Spectral sequences for multidimensional persistence

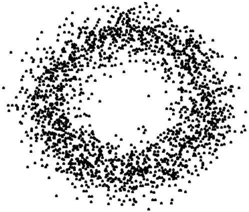
Andrea Guidolin

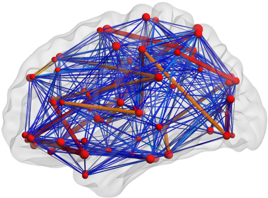
Politecnico di Torino - ISI Foundation

23 January 2017

Topological data analysis

Applications: image analysis, shape study, sensor networks, viral evolution and propagation of epidemics, internet and social networks, biological networks . . .





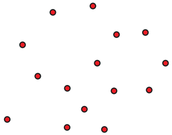
Topological methods:

- Transform data into global topological object (**simplicial complexes**)
- Study these objects through algebraic topology (**persistent homology**)

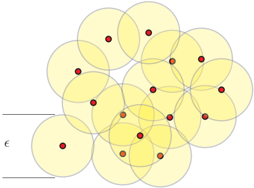
Plan of the talk

1. Persistent homology
   - data $\rightarrow$ simplicial complexes $\rightarrow$ filtrations $\rightarrow$ topological invariants
2. Spectral sequences vs Persistent homology
   - "same topological information"
3. Multidimensional case

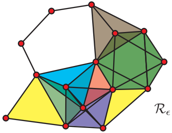
From data to simplicial complexes



(a) Point cloud



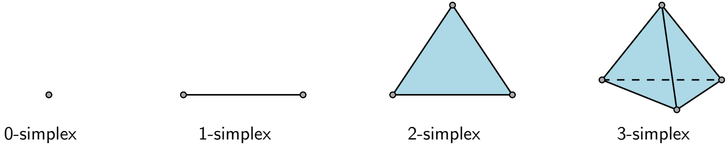
(b) Connections between points



(c) Simplicial complex (Rips)

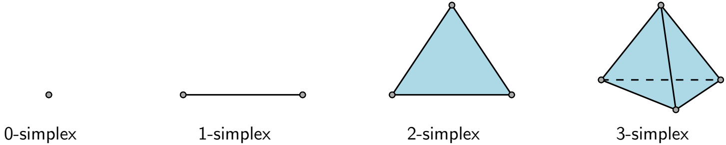
Simplicial complexes

Simplices:

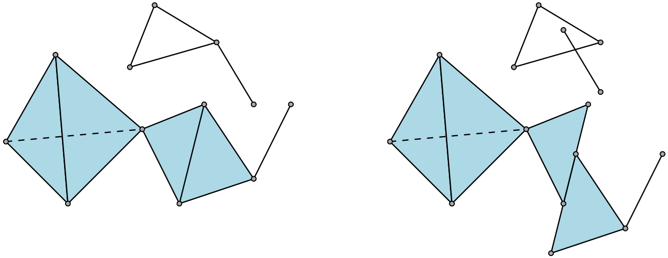


Simplicial complexes

Simplices:

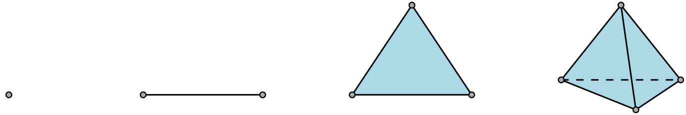


Simplicial complexes:



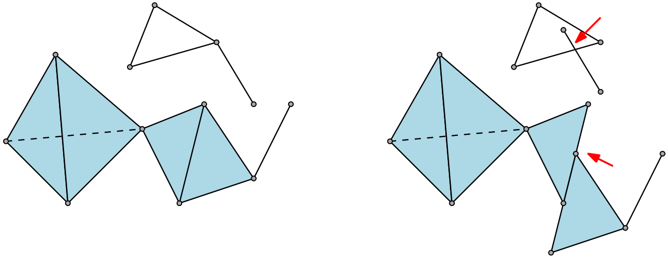
Simplicial complexes

Simplices:



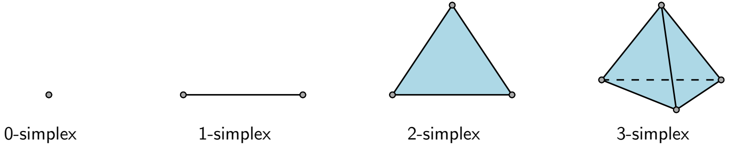
0-simplex 1-simplex 2-simplex 3-simplex

Simplicial complexes:

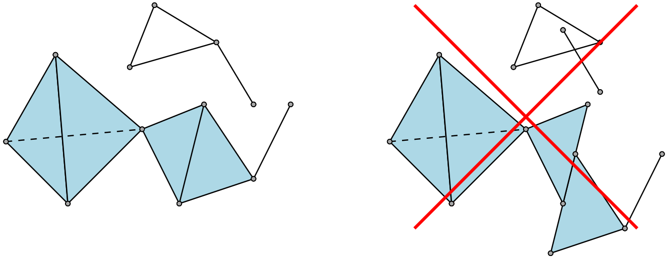


Simplicial complexes

Simplices:

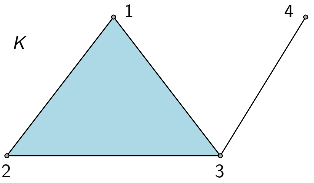


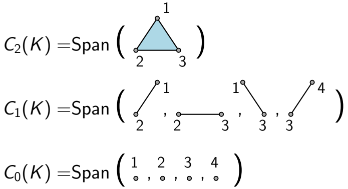
Simplicial complexes:



Homology: Chain complex of a simplicial complex

$C_n(K) := \{$linear combinations of n-simplices of $K\}$





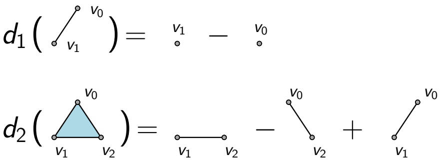
Chain complex of K:

$$C_\bullet(K): \quad \cdots \xrightarrow{d_3} C_2(K) \xrightarrow{d_2} C_1(K) \xrightarrow{d_1} C_0(K) \to 0$$

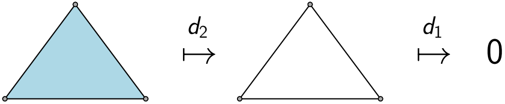
The **boundary operator**

$$d_n : C_n(K) \to C_{n-1}(K)$$

is the linear map defined on n-simplices as the alternating sum of $(n-1)$-faces:



Fundamental property: $d_n d_{n+1} = 0$



Homology groups and Betti numbers

Consider the chain complex of a simplicial complex K

$$C_\bullet(K): \quad \cdots \to C_{n+1}(K) \xrightarrow{d_{n+1}} C_n(K) \xrightarrow{d_n} C_{n-1}(K) \to \cdots$$

The fundamental property $d_n d_{n+1} = 0$ implies

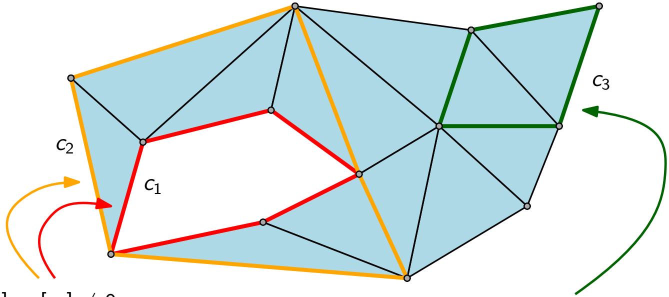
$$\operatorname{Im} d_{n+1} \subseteq \ker d_n \;\; (\subseteq C_n(K)).$$

The n-th **homology group** of K is defined as

$$H_n(K) := \frac{\ker d_n}{\operatorname{Im} d_{n+1}}$$

and its rank β_n is called n-th **Betti number**.

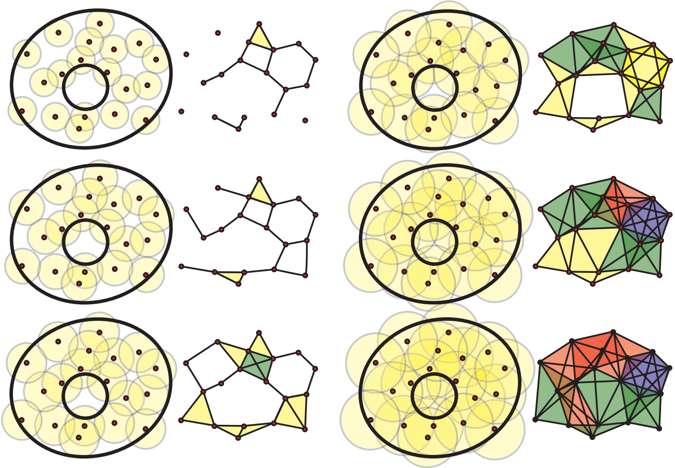
Homology classes



$[c_1] = [c_2] \neq 0$
in $H_1(K)$

$[c_3] = 0$ in $H_1(K)$

Filtration of simplicial complexes



Persistent homology

Given a filtration of simplicial complexes

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_m = K,$$

for every $i \leq j$ the inclusion $K_i \subseteq K_j$ induces maps in homology

$$f_n^{i,j} : H_n(K_i) \to H_n(K_j).$$

For every $i \leq j$, we call the groups

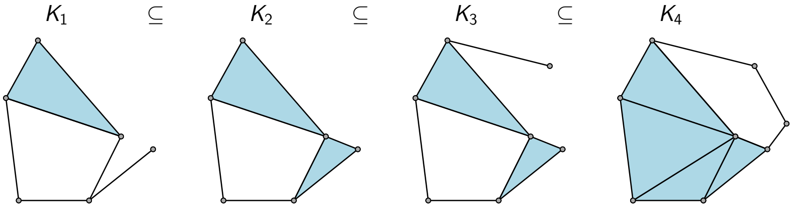
$$H_n^{i,j} := \operatorname{Im}(f_n^{i,j})$$

persistent homology groups. We define

$$\beta_n^{i,j} := \dim_k H_n^{i,j}$$

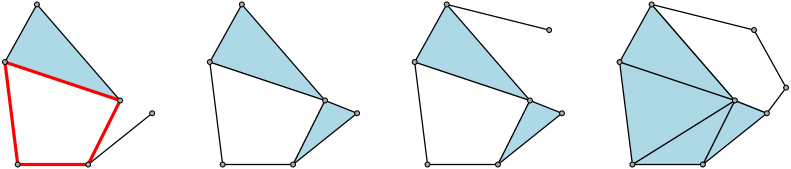
persistent Betti numbers.

Persistent homology



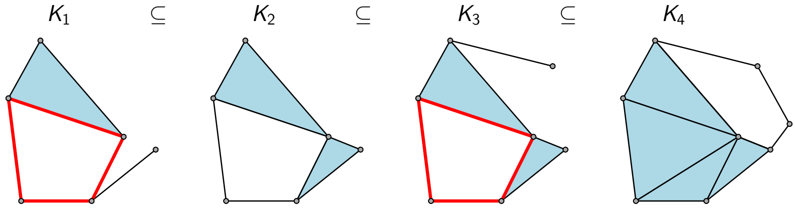
Persistent homology

$K_1 \quad \subseteq \quad K_2 \quad \subseteq \quad K_3 \quad \subseteq \quad K_4$



$[z] \in H_1(K_1)$

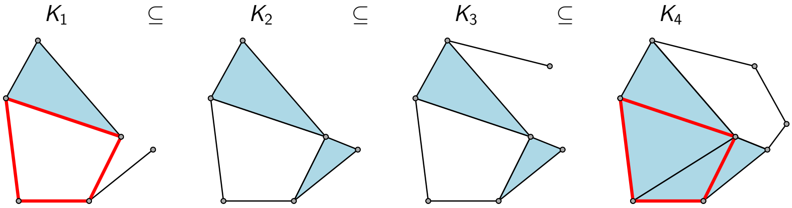
Persistent homology



$[z] \in H_1(K_1)$ $\qquad$ $[z] \in H_1(K_3)$

Persistent Betti numbers: $\beta_1^{1,3} = 1$

Persistent homology

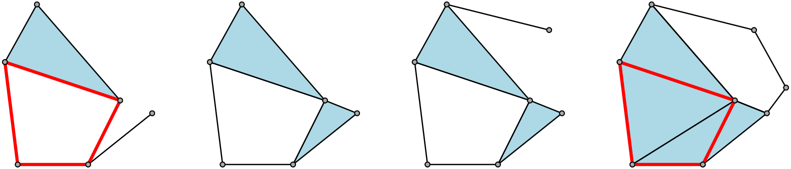


$[z] \in H_1(K_1)$ $\qquad$ $[z] = 0 \in H_1(K_4)$

Persistent Betti numbers: $\beta_1^{1,3} = 1$, $\beta_1^{1,4} = 0$

Persistent homology

$K_1 \subseteq K_2 \subseteq K_3 \subseteq K_4$

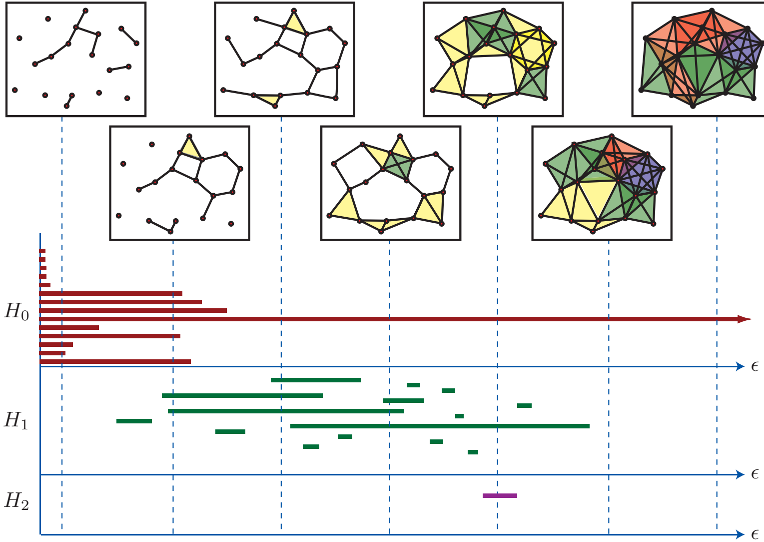


$[z] \in H_1(K_1)$ $\qquad$ $[z] = 0 \in H_1(K_4)$

Persistent Betti numbers: $\beta_1^{1,3} = 1$, $\beta_1^{1,4} = 0$, ...

Barcode for H_1 (counting "holes"):

Barcodes



Plan of the talk

1. Persistent homology
   - data → simplicial complexes → filtrations → topological invariants
2. Spectral sequences vs Persistent homology
   - "same topological information"
3. Multidimensional case

Spectral sequences

Filtration of a simplicial complex:

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq \; K_{p-1} \subseteq \; K_p \subseteq \cdots \subseteq \; K_m = K.$$

Filtration of a chain complex:

$$0 = F_0 \hookrightarrow F_1 \hookrightarrow \ldots \hookrightarrow F_{p-1} \hookrightarrow F_p \hookrightarrow \ldots \hookrightarrow F_m = C_\bullet(K),$$

where we denote $F_p := C_\bullet(K_p)$.

Leray ($\sim$ 1940): spectral sequences as a device for computing the homology $H_n(K)$ using the information contained in the filtration.

Spectral sequences

Filtration of a chain complex:

$$0 = F_0 \hookrightarrow F_1 \hookrightarrow \ldots \hookrightarrow F_{p-1} \hookrightarrow F_p \hookrightarrow \ldots \hookrightarrow F_m = C_\bullet(K)$$

Spectral sequence:

$$\{E_p^r\}_{p,r\in\mathbb{Z}}, \quad r \text{ ``page index''}, p \text{ ``filtration index''},$$

where

$$E_p^1 := H(F_p/F_{p-1})$$

and there exist differentials

$$\cdots E_{p+r}^r \xrightarrow{d} E_p^r \xrightarrow{d'} E_{p-r}^r \cdots \quad \text{with} \quad E_p^{r+1} \cong \ker d' / \operatorname{Im} d.$$

Convergence

There exists a filtration $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_m = H(K)$ such that, for $r \geq r_0$,

$$E_p^r \cong V_p/V_{p-1}.$$

Spectral sequences

Filtration of a chain complex:

$$0 = F_0 \hookrightarrow F_1 \hookrightarrow \ldots \hookrightarrow F_{p-1} \hookrightarrow F_p \hookrightarrow \ldots \hookrightarrow F_m = C_\bullet(K)$$

Spectral sequence:

$$\{E^r_p\}_{p,r\in\mathbb{Z}}, \quad r \text{ ``page index''}, p \text{ ``filtration index''},$$

where

$$E^1_p := H(F_p/F_{p-1})$$

and there exist differentials

$$\cdots E^r_{p+r} \xrightarrow{d} E^r_p \xrightarrow{d'} E^r_{p-r} \cdots \qquad \text{with} \quad E^{r+1}_p \cong \ker d' / \operatorname{Im} d.$$

Convergence

There exists a filtration $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_m = H(K)$ such that, for $r \geq r_0$,

$$E^r_p \cong V_p/V_{p-1}.$$

Spectral sequences vs Persistent homology

Given a filtration

$$0 = F_0 \hookrightarrow F_1 \hookrightarrow \ldots \hookrightarrow F_{p-1} \hookrightarrow F_p \hookrightarrow \ldots \hookrightarrow F_m = C_\bullet(K)$$

study the relationship between the spectral sequence $\{E^r_p\}$ and the persistent homology groups $\{H^{i,j} = \mathrm{Im}(H(F_i) \to H(F_j))\}$:
there are long exact sequences

$$\ldots H^{p-1,p+r-2} \to H^{p,p+r-1} \to E^r_p \to H^{p-r,p-1} \to H^{p-r+1,p} \ldots$$

The spectral sequence and the persistent homology groups "contain the same topological information" :

$$\dim_k E^r_p \quad \leftrightsquigarrow \quad \beta^{i,j} := \dim_k H^{i,j}$$

Plan of the talk

1. Persistent homology
   - data $\rightarrow$ simplicial complexes $\rightarrow$ filtrations $\rightarrow$ topological invariants
2. Spectral sequences vs Persistent homology
   - "same topological information"
3. Multidimensional case

Multidimensional persistent homology

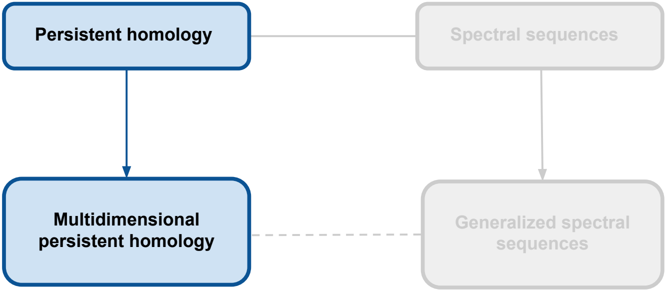
Multidimensional filtration of a simplicial complex:

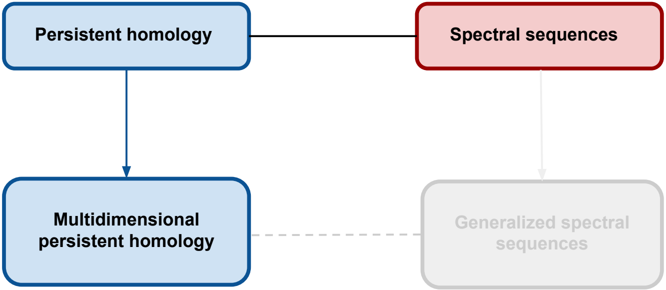
$$
\begin{array}{ccccccc}
K_{m1} & \hookrightarrow & K_{m2} & \hookrightarrow & \cdots & \hookrightarrow & K_{mm'} \\
\uparrow & & \uparrow & & & & \uparrow \\
\cdots & & \cdots & & & & \cdots \\
\uparrow & & \uparrow & & & & \uparrow \\
K_{21} & \hookrightarrow & K_{22} & \hookrightarrow & \cdots & \hookrightarrow & K_{2m'} \\
\uparrow & & \uparrow & & & & \uparrow \\
K_{11} & \hookrightarrow & K_{12} & \hookrightarrow & \cdots & \hookrightarrow & K_{1m'}
\end{array}
$$

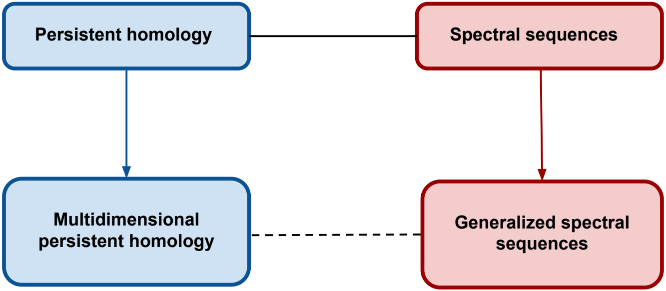
Associated invariant: **rank invariant**

$$\beta_n^{P,Q} := \dim_k \operatorname{Im}(H_n(K_P) \to H_n(K_Q)), \quad P, Q \in \mathbb{Z}^2, P \preceq Q.$$

Unlike the 1-dimensional case (persistent Betti numbers), it is *not* a complete invariant.


Persistent homology
Spectral sequences
Multidimensional persistent homology
Generalized spectral sequences


Persistent homology
Spectral sequences
Multidimensional persistent homology
Generalized spectral sequences


Persistent homology
Spectral sequences
Multidimensional persistent homology
Generalized spectral sequences

Generalized spectral sequences

For a filtration $\{F_p\}_{p\in\mathbb{Z}}$ we have

$$E_p^1 = H(F_p/F_{p-1}) = \frac{\text{rel. cycles}}{\text{rel. boundaries}} = \frac{F_p \cap d^{-1}(F_{p-1})}{d(F_p) + F_{p-1}}$$

$$E_p^r = \frac{F_p \cap d^{-1}(F_{p-r})}{d(F_{p+r-1}) + F_{p-1}}$$

This can be generalized for a **filtration indexed over a poset** I, i.e. a collection of chain complexes $\{F_i\}_{i\in I}$ with $F_i \hookrightarrow F_j \iff i \leq j$, in the following way:

$$S[z,q,p,b] := \frac{F_p \cap d^{-1}(F_z)}{d(F_b) + F_q},$$

for all $z \leq q \leq p \leq b$ in I.

Generalized spectral sequences

For a filtration $\{F_p\}_{p\in\mathbb{Z}}$ we have

$$E_p^1 = H(F_p/F_{p-1}) = \frac{\text{rel. cycles}}{\text{rel. boundaries}} = \frac{F_p \cap d^{-1}(F_{p-1})}{d(F_p) + F_{p-1}}$$

$$E_p^r = \frac{F_p \cap d^{-1}(F_{p-r})}{d(F_{p+r-1}) + F_{p-1}}$$

This can be generalized for a **filtration indexed over a poset** I, i.e. a collection of chain complexes $\{F_i\}_{i\in I}$ with $F_i \hookrightarrow F_j \iff i \leq j$, in the following way:

$$S[z,q,p,b] := \frac{F_p \cap d^{-1}(F_z)}{d(F_b) + F_q},$$

for all $z \leq q \leq p \leq b$ in I.

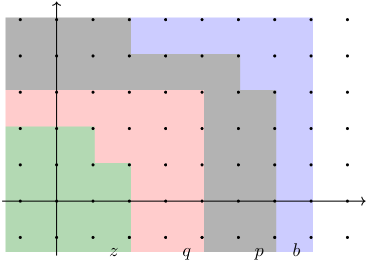
Spectral sequences for multidim. persistent homology

Consider the poset $I := D(\mathbb{Z}^n)$ of **downsets** of $\mathbb{Z}^n$, ordered wrt inclusion $\subseteq$. (A downset of $\mathbb{Z}^n$ is a subset $A \subseteq \mathbb{Z}^n$ such that if $P \in A$ and $Q \preceq P$, then $Q \in A$).

The terms of the generalized spectral sequence are of the form

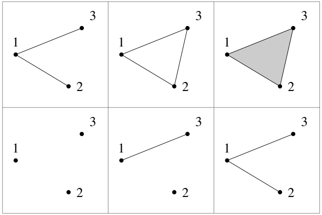
$$S[z, q, p, b] := \frac{F_p \cap d^{-1}(F_z)}{d(F_b) + F_q},$$

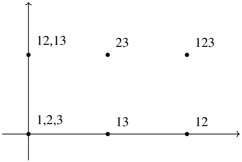
for all $z \subseteq q \subseteq p \subseteq b$ in $I = D(\mathbb{Z}^n)$.



Spectral sequences for multidim. persistent homology

How to define a filtration of chain complexes $\{F_p\}_{p \in I = D(\mathbb{Z}^n)}$ from a multidimensional filtration of simplicial complexes $\{K_P\}_{P \in \mathbb{Z}^n}$.





For every point $P \in \mathbb{Z}^n$ consider the simplices entering the filtration at K_P. Then, for $p \in I = D(\mathbb{Z}^n)$, set

$$F_p := \text{Span}\{\text{simplices entering the filtration at } K_P, \text{ for all } P \in p\}$$

Spectral sequences for multidim. persistent homology

Theorem

Under suitable hypotheses on $z_1, q_1, \dots, b_3 \in I = D(\mathbb{Z}^n)$, there exist differentials

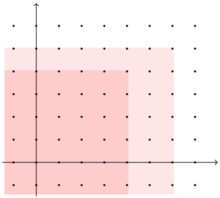
$$S[z_3, q_3, p_3, b_3] \xrightarrow{d} S[z_2, q_2, p_2, b_2] \xrightarrow{d'} S[z_1, q_1, p_1, b_1]$$

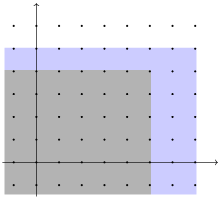
such that

$$\frac{\ker d'}{\operatorname{Im} d} \cong S[q_1, q_2, p_2, p_3].$$

Our idea: consider $\{\dim_k S[z, q, p, b]\}$ as invariant for the multidimensional filtration $\{K_P\}_{P \in \mathbb{Z}^n}$.

Spectral sequences for multidim. persistent homology

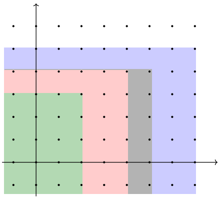




$S[-\infty, -\infty, q, q']$ $\xrightarrow{\text{incl}_*}$ $S[-\infty, -\infty, p, b]$

$\dots$

$S[z, q, p, b]$



Short bibliography

R. Ghrist, *Barcodes: the persistent topology of data* (2008),

G. Carlsson, A. Zomorodian, *The theory of multidimensional persistence* (2007),

A. Romero, J. Heras, J. Rubio, F. Sergeraert, *Defining and computing persistent* $\mathbb{Z}$*-homology in the general case* (2014),

S. Basu, L. Parida, *Spectral sequences, exact couples and persistent homology of filtrations* (2015),

B. Matschke, *Successive spectral sequences* (2013),

A. G., F. Vaccarino, *Multidimensional persistence and spectral sequences*, in preparation.