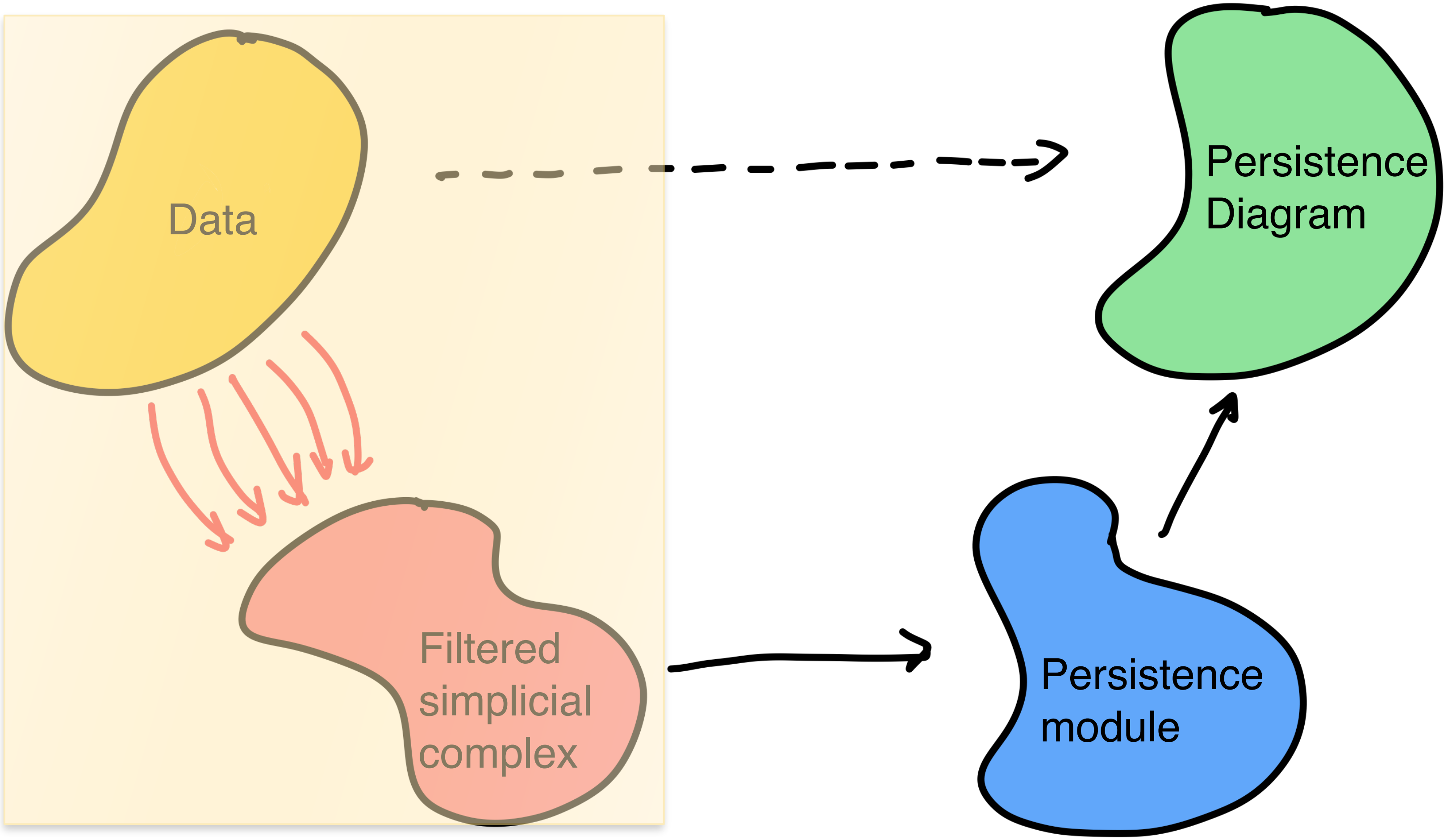
MATH412/COMPSCI434/MATH713
Fall 2025
Topological Data Analysis

Topic 4: Introduction to Persistent Homology - Part 1

Instructor: Ling Zhou

Filtrations

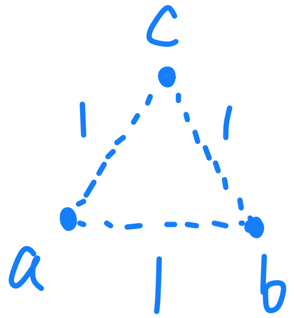
Filtered Simplicial complex



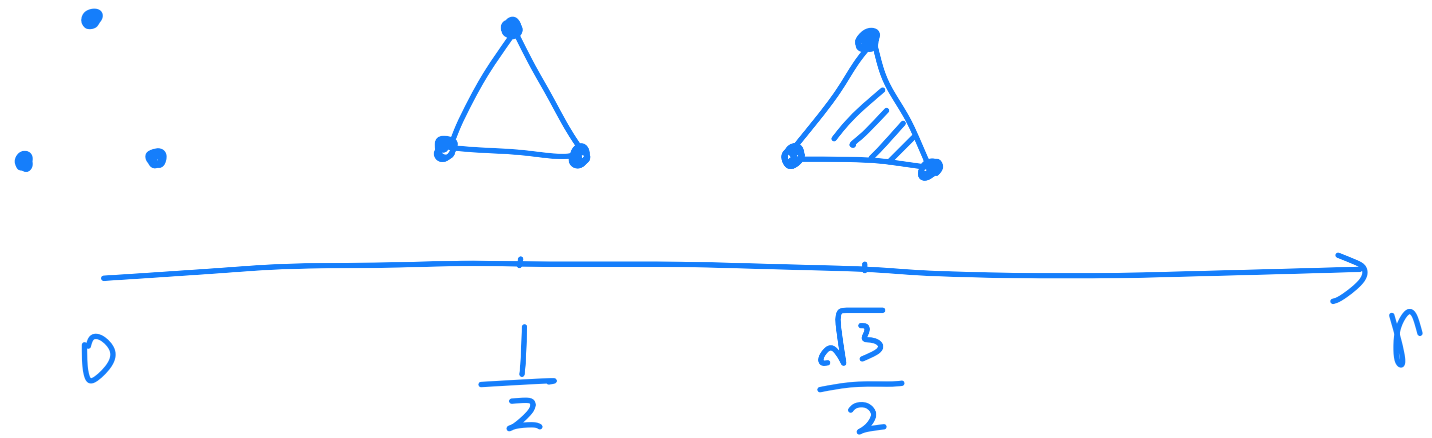
Čech Filtration

$C^r(P) = \{ p_{i_1} \cdots p_{i_k} \mid \bigcap_j B(p_{i_j}, r) \neq \phi \}$

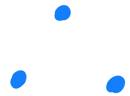
- Given a set of points $P = \{p_1, \ldots, p_n\} \subset \mathbb{R}^d$
- $(C^r(P))_{r \geq 0}$ is called the Čech filtration

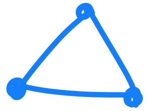


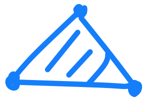
$C^r(\{a,b,c\})$:



$C^r(\{a,b,c\}) =$

 if $r \in [0, \frac{1}{2})$

 if $r \in [\frac{1}{2}, \frac{\sqrt{3}}{2})$

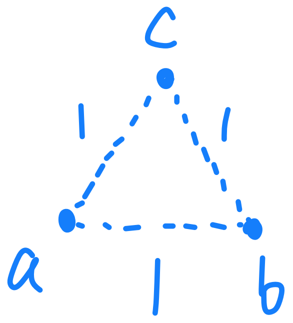
 if $r \geq \frac{\sqrt{3}}{2}$

Vietoris-Rips (Rips) Filtration

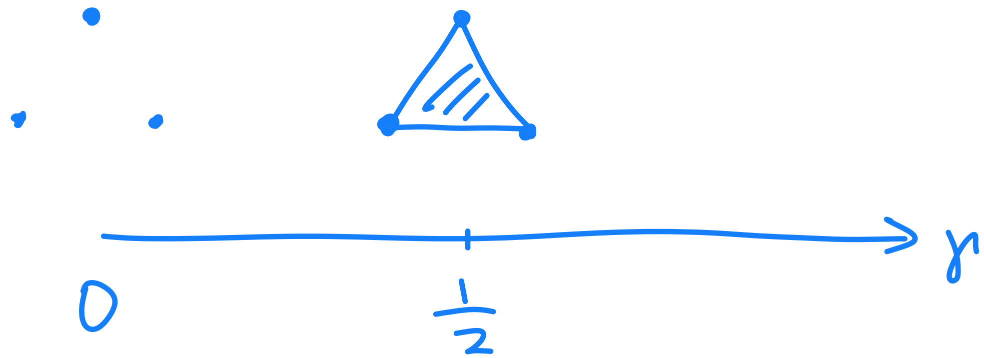
$\mathrm{Rips}^r(P) = \{ p_{i_1} \cdots p_{i_k} \mid d(p_{i_j}, p_{i_\ell}) \leq 2r \}$

- Given a set of points $P = \{p_1, \ldots, p_n\} \subset \mathbb{R}^d$
- $(Rips^r(P))_{r \geq 0}$ is called the Vietoris-Rips (Rips) Filtration

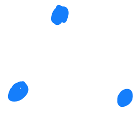
Another convention is to use r instead.



$\mathrm{Rips}^r(\{a,b,c\})$



$\mathrm{Rips}^r(\{a,b,c\}) =$



if $r \in [0, \frac{1}{2})$

if $r \in [\frac{1}{2}, \infty)$

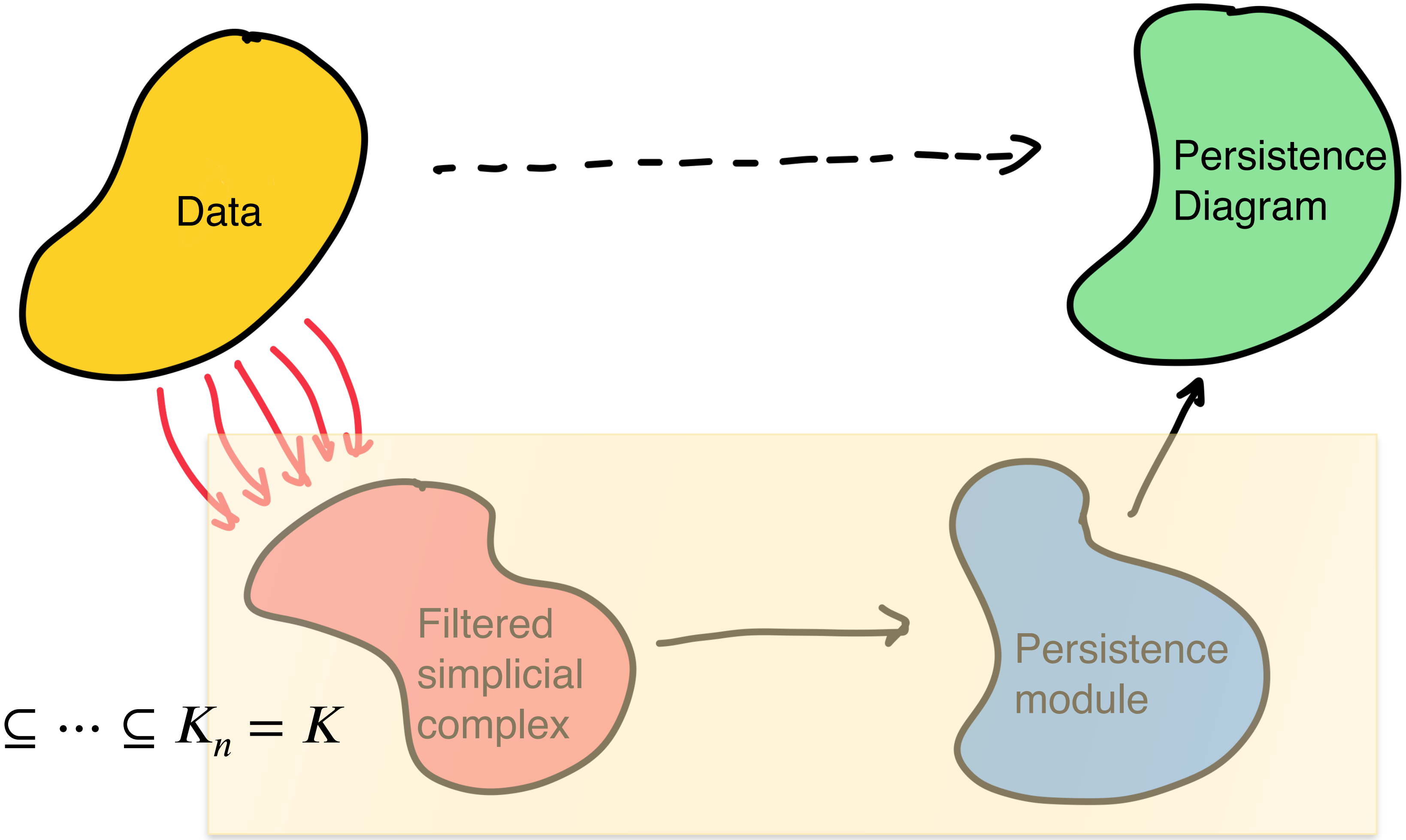
Finitely Presented filtration

- A filtration $(K_t)_{t\in[0,\infty)}$ is called **finitely represented** if
  - There exist $0 = t_0 < t_1 < \cdots < t_n$ such that
  - $K_t = K_{t'}, \quad \forall t_i \leq t < t' < t_{i+1}$ and $i = 0,\ldots,n$ $(t_{n+1} := \infty)$

- So $(K_t)_{t\in[0,\infty)}$ is essentially the same as (or can be reconstructed from) $(K_{t_i})_{i=0,\ldots,n}$: $K_0 \hookrightarrow K_1 \hookrightarrow \ldots \hookrightarrow K_n$
- Both Čech and Rips filtrations of a finite set P are finitely represented

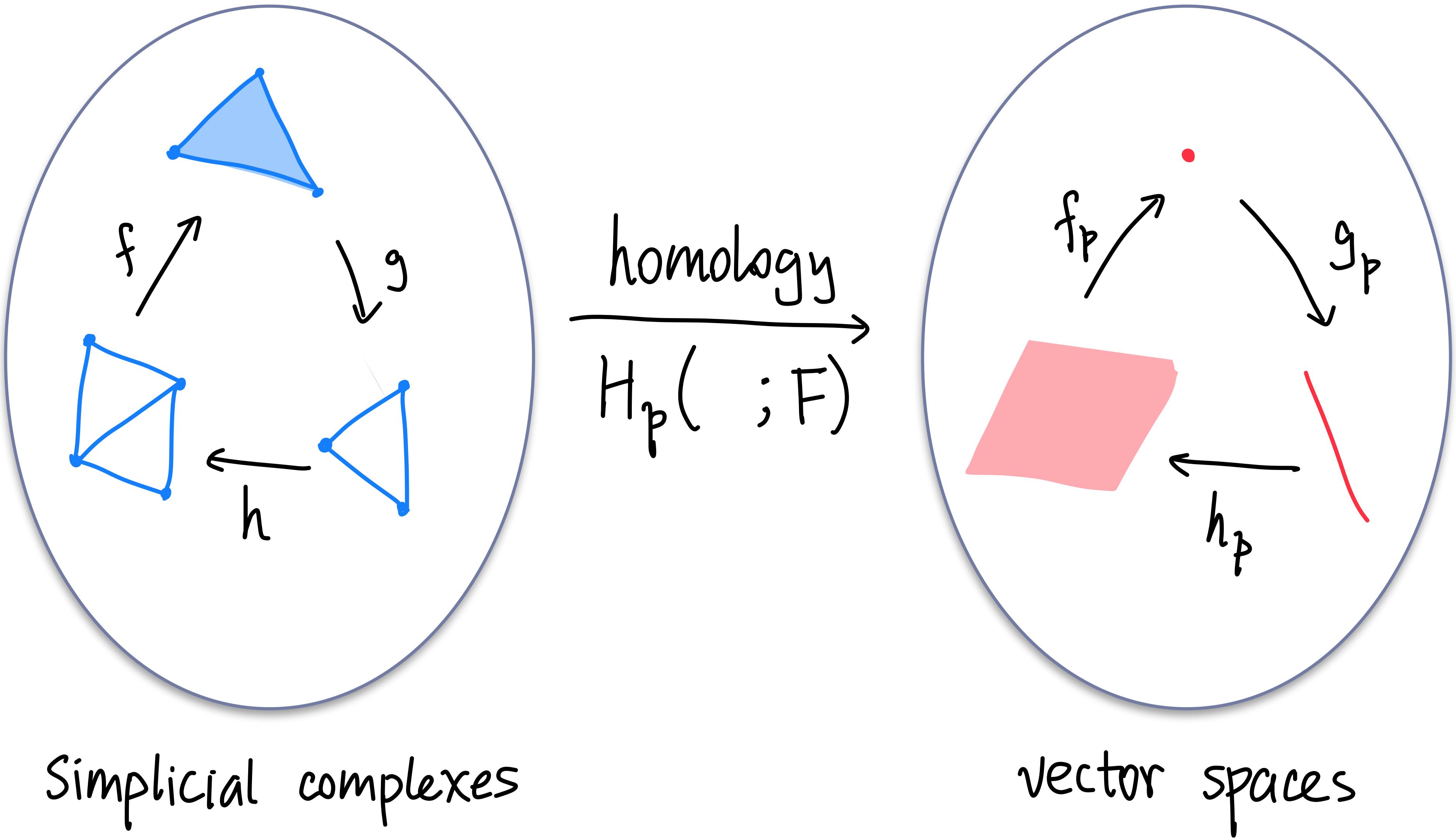
Persistent Homology

Persistence Modules



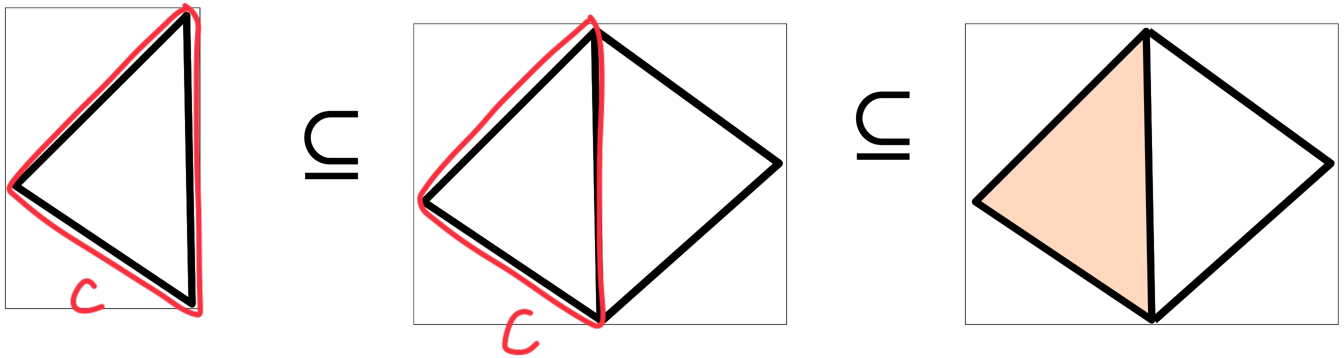
$$K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$$

Recall: Functoriality of homology



Persistent Homology

- $\iota : K \hookrightarrow K' \implies \iota_p : H_p(K) \to H_p(K')$
  - Simplicial maps (e.g. the above inclusion) induce homomorphisms in homology groups (under Z_2-coefficients, linear maps in vector spaces)



$$
\begin{array}{ccccc}
K_1 & \subseteq & K_2 & \subseteq & K_3 \\
H_1(K_1) & \to & H_1(K_2) & \to & H_1(K_3) \\
[c] & \mapsto & [c] & \mapsto & 0
\end{array}
$$

Persistent Homology

- $\imath : K \hookrightarrow K' \implies \imath_p : H_p(K) \to H_p(K')$
  - Simplicial maps (e.g. the above inclusion) induce homomorphisms in homology groups (under Z_2-coefficients, linear maps in vector spaces)

- Let $K_\bullet : K_0 \hookrightarrow K_1 \hookrightarrow \ldots \hookrightarrow K_n$ be a **simplicial filtration**, i.e, a sequence of simplicial complexes connected by inclusions $\imath^{i,i+1} : K_i \hookrightarrow K_{i+1}$.
- Denote $\imath^{i,j} := \imath^{i,i+1} \circ \ldots \circ \imath^{j-1,j} : K_i \hookrightarrow K_j$ and $\imath_*^{i,j} : H_*(K_i) \to H_*(K_j)$
- **Persistent homology** of the filtration $K_\bullet$ is
  - $H_*(K_\bullet) := \left\{ H_*(K_i) \xrightarrow{\imath_*^{i,j}} H_*(K_j) : 0 \le i \le j \le n \right\}$, i.e.
  - $H_*(K_\bullet) : H_*(K_0) \xrightarrow{\imath_*^{0,1}} H_*(K_1) \xrightarrow{\imath_*^{1,2}} \ldots \xrightarrow{\imath_*^{n-1,n}} H_*(K_n)$

Persistence Modules

Persistence Modules

- A **persistence module (or persistent vector space)** $V_\bullet$ over a field $\mathbb{F}$ is
  - a sequence of vector spaces $\{V_i\}_{i=0,\ldots,n}$
  - together with linear maps $L_{i,j} : V_i \to V_j$ for $i \le j$ such that
    - $L_{i,i} = Id_{V_i}$
    - For $i \le j \le k$, $L_{i,k} = L_{j,k} \circ L_{i,j}$
- Write $V_\bullet = \{L_{i,j} : V_i \to V_j\}$ or $V_\bullet : V_0 \xrightarrow{L_{0,1}} V_1 \xrightarrow{L_{1,2}} \ldots \xrightarrow{L_{n-1,n}} V_n$ or simply $V_\bullet = \{V_i\}$
- **Persistent homology is a persistence module**:

$$H_*(K_\bullet) : H_*(K_0) \xrightarrow{\iota_*^{0,1}} H_*(K_1) \xrightarrow{\iota_*^{1,2}} \ldots \xrightarrow{\iota_*^{n-1,n}} H_*(K_n)$$

Persistence Modules

$$e_i = (0, \cdots, 0, \overset{\substack{i\text{-th} \\ \downarrow}}{1}, 0, \cdots, 0)$$

Example

- $\langle e_1 \rangle \longrightarrow \langle e_1, e_2 \rangle$

  $e_1 \longmapsto e_1$

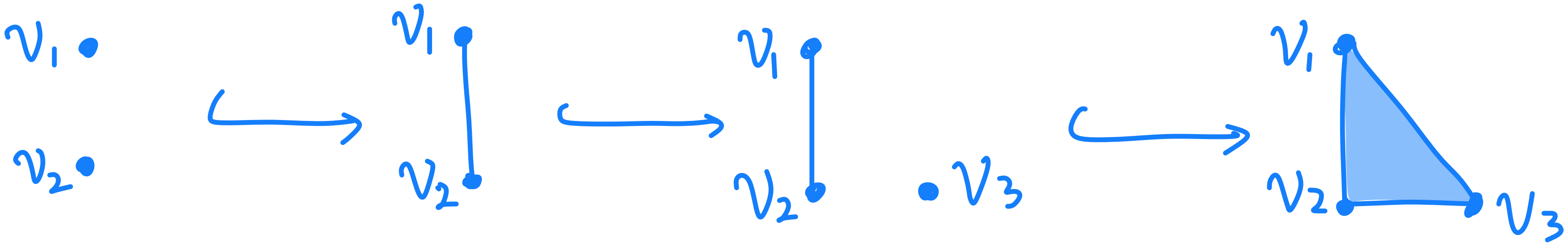
- $\langle e_1, e_2 \rangle \longrightarrow \langle e_1 \rangle$

  $e_1 \longmapsto e_1$

  $e_2 \longmapsto e_1$

- $\langle e_1 \rangle \longrightarrow \langle e_1, e_2, \cdots e_n \rangle \longrightarrow \langle e_2, e_3, e_5 \rangle \longrightarrow 0$

Persistent Homology (PH) as Persistence Modules



$$K_0 \hookrightarrow K_1 \hookrightarrow K_2 \hookrightarrow K_3$$

$$\xRightarrow{PH_0}$$

$$\begin{tikzcd}
H_0(K_0) \arrow[r] \arrow[d, equal] & H_0(K_1) \arrow[r] \arrow[d, equal] & H_0(K_2) \arrow[r] \arrow[d, equal] & H_0(K_3) \arrow[d, equal] \\
\langle [v_1], [v_2] \rangle \arrow[r] & \langle [v_1] \rangle \arrow[r] & \langle [v_1], [v_3] \rangle \arrow[r] & \langle [v_1] \rangle
\end{tikzcd}$$

$$\begin{tikzcd}
{[v_1]} \arrow[r, mapsto] & {[v_1]} \arrow[r, mapsto] & {[v_1]} \arrow[r, mapsto] & {[v_1]} \\
{[v_2]} \arrow[ur, mapsto] & & {[v_3]} \arrow[ur, mapsto] &
\end{tikzcd}$$

Maps Between Persistence Modules

- Let $\{V_i\}$ and $\{W_i\}$ be two persistence modules
  - a sequence of linear maps $\{\varphi_i : V_i \to W_i\}_{i=0,\ldots,n}$ is called a **linear transformation** from $\{V_i\}$ to $\{W_i\}$ if for any $i \leq j$

$$\begin{tikzcd} V_i \arrow[r, "L^V_{i,j}"] \arrow[d, "\varphi_i"'] & V_j \arrow[d, "\varphi_j"] \\ W_i \arrow[r, "L^W_{i,j}"'] & W_j \end{tikzcd}$$

  - φ is called an isomorphism if each φ_i is an **isomorphism**

Direct Sum of Persistence Modules

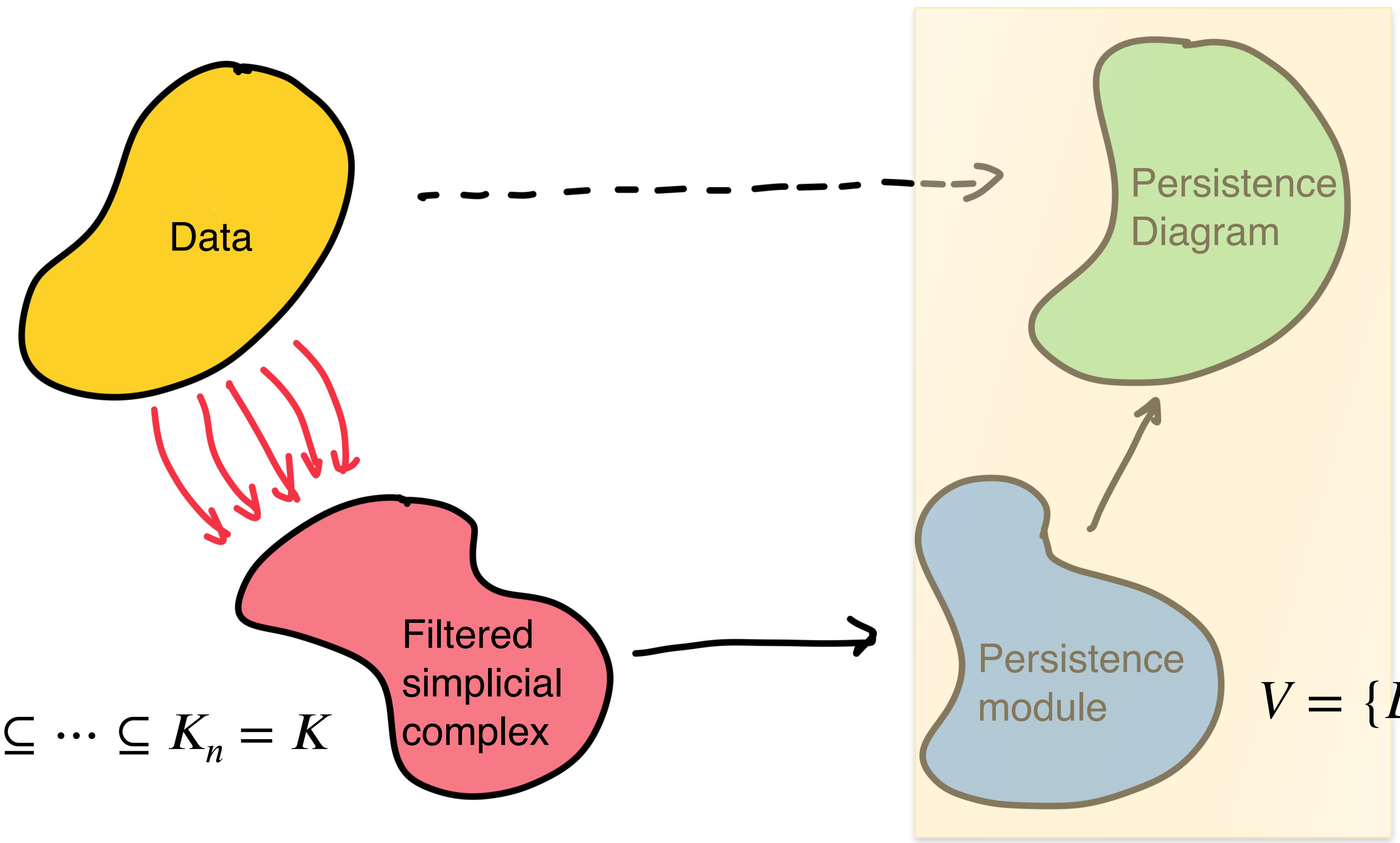
- Let $\{V_i\}$ and $\{W_i\}$ be two persistence modules
- The **direct sum** $V \oplus W$ is the collection
  - the collection of vector spaces $\{(V \oplus W)_i := V_i \oplus W_i\}$, and
  - maps defined by $L_{i,j}^{V \oplus W}(v, w) := (L_{i,j}^{V} \oplus L_{i,j}^{W})(v, w) = (L_{i,j}^{V}(v), L_{i,j}^{W}(w))$

$$
\begin{array}{ccccccc}
V_0 & \longrightarrow & V_1 & \longrightarrow & \cdots & \longrightarrow & V_n \\
\oplus & \longrightarrow & \oplus & \longrightarrow & \oplus & \longrightarrow & \oplus \\
W_0 & \longrightarrow & W_1 & \longrightarrow & \cdots & \longrightarrow & W_n
\end{array}
$$

- Dimension and basis are the most important objects of a vector space
  - $\dim(V) = \dim(W) \iff \exists$ isomorphism $\varphi : V \to W$
  - basis determines a vector space

- What are “dimension” and “basis” for a persistence vector space?
  - $\text{"dim"}(V_\bullet) = \text{"dim"}(W_\bullet) \iff \exists$ isomorphism $\varphi : V_\bullet \to W_\bullet$
  - “basis” determines a persistent vector space

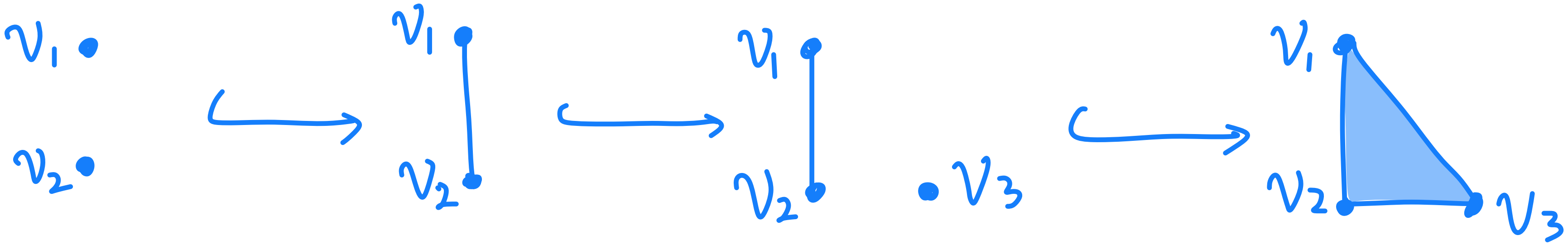
Persistence Diagram



$$K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$$

$$V = \{L_{i,j} : V_i \to V_j\}$$

Decomposition of Persistence Modules



$$K_0 \hookrightarrow K_1 \hookrightarrow K_2 \hookrightarrow K_3$$

$$\begin{tikzcd} \arrow[r, Rightarrow, "PH_0"] & H_0(K_0) \arrow[r] \arrow[d, phantom, "\shortparallel"] & H_0(K_1) \arrow[r] \arrow[d, phantom, "\shortparallel"] & H_0(K_2) \arrow[r] \arrow[d, phantom, "\shortparallel"] & H_0(K_3) \arrow[d, phantom, "\shortparallel"] \\ & \langle [v_1], [v_2] \rangle \arrow[r] & \langle [v_1] \rangle \arrow[r] & \langle [v_1], [v_3] \rangle \arrow[r] & \langle [v_1] \rangle \end{tikzcd}$$

$$\begin{tikzcd} {[v_1]} \arrow[r, mapsto] & {[v_1]} \arrow[r, mapsto] & {[v_1]} \arrow[r, mapsto] & {[v_1]} \\ {[v_2]} \arrow[ur, mapsto] & & {[v_3]} \arrow[ur, mapsto] & \end{tikzcd}$$

Decomposition of Persistence Modules

$$
\begin{tikzcd}
H_0(K_\bullet): & H_0(K_0) \arrow[r] \arrow[d, equal] & H_0(K_1) \arrow[r] \arrow[d, equal] & H_0(K_2) \arrow[r] \arrow[d, equal] & H_0(K_3) \arrow[d, equal] \\
& \langle [v_1], [v_2] \rangle \arrow[r] & \langle [v_1] \rangle \arrow[r] & \langle [v_1], [v_3] \rangle \arrow[r] & \langle [v_1] \rangle
\end{tikzcd}
$$

$$
\begin{tikzcd}
{[v_1]} \arrow[r, mapsto] & {[v_1]} \arrow[r, mapsto] & {[v_1]} \arrow[r, mapsto] & {[v_1]} \\
{[v_2]} \arrow[ur, mapsto] & & {[v_3]} \arrow[ur, mapsto] &
\end{tikzcd}
$$

Decompose

$$
\begin{array}{ccccccccc}
& & \langle [v_1] \rangle & \longrightarrow & \langle [v_1] \rangle & \longrightarrow & \langle [v_1] \rangle & \longrightarrow & \langle [v_1] \rangle \\
& & & & \oplus & & & & \\
H_0(K_\bullet) & = & \langle [v_2 - v_1] \rangle & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 \\
& & & & \oplus & & & & \\
& & 0 & \longrightarrow & 0 & \longrightarrow & \langle [v_3 - v_1] \rangle & \longrightarrow & 0
\end{array}
$$

Decomposition of Persistence Modules

$$\langle[v_1]\rangle \longrightarrow \langle[v_1]\rangle \longrightarrow \langle[v_1]\rangle \longrightarrow \langle[v_1]\rangle$$

$$\oplus$$

$$H_0(K_\bullet) = \langle[v_2 - v_1]\rangle \longrightarrow 0 \longrightarrow 0 \longrightarrow 0$$

$$\oplus$$

$$0 \longrightarrow 0 \longrightarrow \langle[v_3 - v_1]\rangle \longrightarrow 0$$

Discard "basis" to only track "dimension"

$$\mathbb{F} \longrightarrow \mathbb{F} \longrightarrow \mathbb{F} \longrightarrow \mathbb{F}$$

$$\oplus$$

$$H_0(K_\bullet) \cong \mathbb{F} \longrightarrow 0 \longrightarrow 0 \longrightarrow 0$$

$$\oplus$$

$$0 \longrightarrow 0 \longrightarrow \mathbb{F} \longrightarrow 0$$

Interval Persistence Modules

- Given the index set $I = \{0, \ldots, n\}$
- Let $0 \leq b < d \leq n+1$, the **interval persistence module**, denoted by $I[b, d)$ is defined as
- $I[b, d) : 0 \to \cdots \to 0 \to \underset{b\text{th position}}{\mathbb{F}} \to \mathbb{F} \to \cdots \to \underset{d-1\text{th position}}{\mathbb{F}} \to 0 \to \cdots \to 0$
- $I[b, n+1) : 0 \to \cdots \to 0 \to \mathbb{F} \to \mathbb{F} \to \cdots \to \mathbb{F}$ is often written as $I[b, \infty)$

Decomposition Theorem

- Theorem: Let $V_\bullet = \{V_i\}_{i=0}^{n}$ be any persistence vector space. Then, there exist a collection of $0 \le b_j < d_j \le n+1, j = 1,\ldots,M$ such that $V_\bullet \cong I[b_1, d_1) \oplus I[b_2, d_2) \oplus \cdots \oplus I[b_M, d_M)$
- The composition is unique up to reordering the summands.

$$
\begin{array}{lccccccccl}
\text{indices:} & 0 & & 1 & & 2 & & 3 & & \\
 & \mathbb{F} & \longrightarrow & \mathbb{F} & \longrightarrow & \mathbb{F} & \longrightarrow & \mathbb{F} & & I[0,4) \ (\text{or } I[\ ,\infty)) \\
 & & & & \oplus & & & & & \oplus \\
H_0(K_\bullet) \cong & \mathbb{F} & \longrightarrow & 0 & \longrightarrow & 0 & \longrightarrow & 0 & = & I[0,1) \\
 & & & & \oplus & & & & & \oplus \\
 & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{F} & \longrightarrow & 0 & & I[2,3)
\end{array}
$$

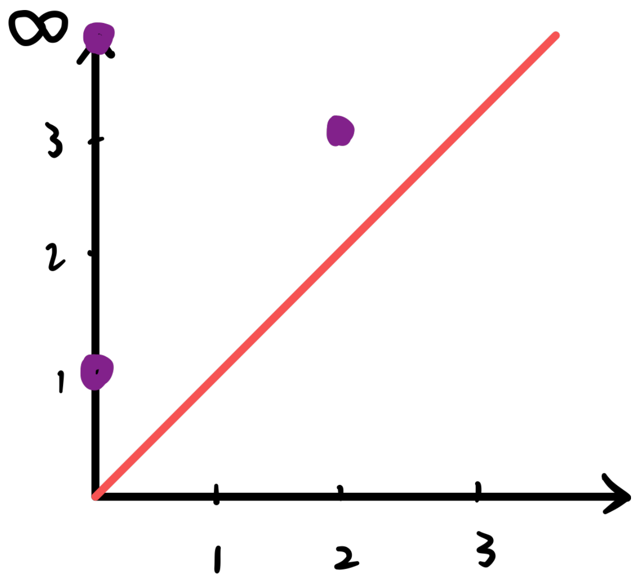
Persistence Diagram and Barcodes

- Assume we have decomposed a persistence module as
  $V_\bullet \cong I[b_1, d_1) \oplus I[b_2, d_2) \oplus \cdots \oplus I[b_M, d_M)$
  - The collection of intervals $\{[b_j, d_j)\}_{j=1,\ldots,M}$ is called the **barcode** of $V_\bullet$
  - The multiset $D = \{(b_j, d_j)\}_{j=1,\ldots,M} \subseteq \mathbb{R}^2$ is called the **persistence diagram** of $V_\bullet$

$H_0(K_\bullet) \cong I[0,\infty) \oplus I[0,1) \oplus I[2,3)$

Barcode

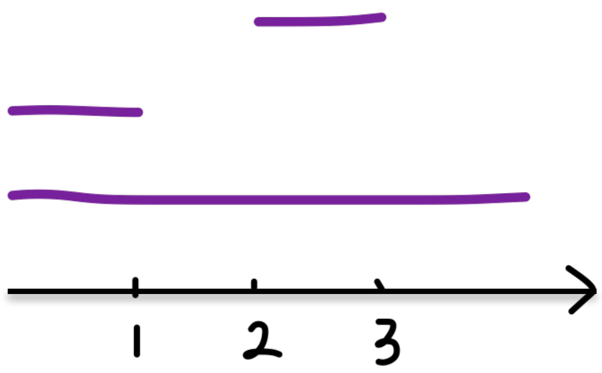
Persistence Diagram



Remark

- Persistence diagrams and barcodes are different ways of representing the same information
- The information they represent serves the role of "dimension" of persistence modules

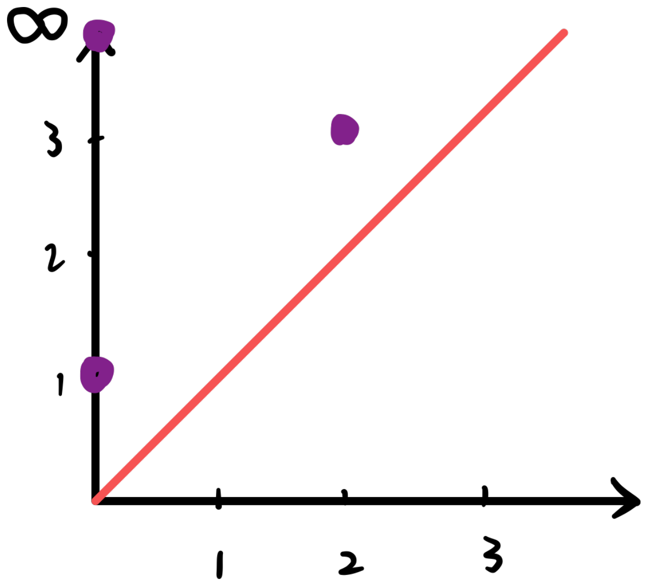
Barcode



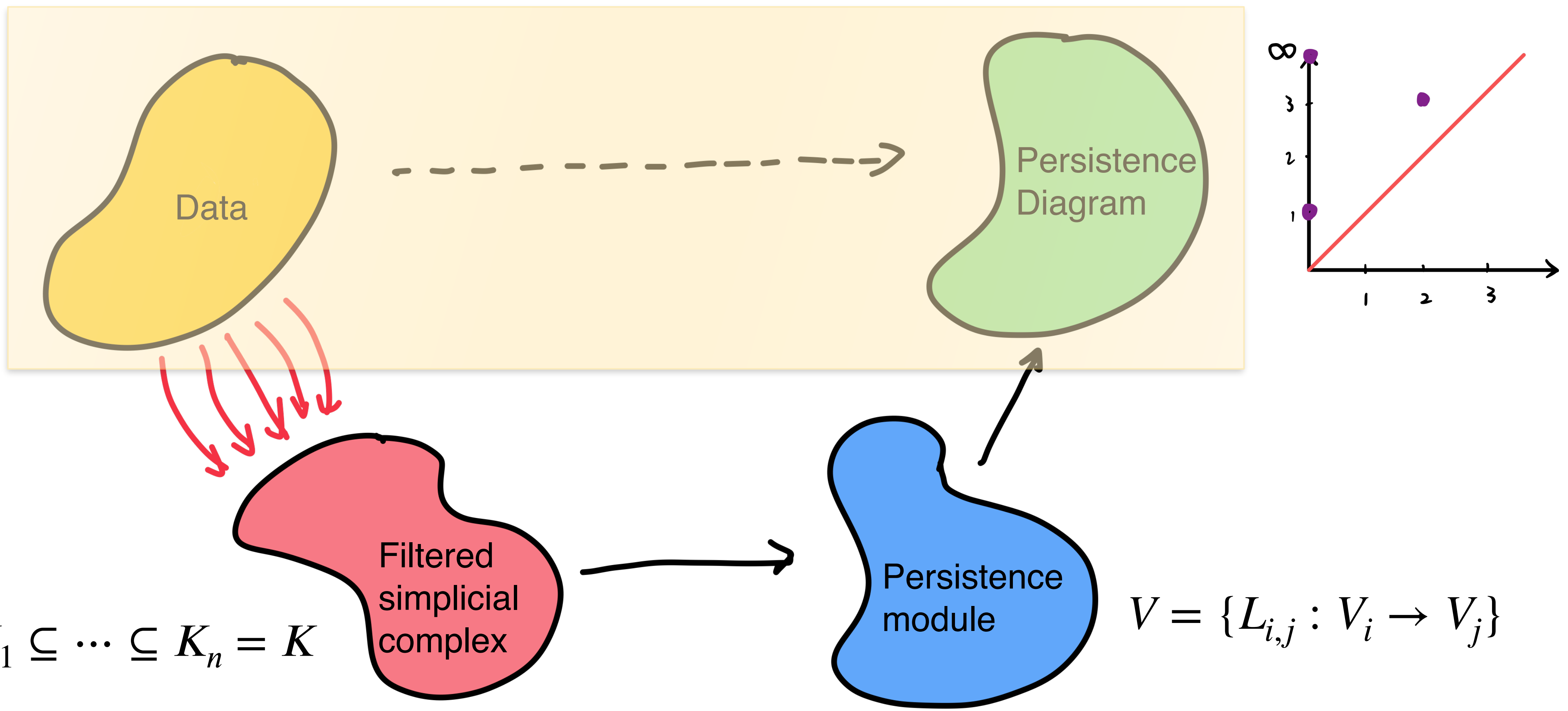
interval   2D point

$[b, d) \longleftrightarrow (b, d)$

Persistence Diagram



TDA in a nutshell



$$K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$$

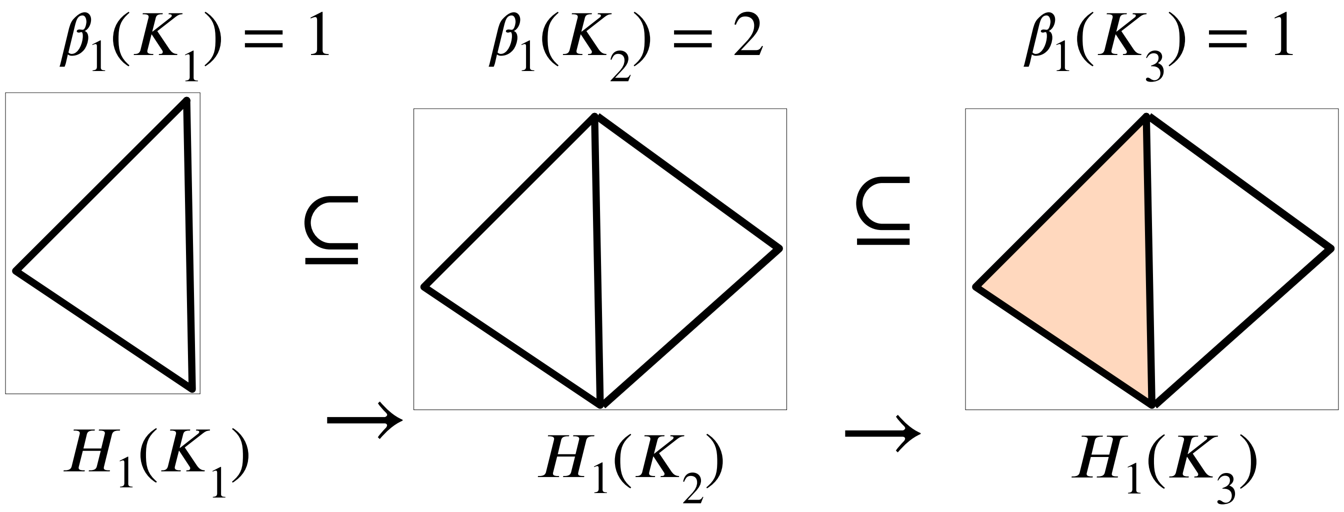
$$V = \{L_{i,j} : V_i \to V_j\}$$

Summary

- Create a filtered simplicial complex $K_\bullet = \{K_i\}$ out of data
- Let $V_\bullet = \{V_i = H_p(K_i)\}_{i=0}^n$ be the p-dim persistence homology of $K_\bullet$
- Decompose $V_\bullet \cong I[b_1, d_1) \oplus I[b_2, d_2) \oplus \cdots \oplus I[b_M, d_M)$
- The multiset $Dgm_p(K_\bullet) = \{(b_j, d_j)\}_{j=1,\ldots,M} \subseteq \mathbb{R}^2$ is called the **degree p persistence diagram** of $K_\bullet$
    - The intervals can repeat themselves

Persistent Betti Number

- p-th **persistent homology group** from i to j: $H_p^{i,j} = \mathrm{Im}(\iota_p^{i,j})\ (\subset H_p(K_j))$
  - Subgroup of $H_p\left(K_j\right)$ that ``existed'' in $H_p\left(K_i\right)$
- p-th **persistent betti number** from i to j: $\beta_p^{i,j} = \dim H_p^{i,j}$
- $\beta_p^{i,j}$ denotes the number of homology classes co-existing at both K_i and K_j



$\beta_1^{1,3} = 0$

$\beta_1^{1,2} = 1$

Persistent Betti Number vs Barcode

- Let $V_\bullet = \{V_i = H_p(K_i)\}_{i=0}^{n}$ be the p-th persistence homology of $K_\bullet$
- Assume that $V_\bullet \cong I[b_1, d_1) \oplus I[b_2, d_2) \oplus \cdots \oplus I[b_M, d_M)$
- $\beta_p^{i,j}$ = # of bars crossing both vertical lines at i and at j

$$\beta_p^{i,j} = \dim H_p^{i,j}$$

$$\beta_p^{2,4} = 3$$