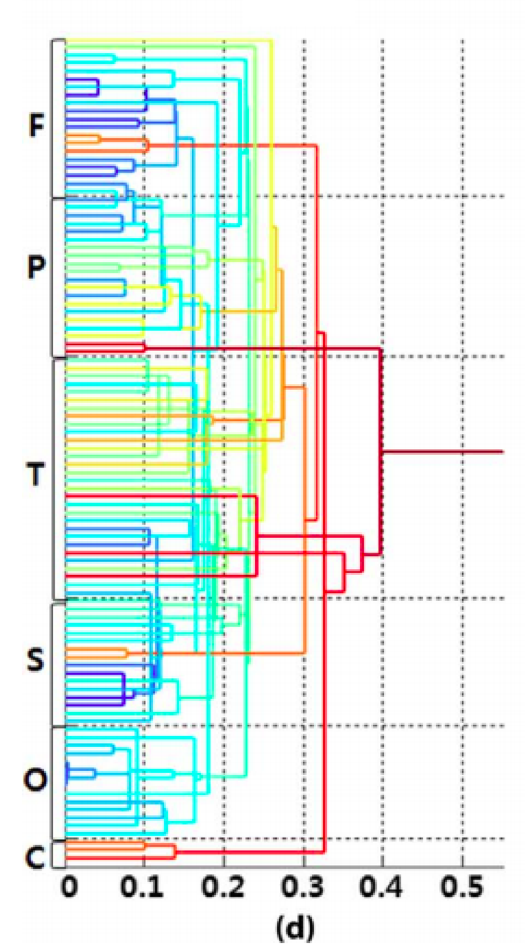
Average persistence diagrams (log-scale) for a Voronoï evolution model

Hyekyoung Lee, Hyejin Kang, Moo K Chung, Bung-Nyun Kim, Dong Soo Lee, Persistent brain network homology from the perspective of dendrogram, 2012

http://pages.stat.wisc.edu/~mchung/papers/lee.2012.TMI.pdf

⟶ H_0-persistent homology induces a hierarchical clustering





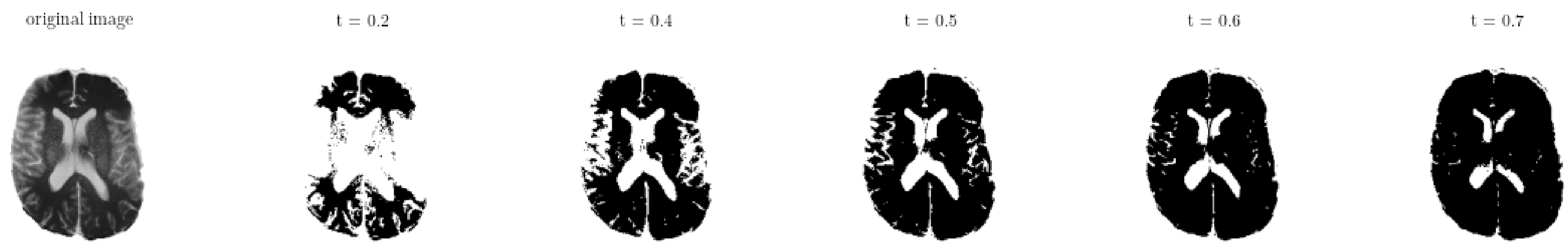
Análise de imagens médicas

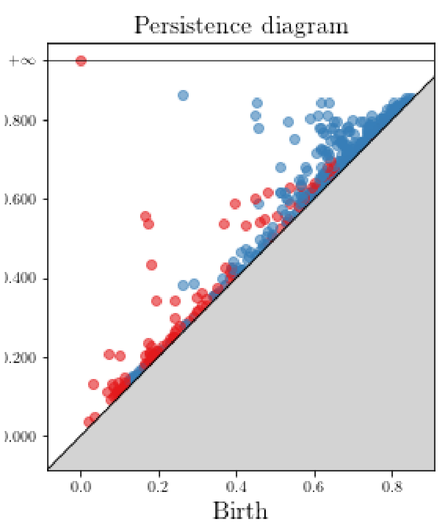
Em colaboração com Anton François.

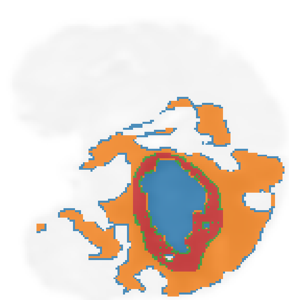
O **glioblastoma** é o tumor cerebral mais comum, difuso, de grau variável de agressividade, e cujo prognóstico médico é difícil de estabelecer.

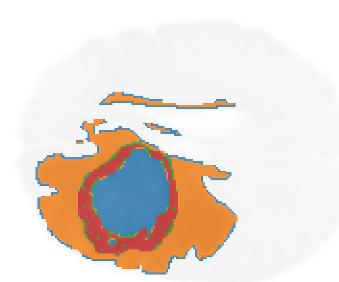
Neste contexto, o problema da **segmentação** consiste na demarcação automática das três regiões que formam o tumor (edema, núcleo necrótico e enhancing tumor).

Podemos usar a **homologia persistente cúbica**, especialmente definida para imagens.









Kernels for machine learning

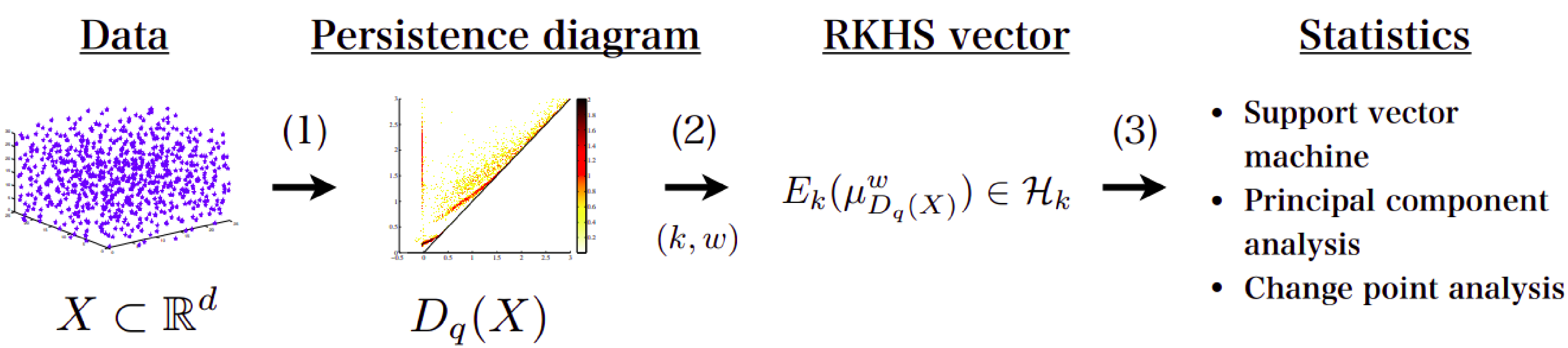
Mathieu Carrière, Marco Cuturi, Steve Oudot, Sliced Wasserstein Kernel for Persistence Diagrams, 2017

https://arxiv.org/abs/1706.03358

Genki Kusano, Kenji Fukumizu, Yasuaki Hiraoka, Kernel Method for Persistence Diagrams via Kernel Embedding and Weight Factor, 2018

https://www.jmlr.org/papers/volume18/17-317/17-317.pdf

⟶ Barcodes are not subsets of some Euclidean space, hence usual machine learning methods cannot be used directly



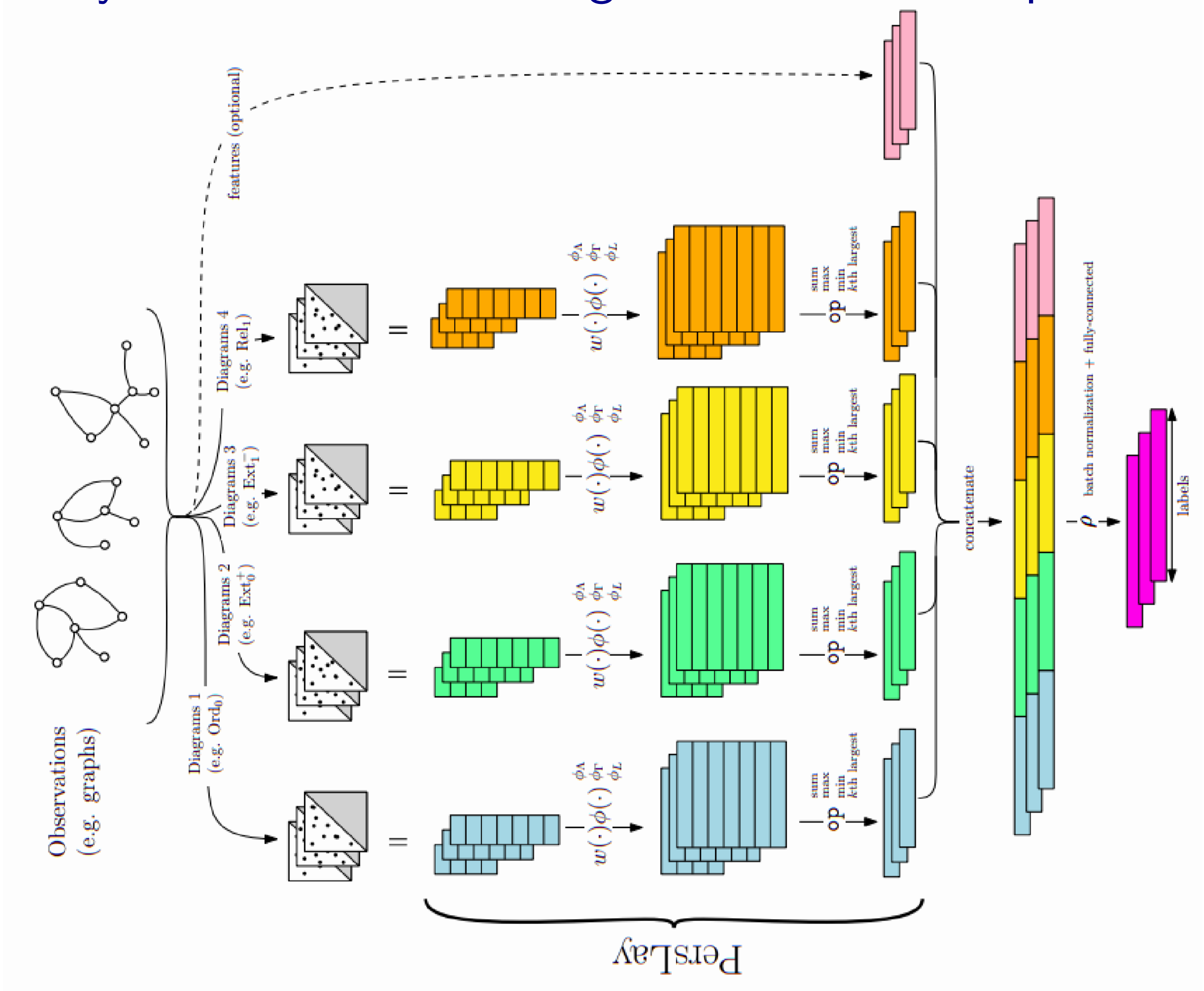
Topological layer in Neural Networks

Rickard Brüel-Gabrielsson, Bradley J. Nelson, Anjan Dwaraknath, Primoz Skraba, Leonidas J. Guibas, Gunnar Carlsson, A Topology Layer for Machine Learning, 2019

https://arxiv.org/abs/1905.12200

Mathieu Carrière, Frédéric Chazal, Yuichi Ike, Théo Lacombe, Martin Royer, Yuhei Umeda, PersLay: A Neural Network Layer for Persistence Diagrams and New Graph Topological Signatures, 2019

https://arxiv.org/abs/1904.09378

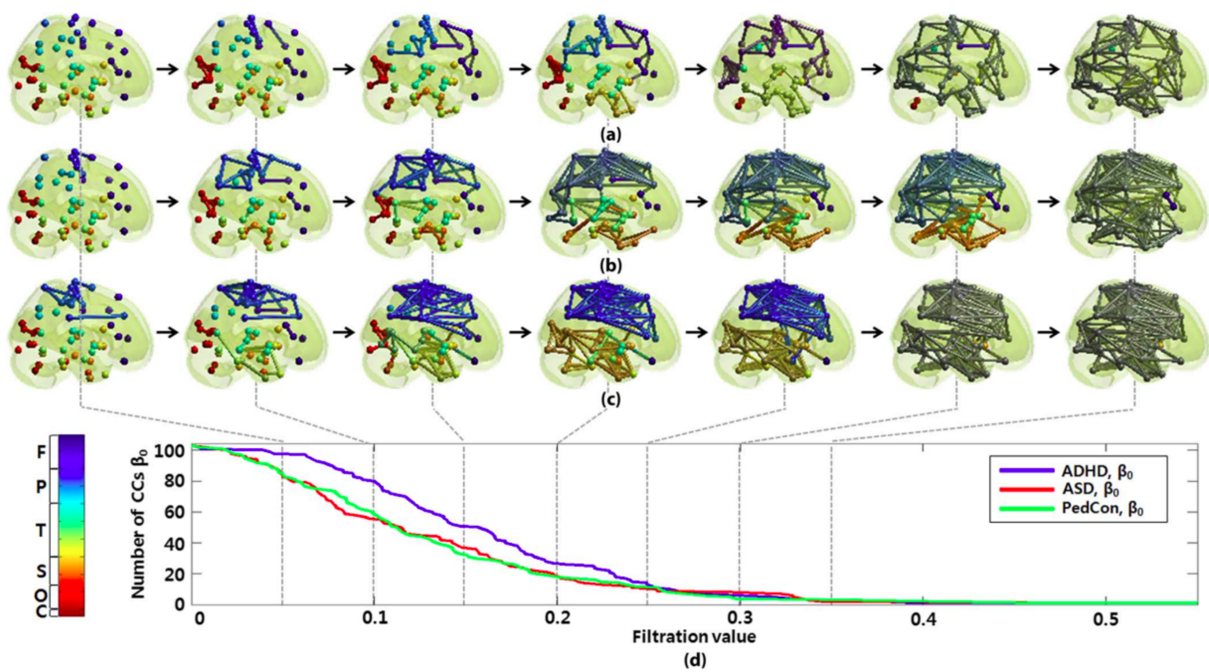


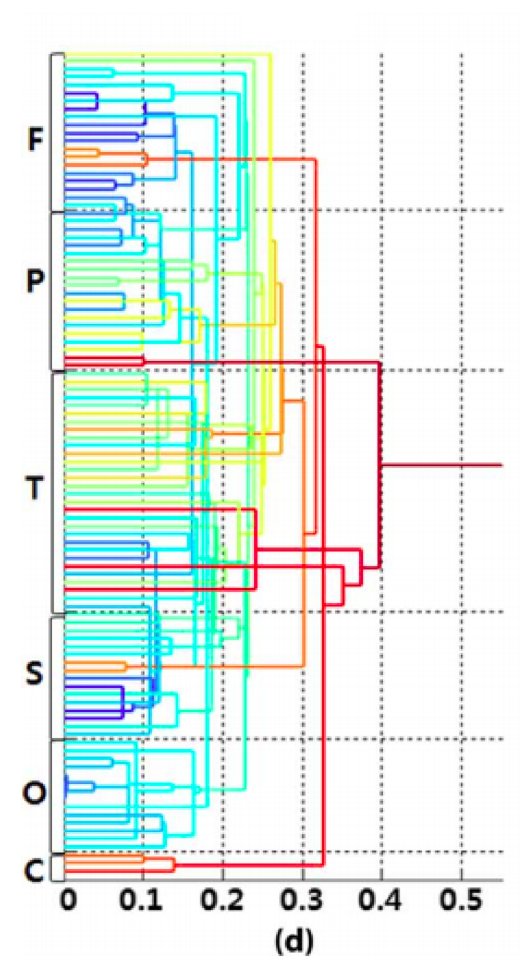
Hierarchical clustering

Hyekyoung Lee, Hyejin Kang, Moo K Chung, Bung-Nyun Kim, Dong Soo Lee, Persistent brain network homology from the perspective of dendrogram, 2012

http://pages.stat.wisc.edu/~mchung/papers/lee.2012.TMI.pdf

⟶ H_0-persistent homology induces a hierarchical clustering





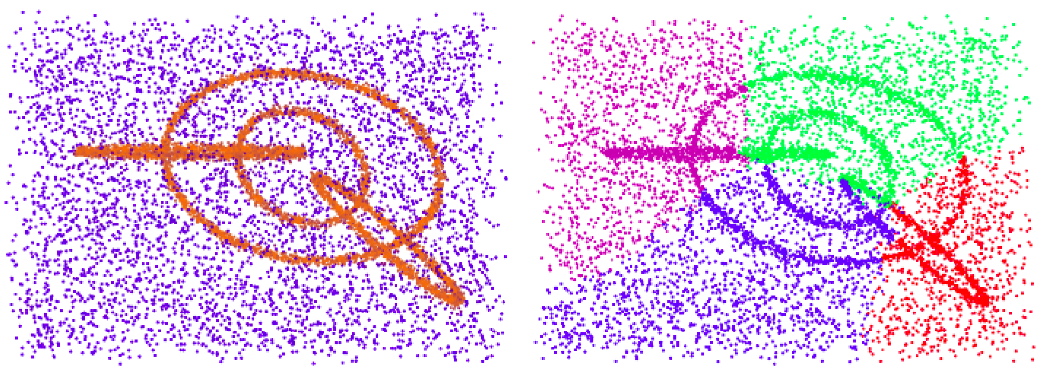
Classification

Frédéric Chazal, Steve Oudot, Primoz Skraba, Leonidas J. Guibas, Persistence-Based Clustering in Riemannian Manifolds, 2011

https://geometrica.saclay.inria.fr/team/Fred.Chazal/papers/cgos-pbc-09/cgos-pbcrm-11.pdf

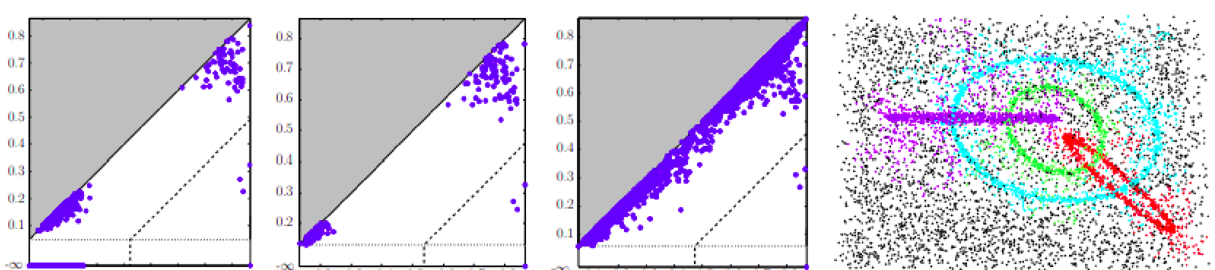
Chunyuan Li, Maks Ovsjanikov, Frederic Chazal, Persistence-based Structural Recognition, 2014

https://geometrica.saclay.inria.fr/team/Fred.Chazal/papers/loc-pbsr-14/CVPR2014.pdf

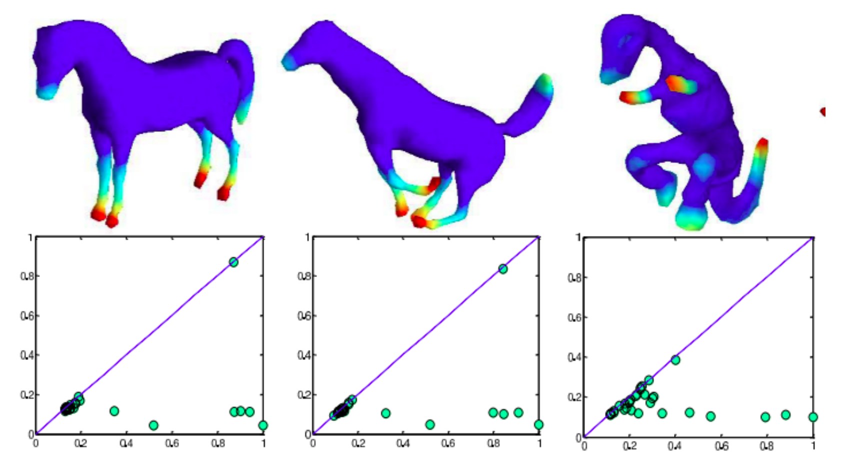


(a) (b)

Figure 7: (a) The rings data set with the estimated density function. (b) The result obtained using spectral clustering.



(a) (b) (c) (d)



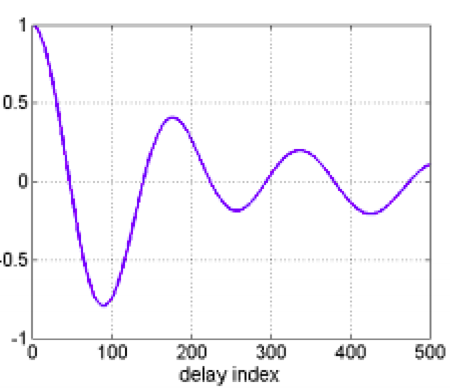
Saba Emrani, Thanos Gentimis, Hamid Krim Persistent Homology of Delay Embeddings and its Application to Wheeze Detection, 2014

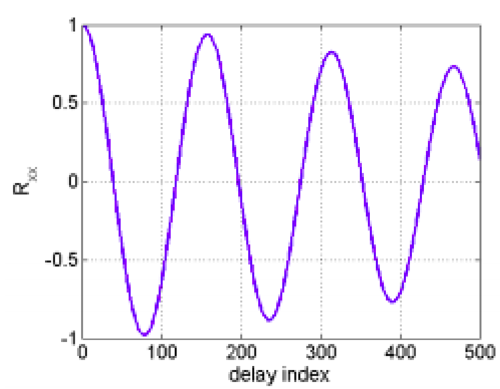
https://www.researchgate.net/publication/260523931_Persistent_Homology_of_Delay_Embeddings_and_its_Application_to_Wheeze_Detection

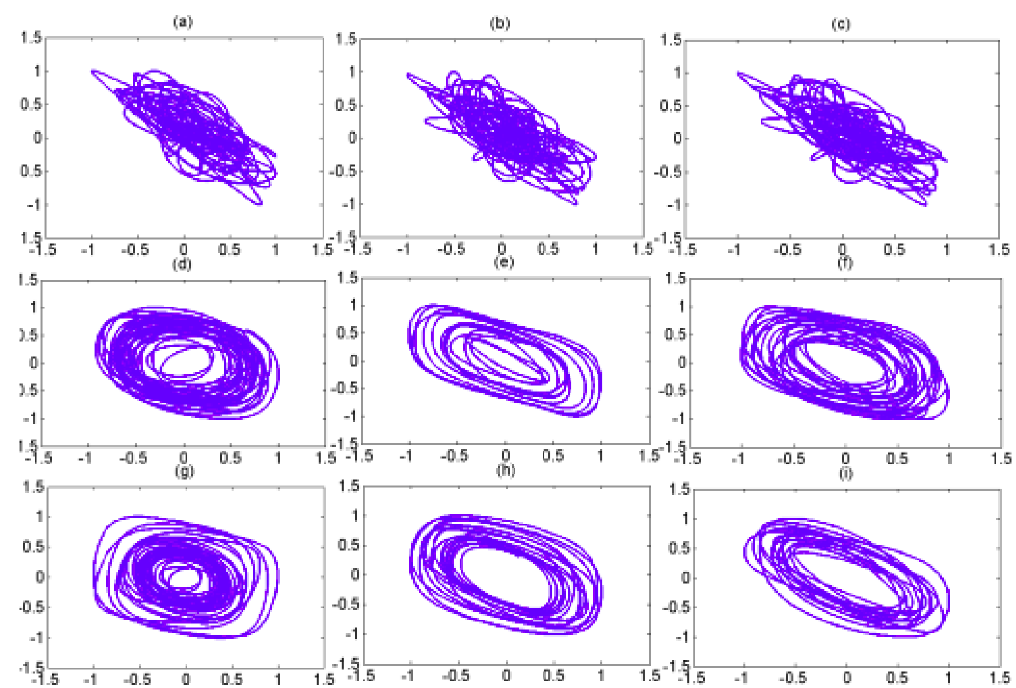
⟶ a time series $(x_1, x_2, x_3, \dots)$ does not contain topology...

turn it into a point cloud of $\mathbb{R}^n$ via **time delay embedding**!

$X = \{\overline{x}_1, \overline{x}_2, \overline{x}_3, \dots\} \subset \mathbb{R}^n$ where $\quad \overline{x}_k = (x_k, x_{k+1}, \dots, x_{k+n-1})$



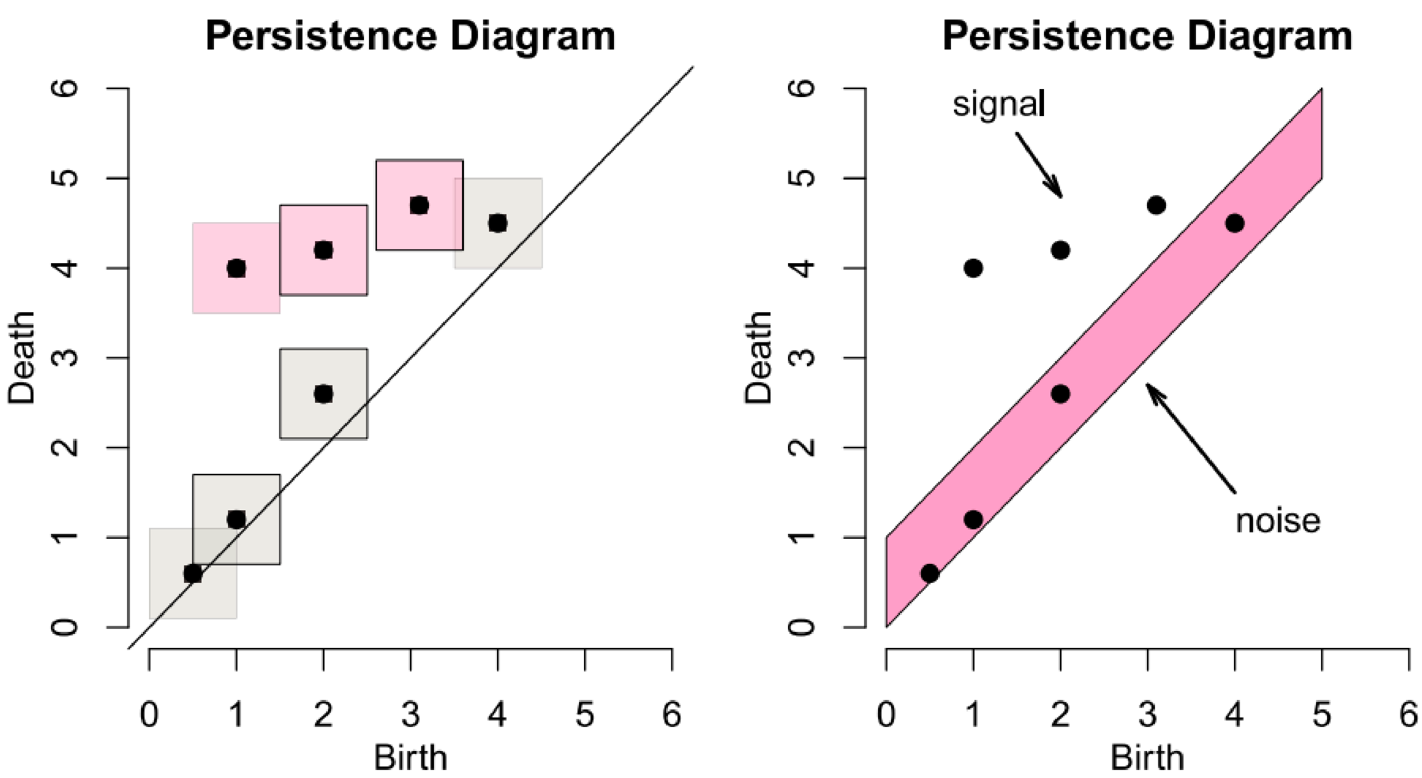




Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, Larry Wasserman, Sivaraman Balakrishnan and Aarti Singh, Confidence sets for persistence diagrams, 2014

https://arxiv.org/pdf/1303.7117.pdf

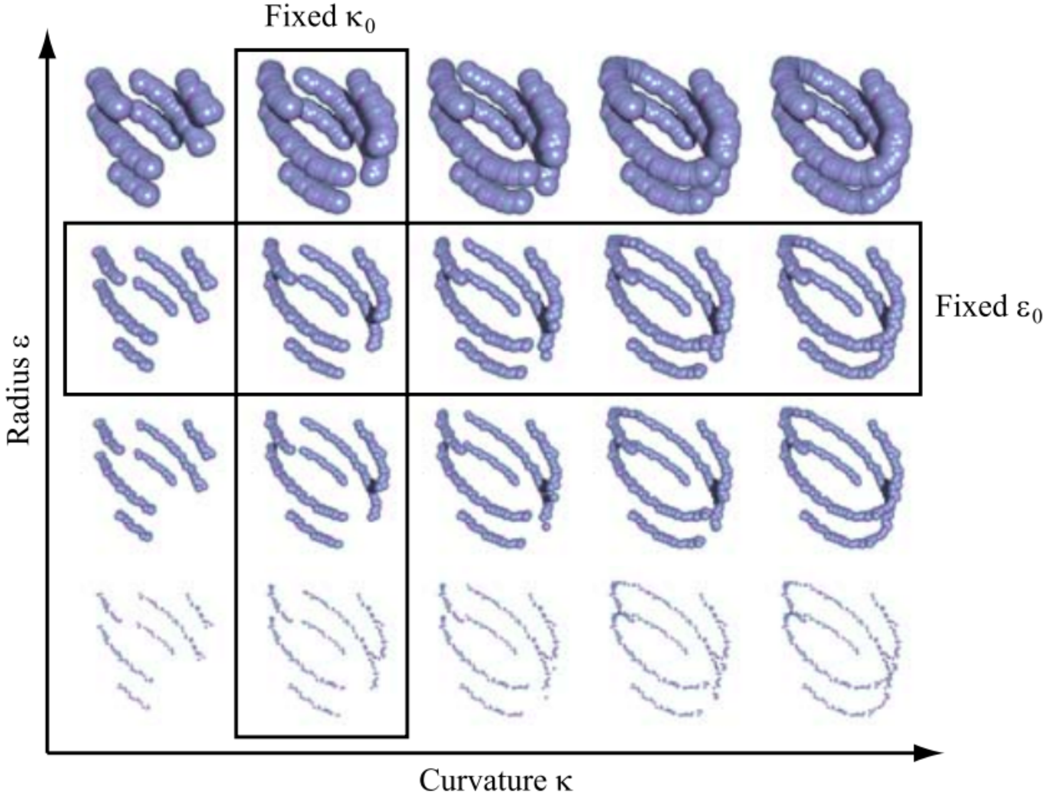
⟶ Given a barcode, how to determine statistically what is noise and what is not?

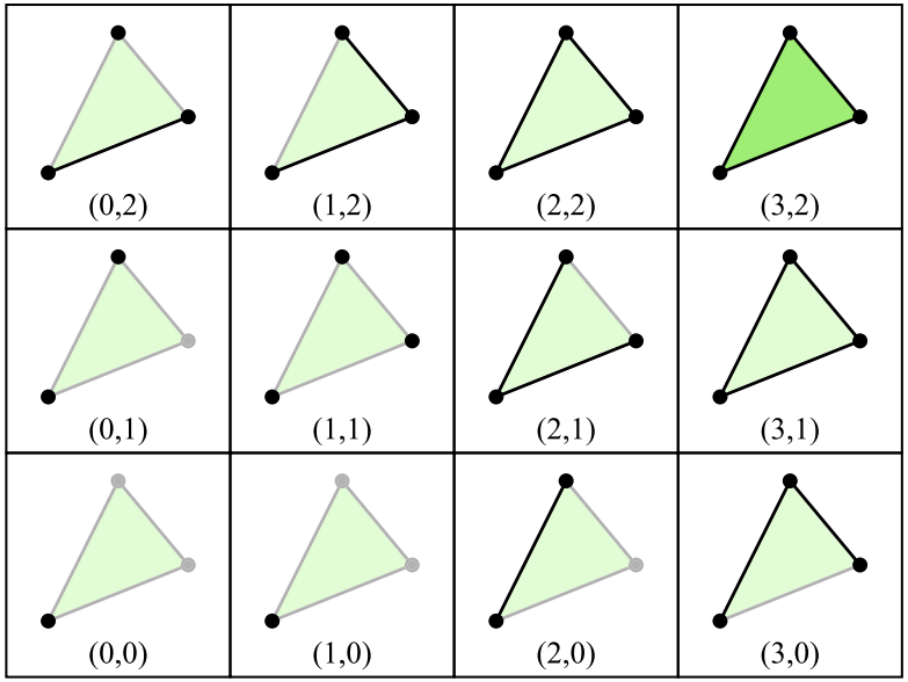


Gunnar Carlsson, Afra Zomorodian, The Theory of Multidimensional Persistence, 2009

https://link.springer.com/article/10.1007/s00454-009-9176-0

→ What if our filtration is not indexed only by $t \in \mathbb{R}^+$?



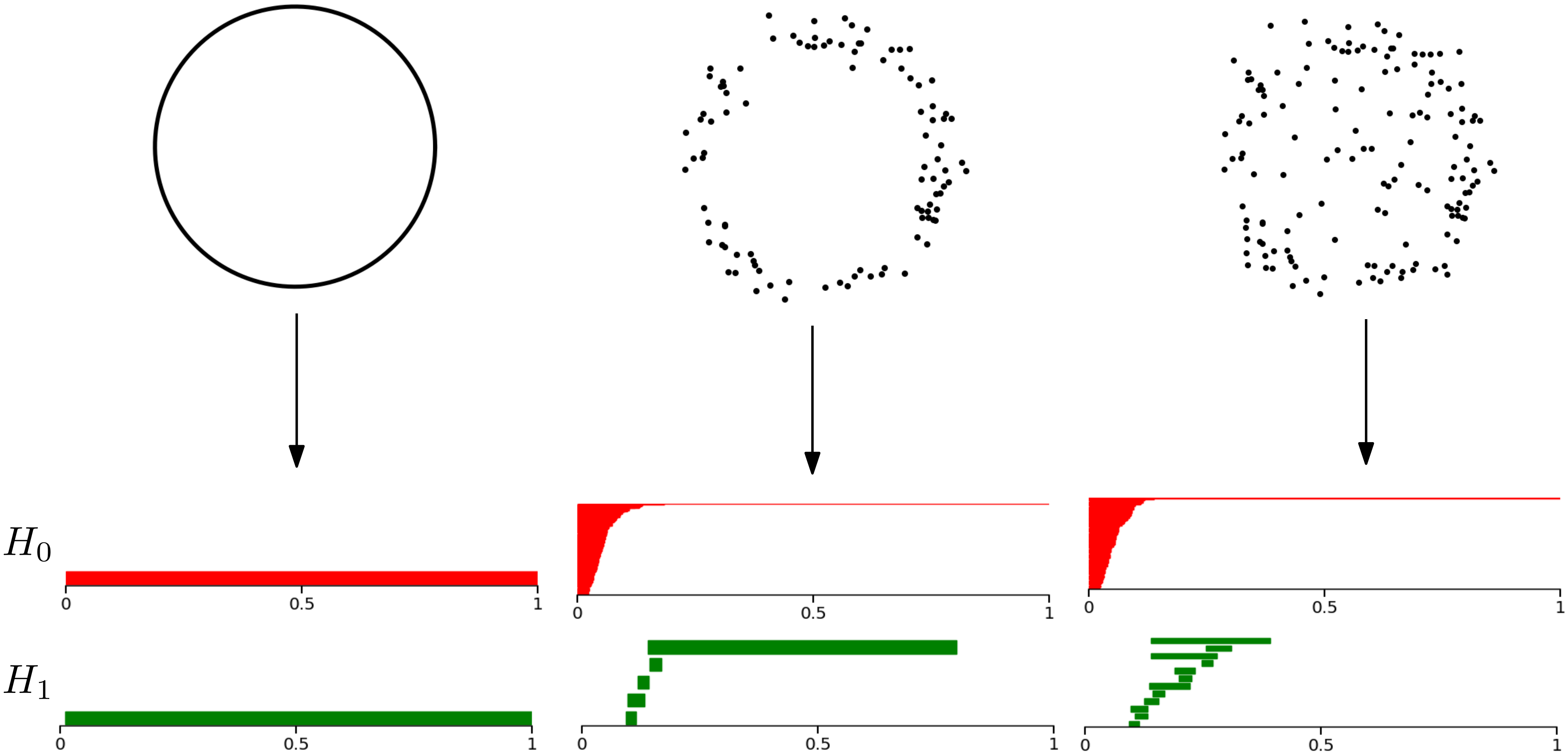


Wasserstein stability

Hirokazu Anai, Frédéric Chazal, Marc Glisse, Yuichi Ike, Hiroya Inakoshi, Raphaël T., Yuhei Umeda, DTM-based filtrations, 2020

https://arxiv.org/abs/1811.04757

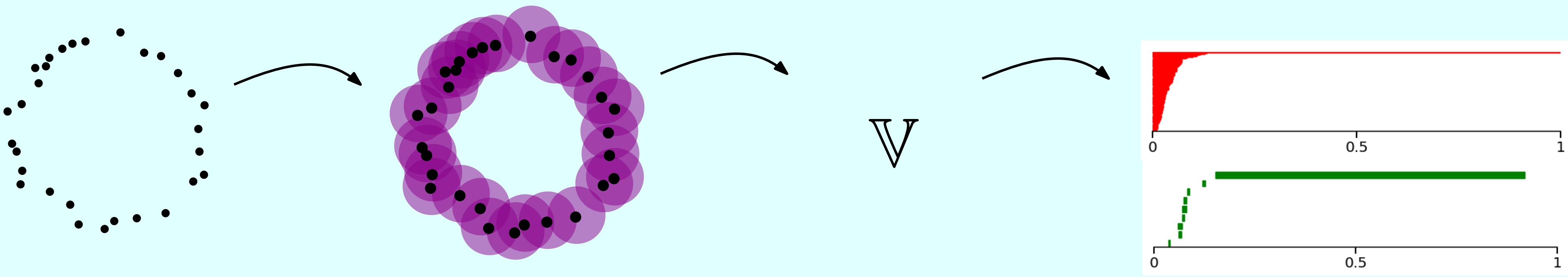
→ When our dataset is not close to an uderlying object in **Hausdorff distance**



Wasserstein stability

Conclusão

A homologia persistente permite uma estimativa **multi-escala** e **estável** da homologia dos conjuntos de dados.



nuvem de pontos | filtração de Čech | módulo de persistência | código de barras

Permite analisar os dados a partir de uma nova perspectiva.

Um curso de TDA: `https://raphaeltinarrage.github.io/EMAp.html`

REPÚBLICA FEDERATIVA DO BRASIL

MINISTÉRIO DA EDUCAÇÃO - MEC

UNIVERSIDADE FEDERAL DO TDA

O Reitor da UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. no uso de suas atribuições e tendo em vista a conclusão do Curso de *Engenharia da Computação*, em 21/05/2002 confere o titulo de *Bacharel em Engenharia da Computação* a

Tu Nombre

cédula de identidade nº. 182910 (órgão expedidor) I.F.P. - RJ

nascido (a) a 13 de Dezembro de 1972 natural Rio de Janiero

e outorga - lhe o presente Diploma, a fim de que possa gozar de todos os direitos e prerrogativas legais.

Rio de Janeiro, 21 de Maio de 2002

Reitor

Secretario

Diretor