Week 13

Boundary matrix of a filtration

Let $\mathcal{F}: K_0 \subseteq K_1 \subseteq \cdots \subseteq K_r = K$ be a filtration for K. The boundary matrix associated to $\mathcal{F}$ stores info about the faces of every simplex in K.

We begin by placing an ordering on the simplices of K that is compatible with the filtration:

- A face of a simplex precedes the simplex.
- If $i<j$, a simplex in K_i precedes simplices in K_j which are not in K_i.

Let n denote the total number of simplices in K, and let

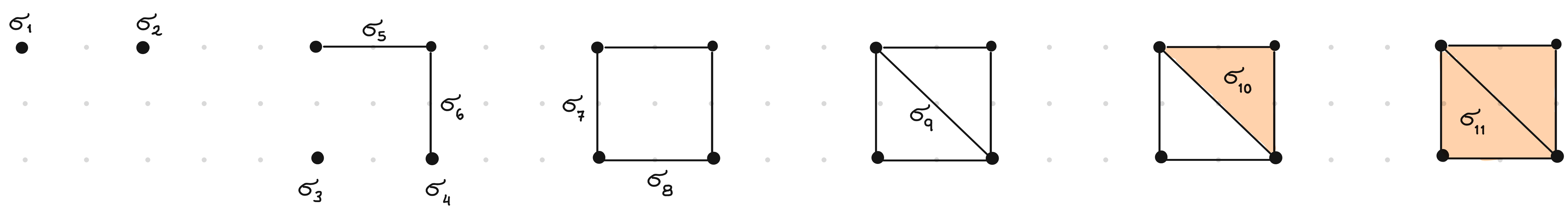
$\sigma_1, \sigma_2, \dots, \sigma_n$ denote the simplices with respect to this ordering.

Terminology: If a simplex τ is a face of a simplex σ, we say τ is a face of codimension k if $\dim\sigma - \dim\tau = k$.

We construct a square matrix δ of dim $n \times n$ as follows

$$\delta = \begin{bmatrix} \delta_{11} & \cdots & \delta_{1n} \\ \vdots & \ddots & \vdots \\ \delta_{n1} & \cdots & \delta_{nn} \end{bmatrix} \qquad \delta_{ij} = \begin{cases} 1, & \text{if } \sigma_i \text{ is a face of } \sigma_j \text{ of codimension } 1. \\ 0, & \text{otherwise.} \end{cases}$$

Ex: $K_0 \hookrightarrow K_1 \hookrightarrow K_2 \hookrightarrow K_3 \hookrightarrow K_4 \hookrightarrow K_5$



Observe:

- Columns 1 to 4 are trivial for 2 reasons:

  A vertex is a face of itself of codim zero.

  For all $5 \leq i \leq 10$, δ_i is not a face of δ_{i+1}.

- Rows 10 and 11 are trivial for 2 reasons:

  For all $1 \leq j \leq 9$, δ_i is not a face of δ_j for $i = 10, 11$

  σ_i is a face of itself of codim zero for $i = 10, 11$

$$\delta = \begin{array}{cc} & \begin{matrix} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 & \sigma_6 & \sigma_7 & \sigma_8 & \sigma_9 & \sigma_{10} & \sigma_{11} \end{matrix} \\ \begin{matrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \\ \sigma_7 \\ \sigma_8 \\ \sigma_9 \\ \sigma_{10} \\ \sigma_{11} \end{matrix} & \begin{bmatrix} & & & & 1 & & 1 & & 1 & & \\ & & & & 1 & 1 & & & & & \\ & & & & & & 1 & 1 & & & \\ & & & & & 1 & & 1 & 1 & & \\ & & & & & & & & & 1 & \\ & & & & & & & & & 1 & \\ & & & & & & & & & & 1 \\ & & & & & & & & & & 1 \\ & & & & & & & & & 1 & 1 \\ & & & & & & & & & & \\ & & & & & & & & & & \end{bmatrix} \end{array}$$

Standard algorithm

Once one has constructed the boundary matrix, one has to reduce it using Gaussian elimination. We will explain the standard algorithm for PH with coeff. in $\mathbb{F}_2$.

For every $j \in \{1, \dots, n\}$, define

$$\text{low}(j) := \text{largest index value } i \text{ such that } \delta_{ij} \neq 0.$$

fix a column, choose the largest index of rows s.t. $\delta_{ij} = 1$

low(j) is undefined if column j is trivial.

We say δ is reduced if the map low is injective on its domain:

Let $\{i_1, \dots, i_m\}$ be the domain of low, where $1 \leq i_j \leq n$ for all j's.

If $low(i_a) = low(i_b)$, then $i_a = i_b$. In other words, for a fixed row were low is defined, there should only be one entry equals to 1.

```
for j = 1 to n do
    while there exists i < j with low(i) = low(j) do
        add column i to column j
    end while
end for
```

In worst case, the complexity of the standard algorithm is cubic in the number of simplices.

Ex:

$$\delta = \begin{array}{cc}
 & \begin{array}{ccccccccccc} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 & \sigma_6 & \sigma_7 & \sigma_8 & \sigma_9 & \sigma_{10} & \sigma_{11} \end{array} \\
\begin{array}{c} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \\ \sigma_7 \\ \sigma_8 \\ \sigma_9 \\ \sigma_{10} \\ \sigma_{11} \end{array} &
\begin{bmatrix}
 & & & & 1 & & 1 & & 1 & & \\
 & & & & 1 & 1 & & & & & \\
 & & & & & & 1 & 1 & & & \\
 & & & & & 1 & & 1 & 1 & & \\
 & & & & & & & & & 1 & \\
 & & & & & & & & & 1 & \\
 & & & & & & & & & & 1 \\
 & & & & & & & & & & 1 \\
 & & & & & & & & & 1 & 1 \\
 & & & & & & & & & & \\
 & & & & & & & & & &
\end{bmatrix}
\end{array}$$

low is undefined for $i = 1, 2, 3, 4$

$$\{5, \dots, 11\} \xrightarrow{\text{low}} \{1, \dots, 9\}$$

$$\begin{array}{ccc}
5 & \longmapsto & 2 \\
6 & \longmapsto & 4 \\
7 & \longmapsto & 3 \\
8 & \longmapsto & 4 \\
9 & \longmapsto & 4 \\
10 & \longmapsto & 9 \\
11 & \longmapsto & 9
\end{array}$$

$j = 1, 2, 3, 4,$ nothing happens.

$j = 5$, $\quad i < 5 \quad$ low(i) undefined $\quad$ nothing happens

$j = 6$, $\quad 5 < 6 \quad$ low$(5) \neq$ low(6) $\quad$ nothing happens

$j = 7$, $\quad 5 < 7 \quad$ low$(5) \neq$ low(7) $\quad$ nothing happens

$\qquad\quad 6 < 7 \quad$ low$(5) \neq$ low(7) $\quad$ nothing happens

$j = 8$, $\quad 5 < 8 \quad$ low$(5) \neq$ low(8) $\quad$ nothing happens

$\qquad\quad 6 < 8 \quad$ low$(6) =$ low(8)

add column 6 to column 8

$\qquad\quad 7 < 8 \quad$ low$(7) =$ new low(8)

add column 7 to column 8

$\qquad\quad 5 < 8 \quad$ low$(5) =$ new low(8)

add column 5 to column 8

$$
\begin{array}{c|ccccccc|}
 & \sigma_5 & \sigma_6 & \sigma_7 & \sigma_8 & \sigma_9 & \sigma_{10} & \sigma_{11} \\
\hline
\sigma_1 & 1 & & 1 & & 1 & & \\
\sigma_2 & 1 & 1 & & & & & \\
\sigma_3 & & & 1 & & & & \\
\sigma_4 & & 1 & & & 1 & & \\
\sigma_5 & & & & & & 1 & \\
\sigma_6 & & & & & & 1 & \\
\sigma_7 & & & & & & & 1 \\
\sigma_8 & & & & & & & 1 \\
\sigma_9 & & & & & & 1 & 1 \\
\hline
\end{array}
$$

$j = 9,$ $\quad 5 < 9 \quad$ low(5) $\neq$ low(9) nothing happens

$6 < 9 \quad$ low(6) $=$ low(9)

add column 6 to column 9

$7 < 9 \quad$ low(7) $\neq$ new low(9) nothing happens

$8 < 9 \quad$ low(8) $\neq$ new low(9) nothing happens

$5 < 9 \quad$ low(5) $=$ new low(9)

add column 5 to column 9

$$
\begin{array}{c|c}
 & \sigma_5 \;\; \sigma_6 \;\; \sigma_7 \;\; \sigma_8 \;\; \sigma_9 \;\; \sigma_{10} \;\; \sigma_{11} \\
\begin{array}{c} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \\ \sigma_7 \\ \sigma_8 \\ \sigma_9 \end{array} &
\begin{bmatrix}
1 & & 1 & & & & \\
1 & 1 & & & & & \\
 & & 1 & & & & \\
 & 1 & & & & & \\
 & & & & & 1 & \\
 & & & & & 1 & \\
 & & & & & & 1 \\
 & & & & & & 1 \\
 & & & & & 1 & 1
\end{bmatrix}
\end{array}
$$

$j = 10,$ $\quad i < 10 \quad$ low(i) $\neq$ low(10) nothing happens

$i = 5, \dots, 9$

$j = 11,$ $\quad i < 11 \quad$ low(i) $\neq$ low(11) nothing happens

$i = 5, \dots, 9$

$10 < 11 \quad$ low(10) $=$ low(11)

add column 10 to column 11

$$
\begin{array}{c|c}
 & \\
\begin{array}{c} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \\ \sigma_7 \\ \sigma_8 \\ \sigma_9 \end{array} &
\begin{bmatrix}
1 & & 1 & & & & \\
1 & 1 & & & & & \\
 & & 1 & & & & \\
 & 1 & & & & & \\
 & & & & & 1 & 1 \\
 & & & & & 1 & 1 \\
 & & & & & & 1 \\
 & & & & & & 1 \\
 & & & & & 1 &
\end{bmatrix}
\end{array}
$$

Therefore the reduction of δ is

$$
\overline{\delta} = \begin{array}{c}
\\ \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \\ \sigma_7 \\ \sigma_8 \\ \sigma_9 \\ \sigma_{10} \\ \sigma_{11}
\end{array}
\begin{array}{c}
\begin{array}{ccccccccccc} \sigma_1 & \sigma_2 & \sigma_3 & \sigma_4 & \sigma_5 & \sigma_6 & \sigma_7 & \sigma_8 & \sigma_9 & \sigma_{10} & \sigma_{11} \end{array} \\
\begin{bmatrix}
 & & & & 1 & & 1 & & & & \\
 & & & & 1 & 1 & & & & & \\
 & & & & & & 1 & & & & \\
 & & & & & 1 & & & & & \\
 & & & & & & & & & 1 & 1 \\
 & & & & & & & & & 1 & 1 \\
 & & & & & & & & & & 1 \\
 & & & & & & & & & & 1 \\
 & & & & & & & & & 1 & \\
 & & & & & & & & & & \\
 & & & & & & & & & &
\end{bmatrix}
\end{array}.
$$

Reading off persistence pairs

- If $low(j)$ is undefined, then the entrance of σ_j in $\mathcal{F}$ causes the birth of a class. If there exists k s.t. $low(k) = j$, then σ_j is paired with with σ_k, whose entrance in $\mathcal{F}$ causes the death of the class. If no such k exists, then σ_j is unpaired.
- If $low(j) = i$, then the simplex σ_j is paired with σ_i, and the entrance of σ_i in $\mathcal{F}$ causes the birth of a class that dies with the entrance of σ_j.

Terminology: Given a simplex σ in K, we define $dg(\sigma)$ to be the smallest number ℓ such that $\sigma \in K_\ell$.

- A pair (σ_i, σ_j) gives the persistence pair $(dg(\sigma_i), dg(\sigma_j))$.
- An unpaired simplex σ_k gives the persistence pair $(dg(\sigma_k), \infty)$.

Ex: Pairs from matrix $\bar{\partial}$ | Persistence pair

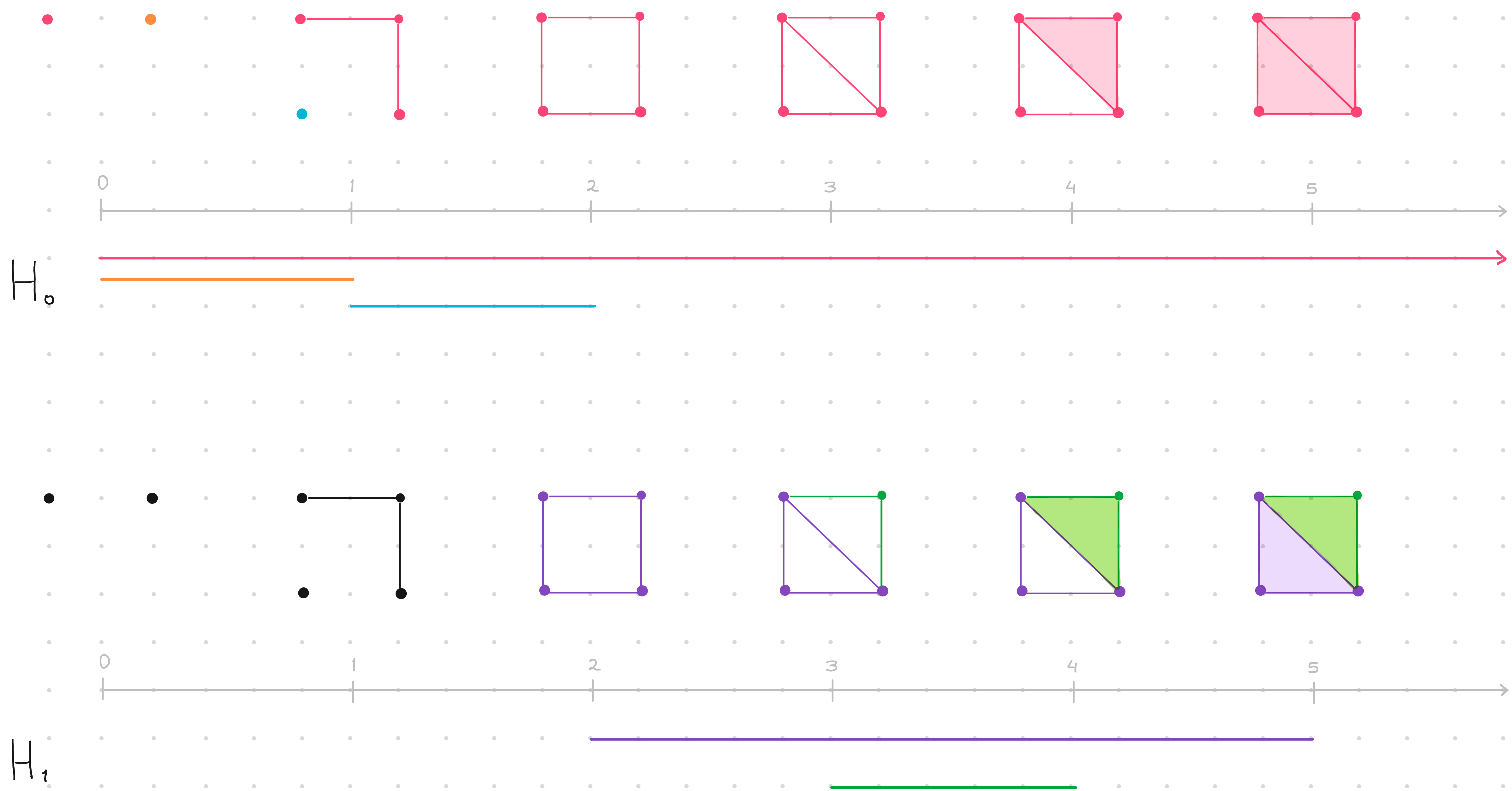
$j=1$ there is no k with $low(k) = 1$ entrance of σ_1 causes the birth of class | $[\sigma_1]$ $(0, \infty)$

$j=2$ $\underset{\text{exists}}{low(5)} = 2 \Rightarrow (\sigma_2, \sigma_5)$ entrance of σ_2 causes the birth of a class & entrance of σ_5 causes its death. | $[\sigma_2]$ $(0, 1)$

$j=3$ $\underset{\text{exists}}{low(7)} = 3 \Rightarrow (\sigma_3, \sigma_7)$ entrance of σ_3 causes the birth of a class & entrance of σ_7 causes its death. | $[\sigma_3]$ $(1, 2)$

$j=4$ $\underset{\text{exists}}{low(6)} = 4 \Rightarrow (\sigma_4, \sigma_6)$ σ_4 and σ_6 are born at the same time | no pair

$j = 5$	$low(5) = 2$	$\Rightarrow (\sigma_2, \sigma_5)$	entrance of σ_2 causes the birth of a class that dies with the entrance of σ_5	$[\sigma_2]$	$(0, 1)$
$j = 6$	$low(6) = 4$	$\Rightarrow (\sigma_4, \sigma_6)$		no pair	
$j = 7$	$low(7) = 3$	$\Rightarrow (\sigma_3, \sigma_7)$	entrance of σ_3 causes the birth of a class that dies with the entrance of σ_7	$[\sigma_3]$	$(1, 2)$
$j = 8$	$low(11) = 8$ exists	$\Rightarrow (\sigma_8, \sigma_{11})$	entrance of σ_8 causes the birth of a class & entrance of σ_{11} causes its death.	$\sum_{i=5}^{8} [\sigma_i]$	$(2, 5)$
$j = 9$	$low(10) = 9$ exists	$\Rightarrow (\sigma_9, \sigma_{10})$		$\sum_{i=7}^{9} [\sigma_i]$	$(3, 4)$
$j = 10$	$low(10) = 9$	$\Rightarrow (\sigma_9, \sigma_{10})$		$\sum_{i=7}^{9} [\sigma_i]$	$(3, 4)$
$j = 11$	$low(11) = 8$	$\Rightarrow (\sigma_8, \sigma_{11})$		$\sum_{i=5}^{8} [\sigma_i]$	$(2, 5)$


0
1
2
3
4
5
H_0
0
1
2
3
4
5
H_1

There exist several algorithms to reduce δ:

Standard algorithm

Spectral - sequence algorithm

Twist algorithm

Chunk algorithm

Multified algorithms

Dual algorithm

Distributed algorithm

A roadmap for the computation of PH, Otter et. al.

After the introduction of the standard algorithm, several new algorithms were developed. Each of these algorithms gives the same output for the computation of PH, so we only give a brief overview and references to these algorithms, as one does not need to know them to compute PH with one of the publicly-available software packages. In Section 7.2, we indicate which implementation of these libraries is best suited to which data set.

As we mentioned in Section 5.3.1, in the worst case, the standard algorithm has cubic complexity in the number of simplices. This bound is sharp, as Morozov gave an example of a complex with cubic complexity in [123]. Note that in cases such as when matrices are sparse, complexity is less than cubic. Milosavljević, Morozov, and Skraba [124] introduced an algorithm for the reduction of the boundary matrix in $\mathcal{O}(n^\omega)$, where ω is the matrix-multiplication coefficient (i.e., $\mathcal{O}(n^\omega)$ is the complexity of the multiplication of two square matrices of size n). At present, the best bound for ω is 2.376 [125]. Many other algorithms have been proposed for the reduction of the boundary matrix. These algorithms give a heuristic speed-up for many data sets and complexes (see the benchmarkings in the forthcoming references), but they still have cubic complexity in the number of simplices. Sequential algorithms include the `twist algorithm` [126] and the `dual algorithm` [72, 127]. (Note that the dual algorithm is known to give a speed-up when one computes PH with the VR complex, but not necessarily for other types of complexes (see also the results of our benchmarking for the **vertebra** data set in Additional file 1 of the SI).) Parallel algorithms in a shared setting include the `spectral-sequence algorithm` (see Section VII.4 of [80]) and the `chunk algorithm` [128]; parallel algorithms in a distributed setting include the `distributed algorithm` [74]. The `multifield algorithm` is a sequential algorithm that allows the simultaneous computation of PH over several fields [129].