

Quiver Representations, Quiver Varieties and Combinatorics
Bologna, May 2023

Quiver Representations in Topological Data Analysis (TDA)

Steve Oudot

Inria

References

- 1-parameter persistence theory:
 - O. (2015): Persistence Theory: from Quiver Representations to Data Analysis.
- Multi-parameter persistence theory:
 - Botnan, Lesnick (2022): An Introduction to Multi-Parameter Persistence.
- Algorithmic aspects:
 - Dey, Wang (2022): Computational Topology for Data Analysis.
- Statistical aspects:
 - Chazal, Michel (2021): An Introduction to Topological Data Analysis.
- Connection to Machine Learning:
 - Hensel, Moor, Rieck (2021): A Survey of Topological Machine Learning Methods.
- Software: *Gudhi*, *PHAT*, *Ripser*, *Eirene*, *Persistable*, ...

Data featurization

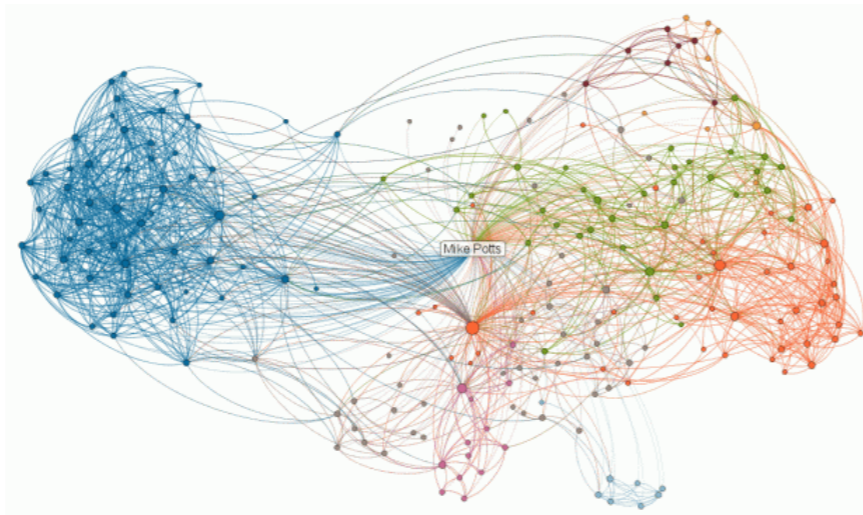
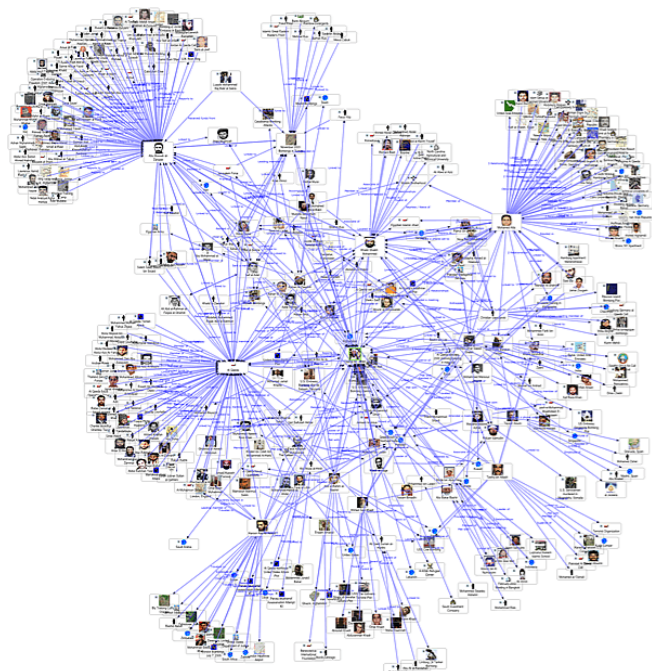
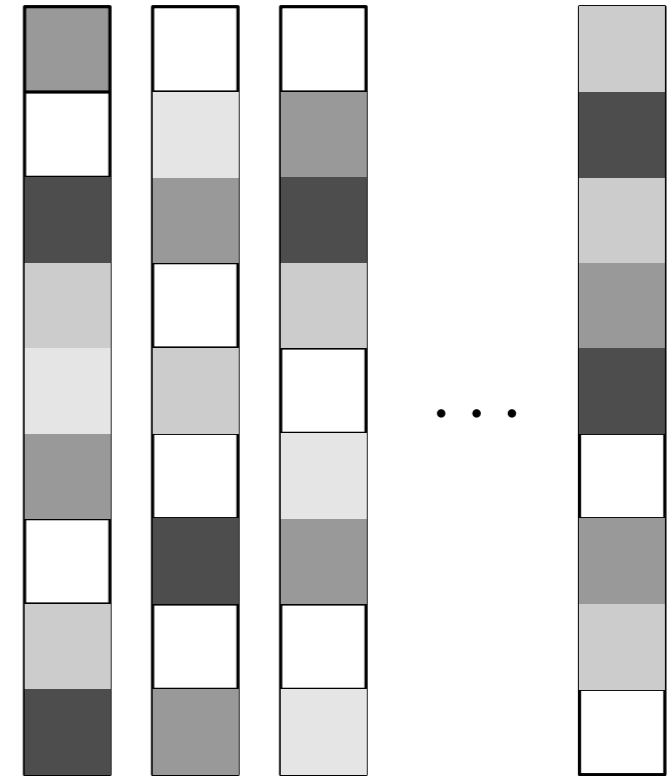
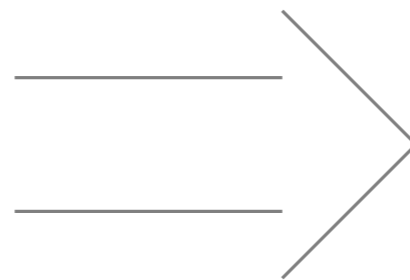


Data

Features

$\in \mathbb{R}^n$

(feature design or learning)

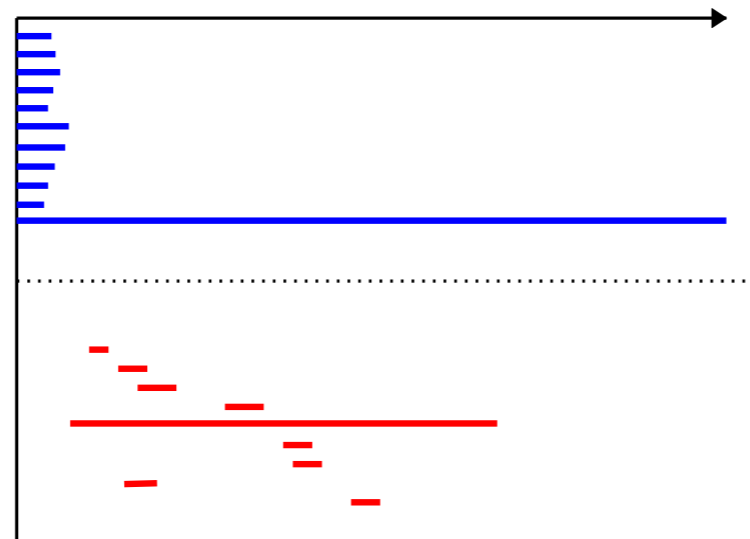
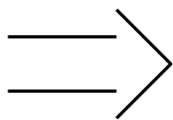


- bag of words, word2vec
- shape contexts, heat kernels
- node2vec, Laplacian fact., rand. walks
- dim. reduction, auto-encoders, etc.

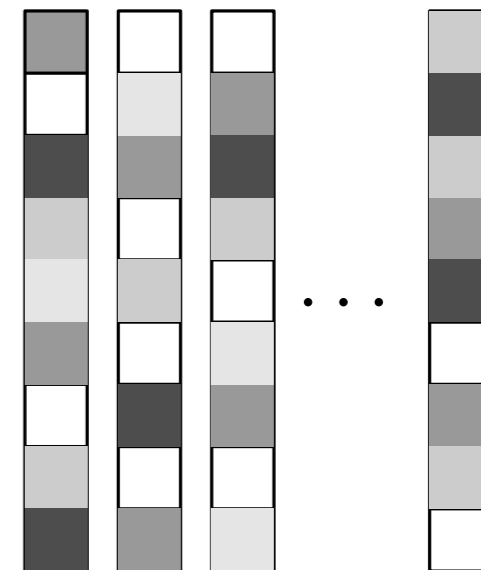
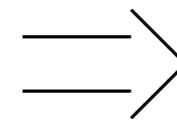
Topological Data Analysis pipeline



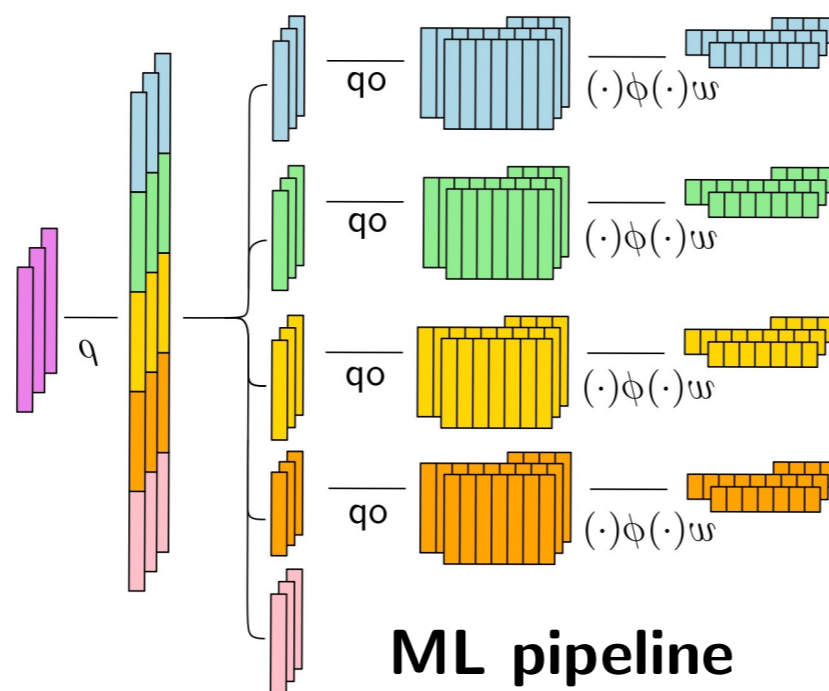
data



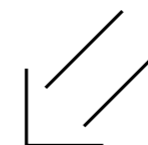
invariant (barcode)



features (vectors)



ML pipeline
(e.g. neural net.)

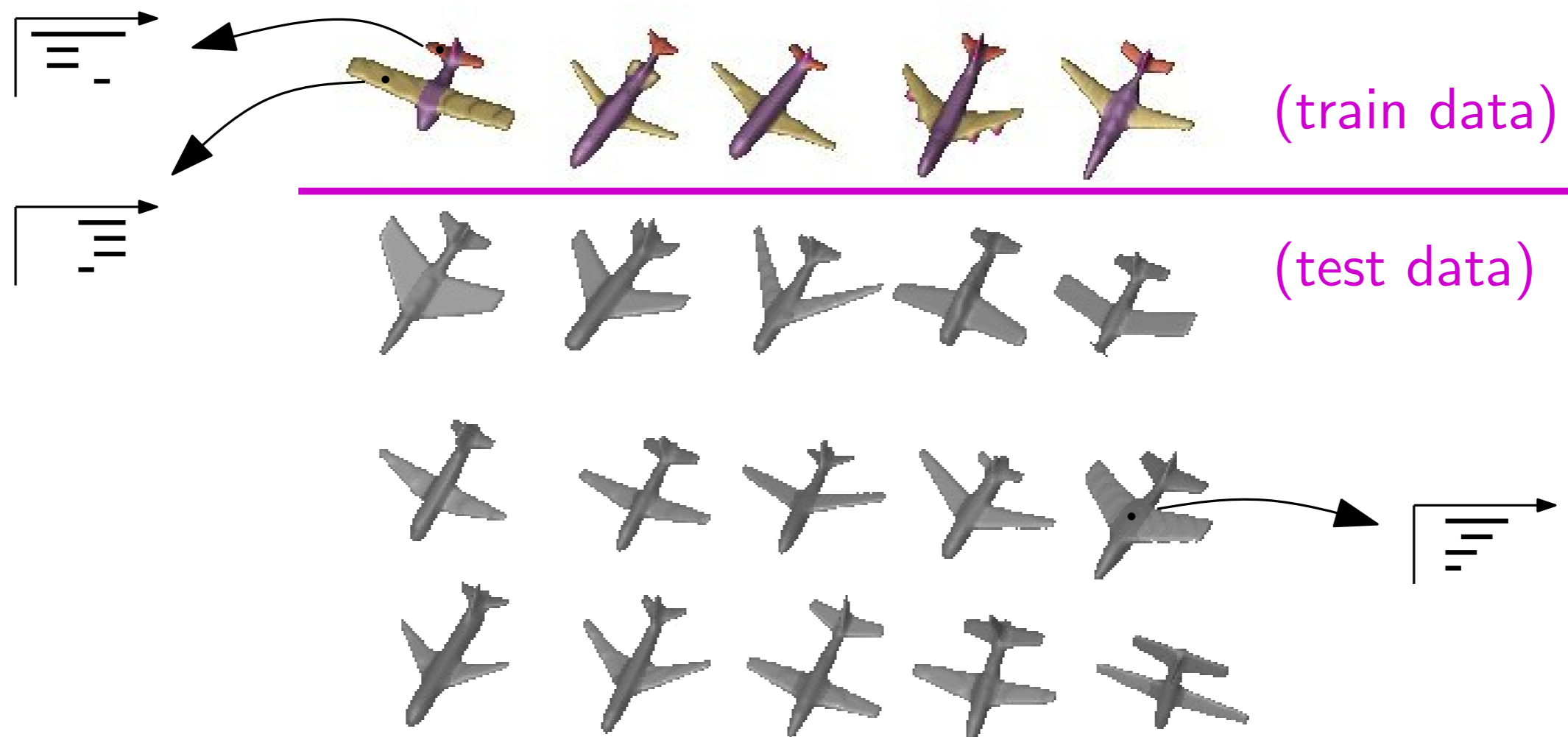


Example of application: shape segmentation

Goal: segment 3d shapes based on examples

Approach:

- train a predictor on **barcodes** extracted from the training shapes
- apply the predictor to **barcodes** extracted from the query shape



Example of application: shape segmentation

Goal: segment 3d shapes based on examples

Approach:

- train a predictor on **barcodes** extracted from the training shapes
- apply the predictor to **barcodes** extracted from the query shape

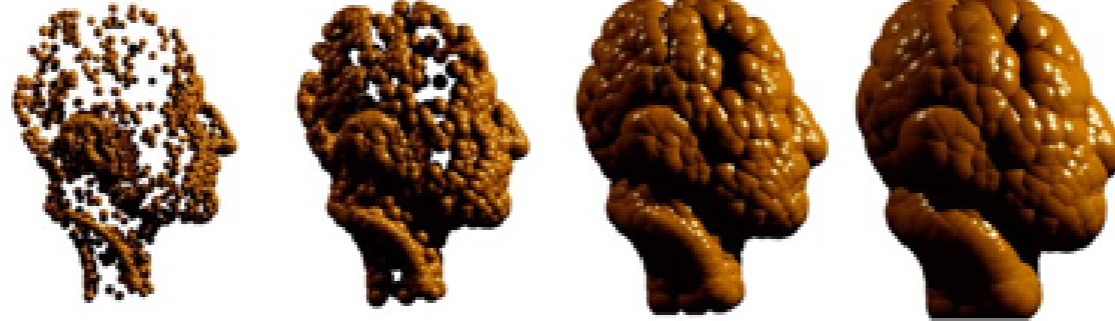
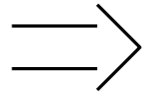
Error rates (%):

	TDA features	geom/stat features	TDA + geom/stat features
Human	26.0	21.3	11.3
Airplane	27.4	18.7	9.3
Ant	7.7	9.7	1.5
FourLeg	27.0	25.6	15.8
Octopus	14.8	5.5	3.4
Bird	28.0	24.8	13.5
Fish	20.4	20.9	7.7

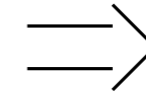
Topological Data Analysis pipeline (again)



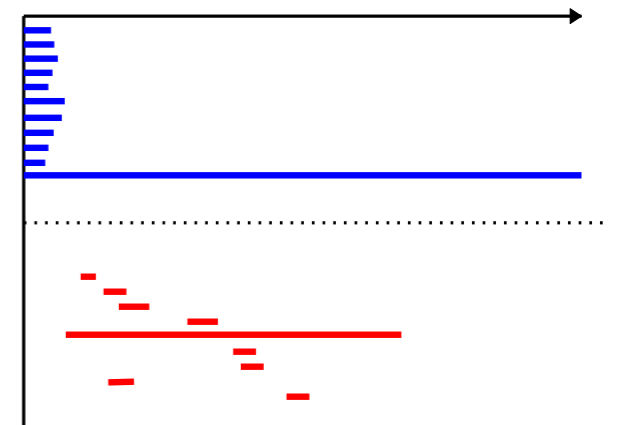
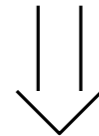
data



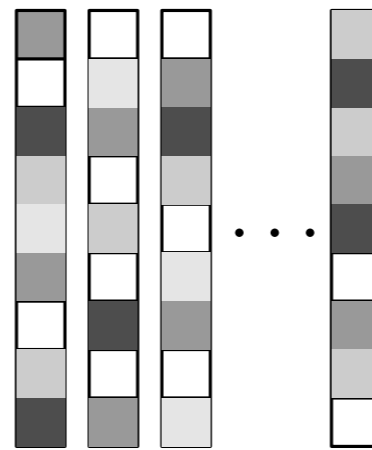
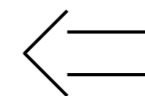
filtration $F : P \subseteq \mathbb{R}^d \rightarrow \text{Top}$
(sublevel sets of distance to data)



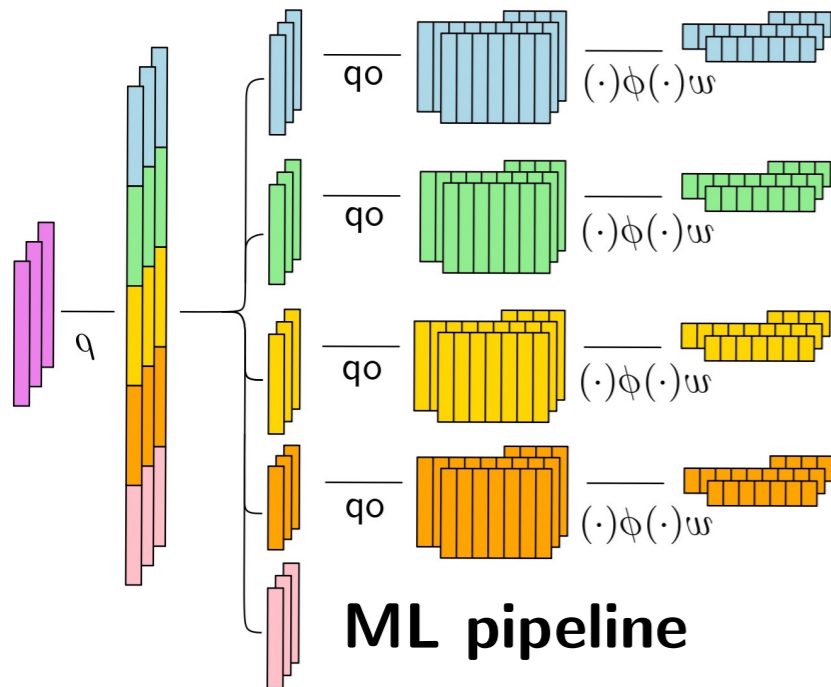
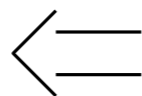
persistence module
 $M = HF : P \rightarrow \text{vect}_k$
(singular homology)



invariant (Bar M)



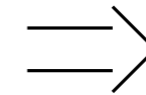
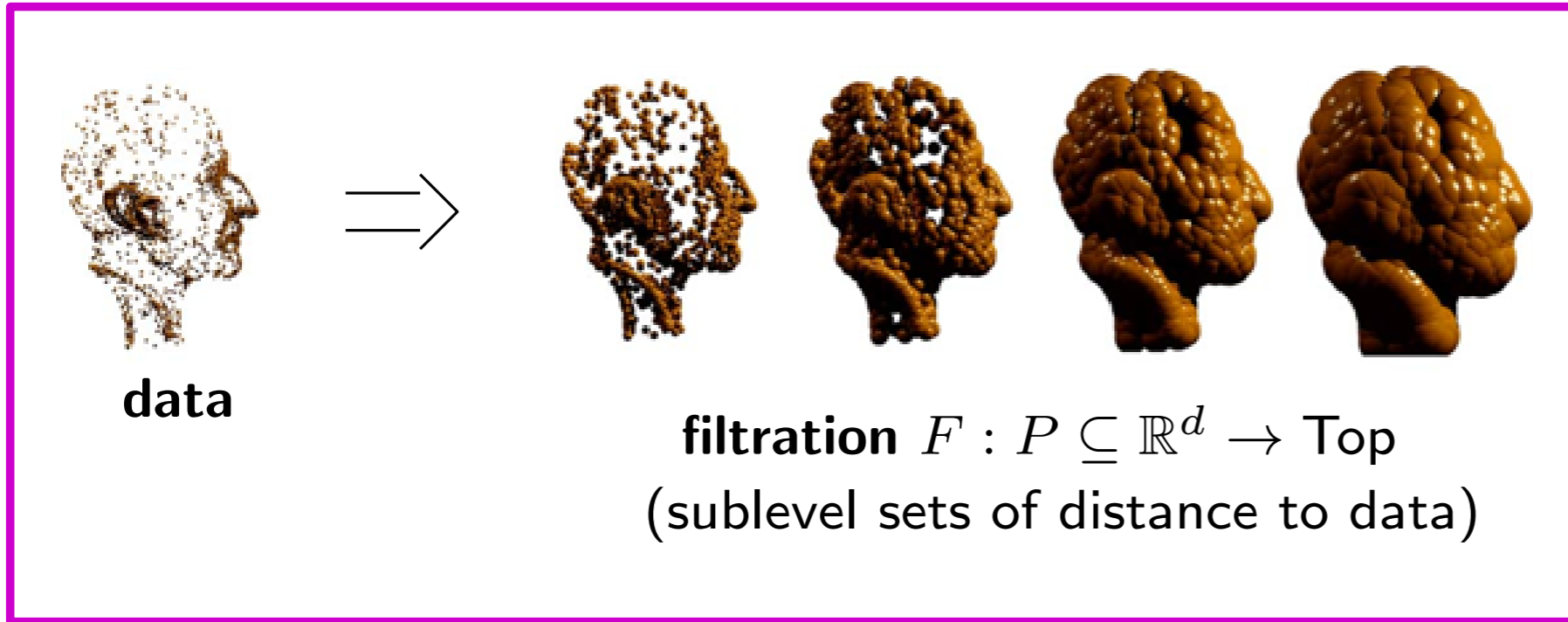
features (vectors)



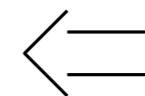
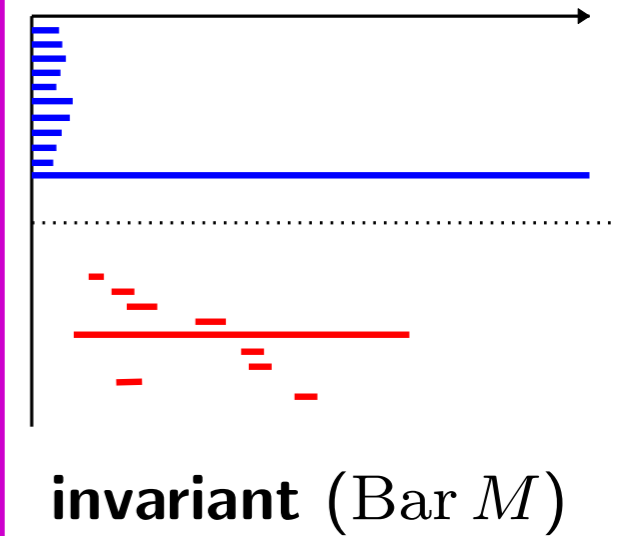
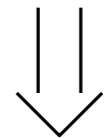
ML pipeline

(e.g. neural net.)

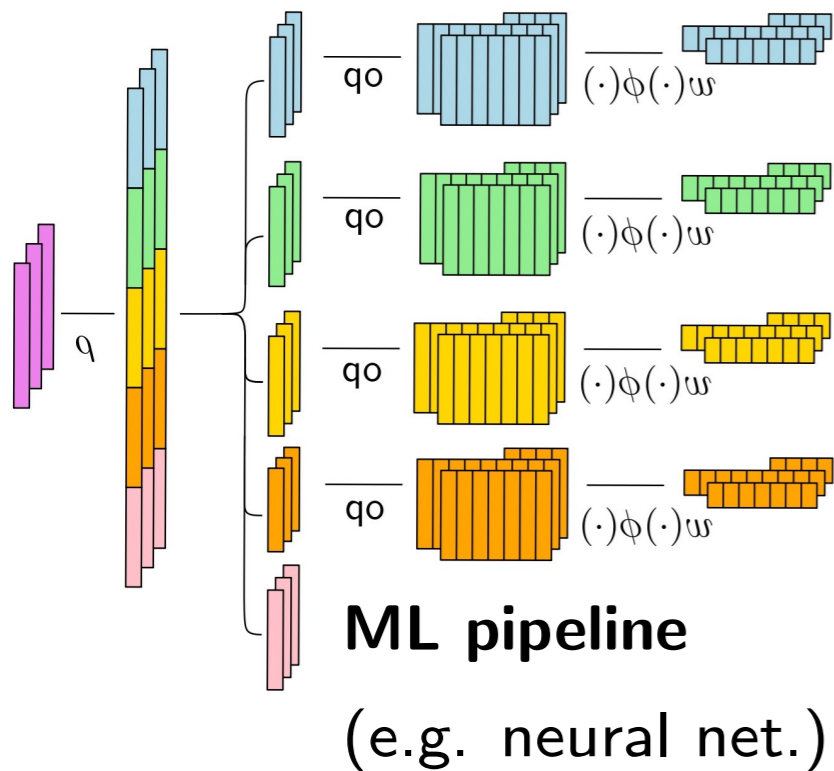
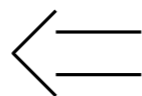
Topological Data Analysis pipeline (again)



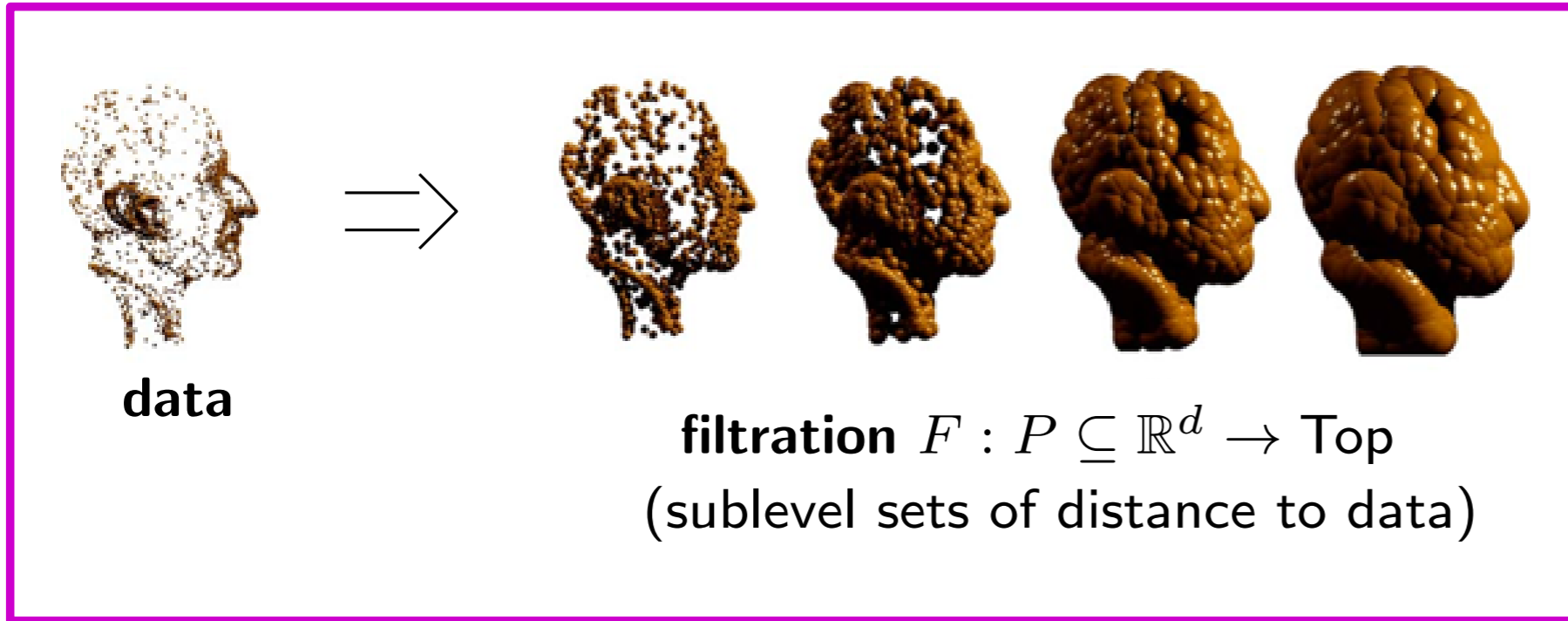
persistence module
 $M = HF : P \rightarrow \text{vect}_k$
 (singular homology)



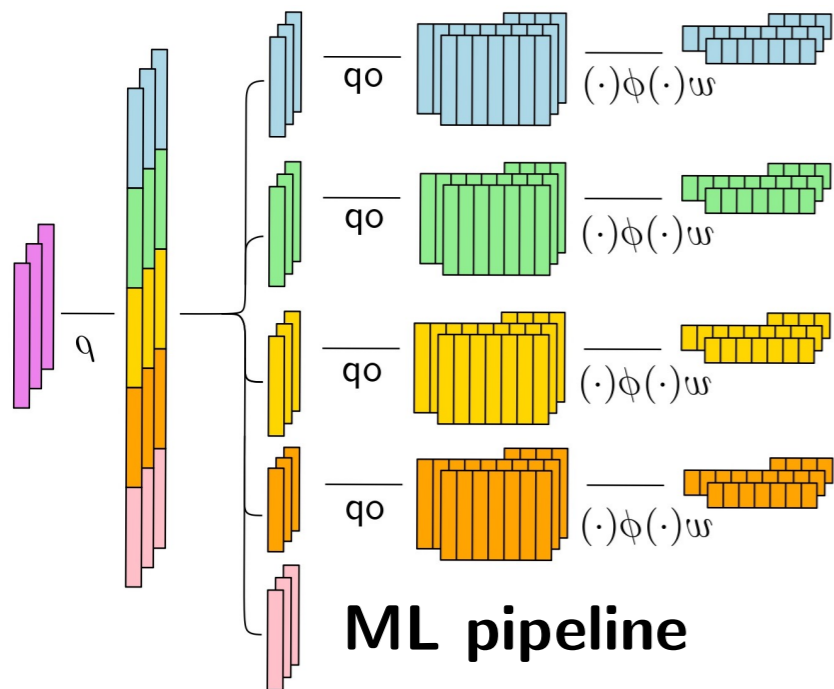
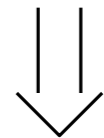
features (vectors)



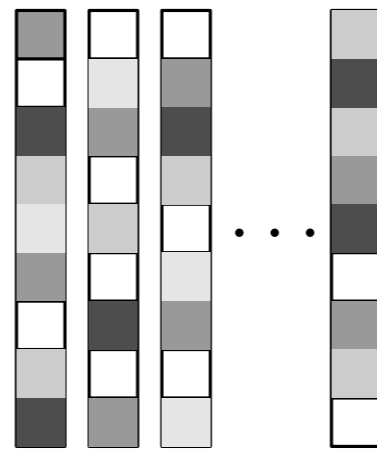
Topological Data Analysis pipeline (again)



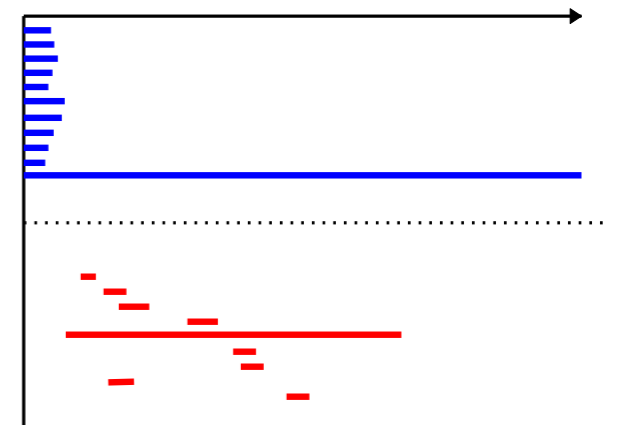
persistence module
 $M = HF : P \rightarrow \text{vect}_k$
 (singular homology)



(e.g. neural net.)



features (vectors)



invariant (Bar M)

Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

Filtration: functor $(P, \leq) \rightarrow \mathbf{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

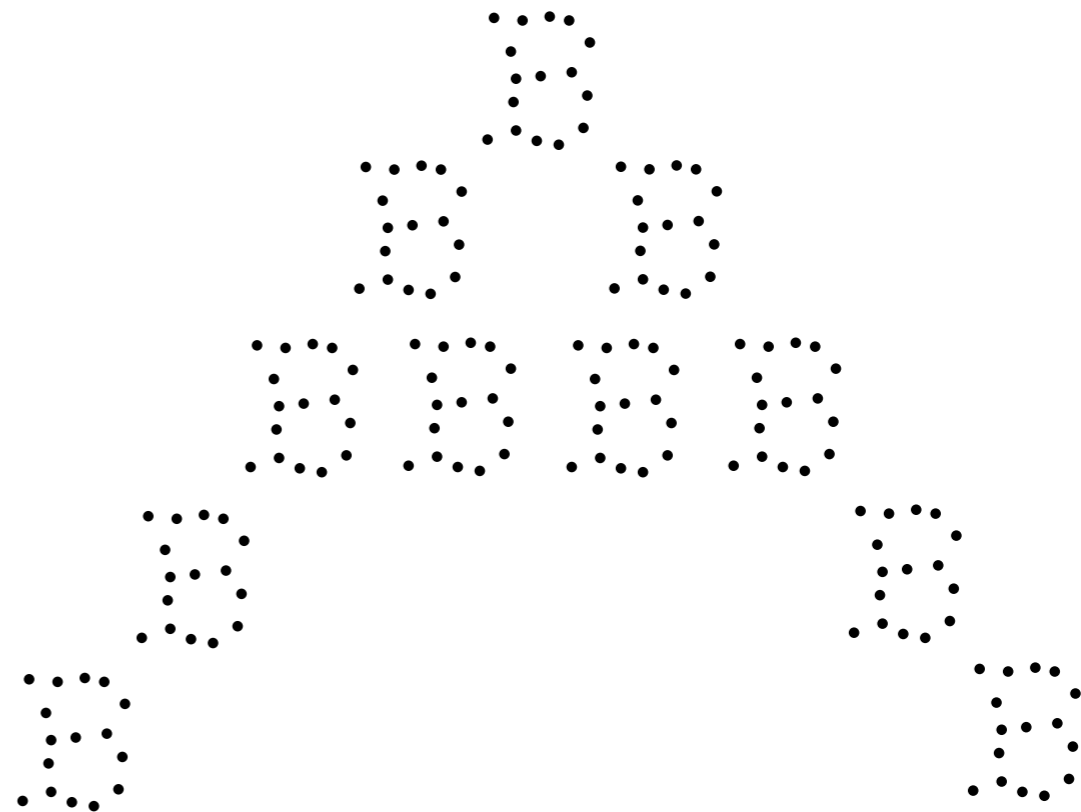
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

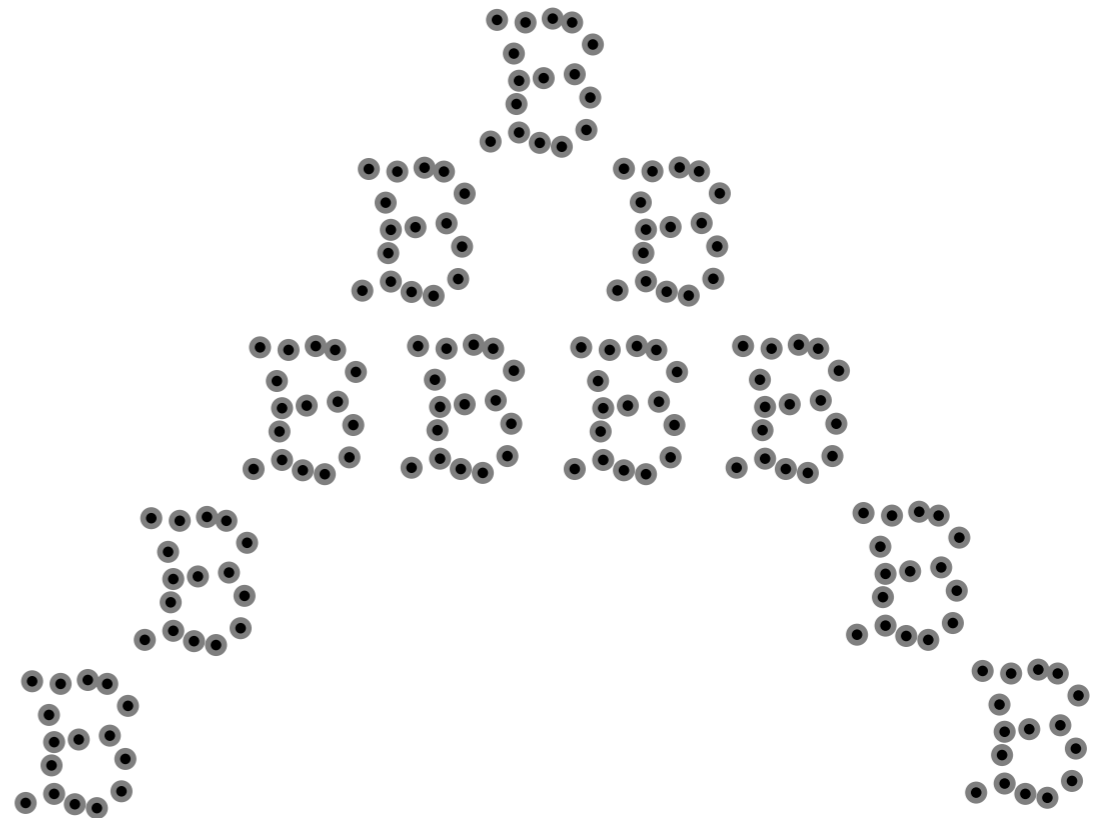
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

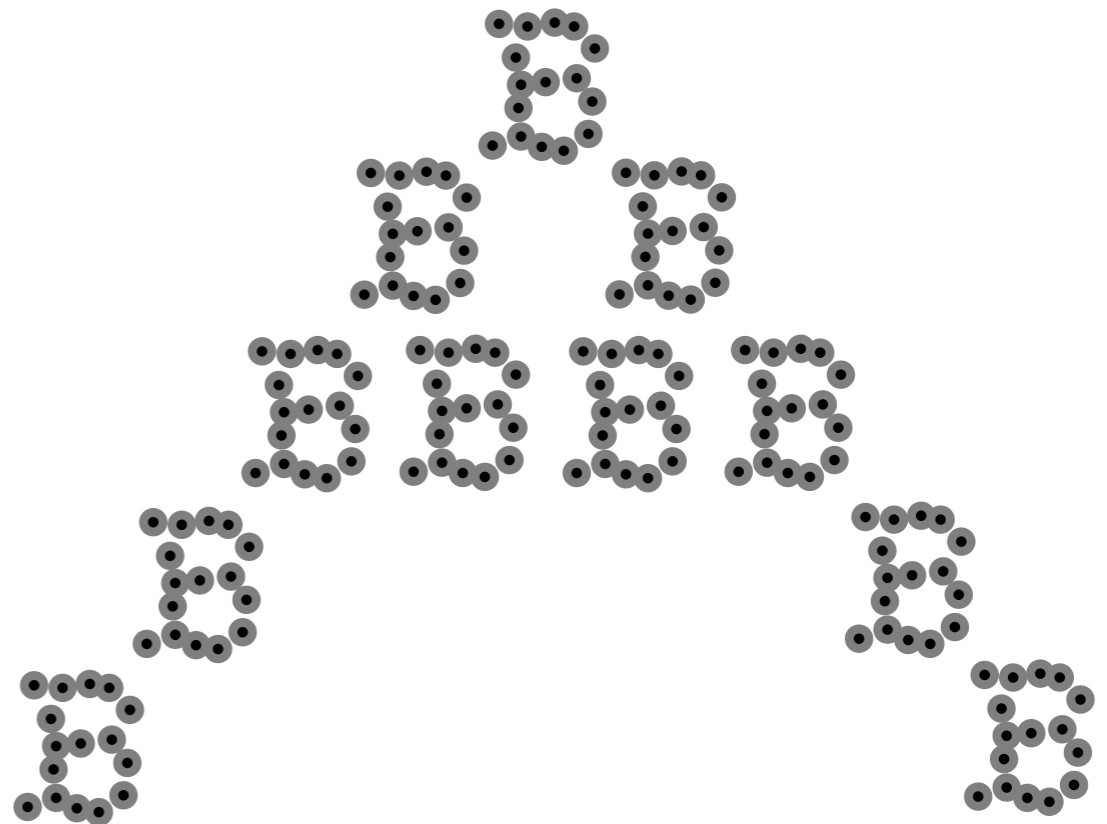
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

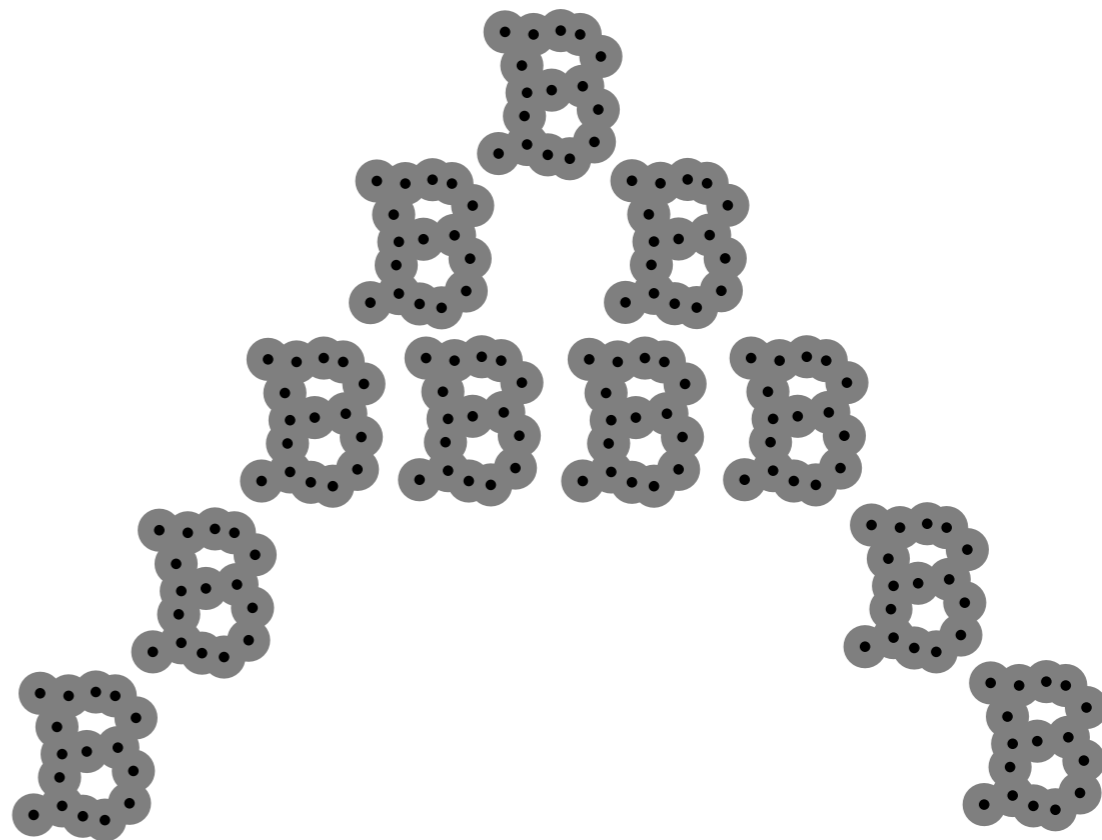
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

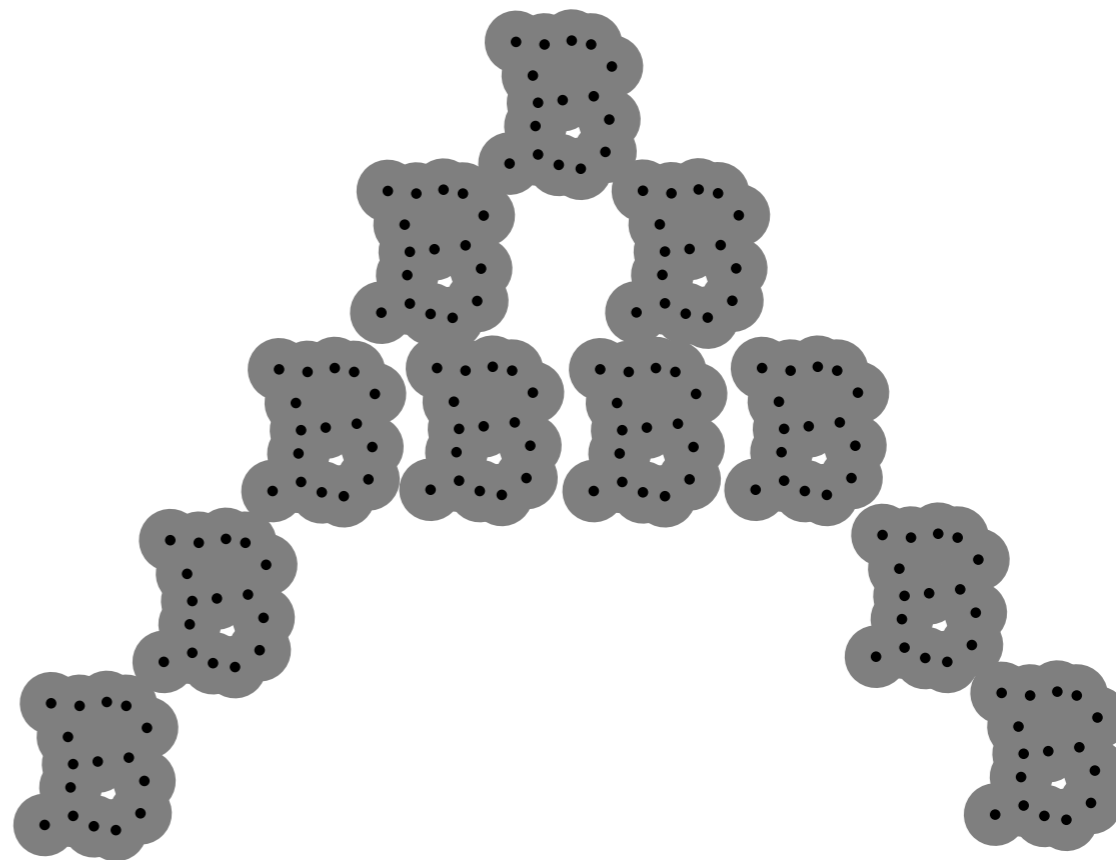
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

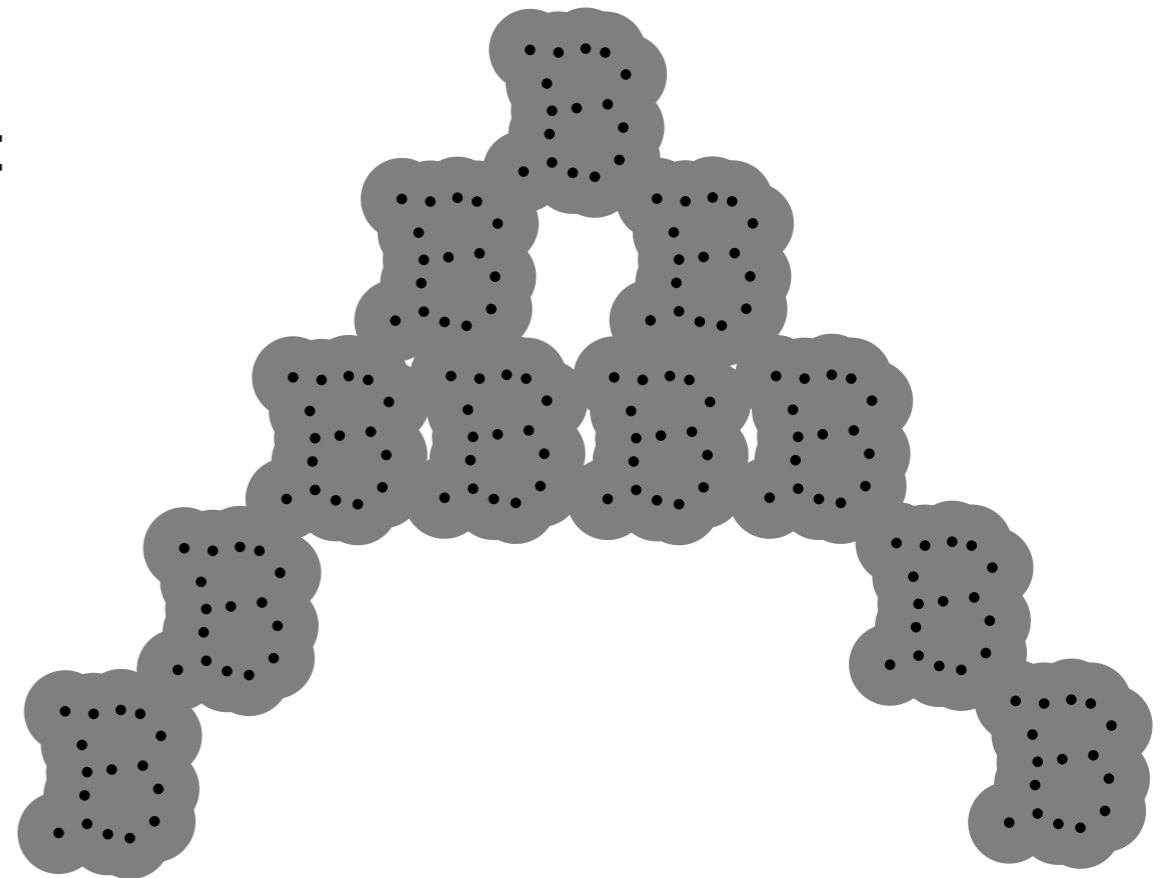
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

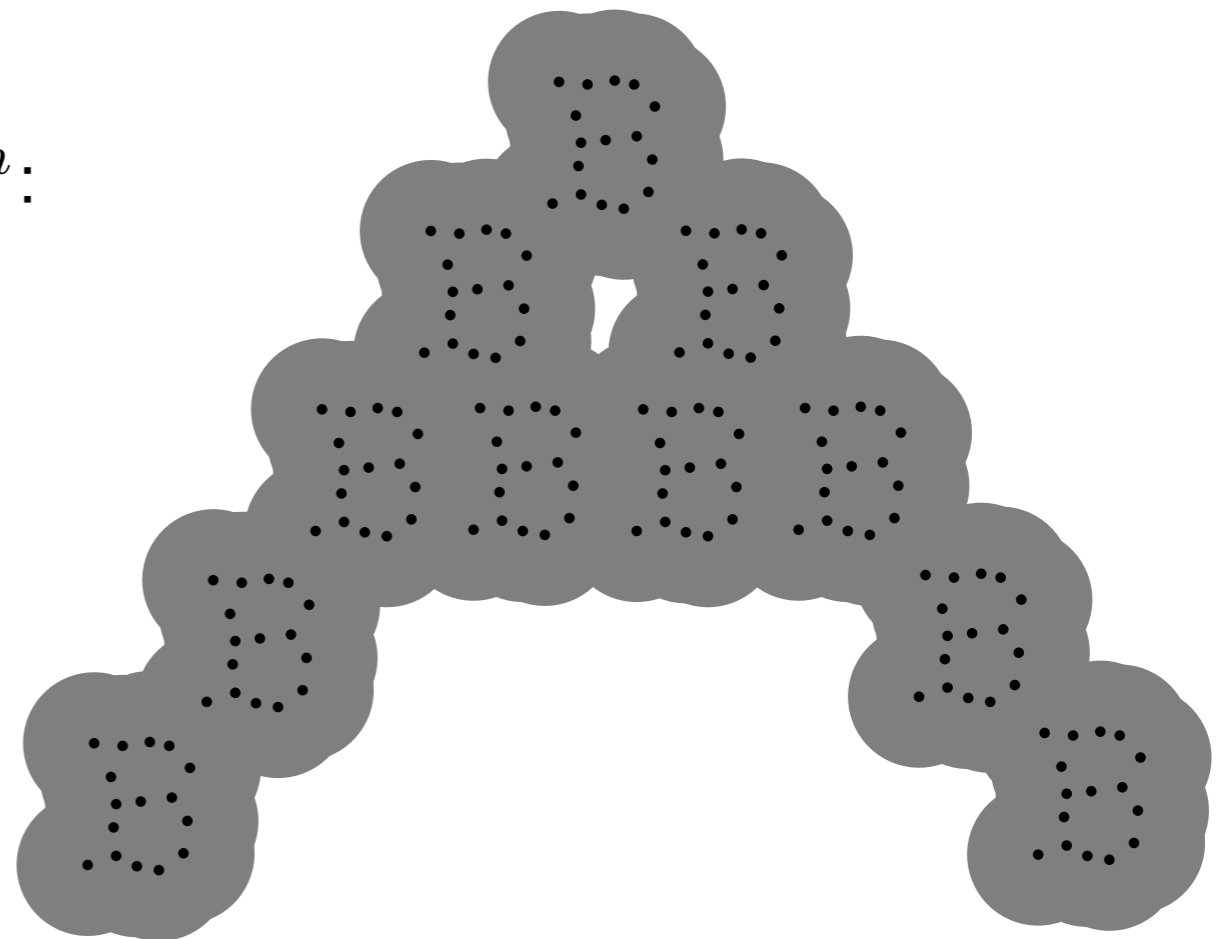
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

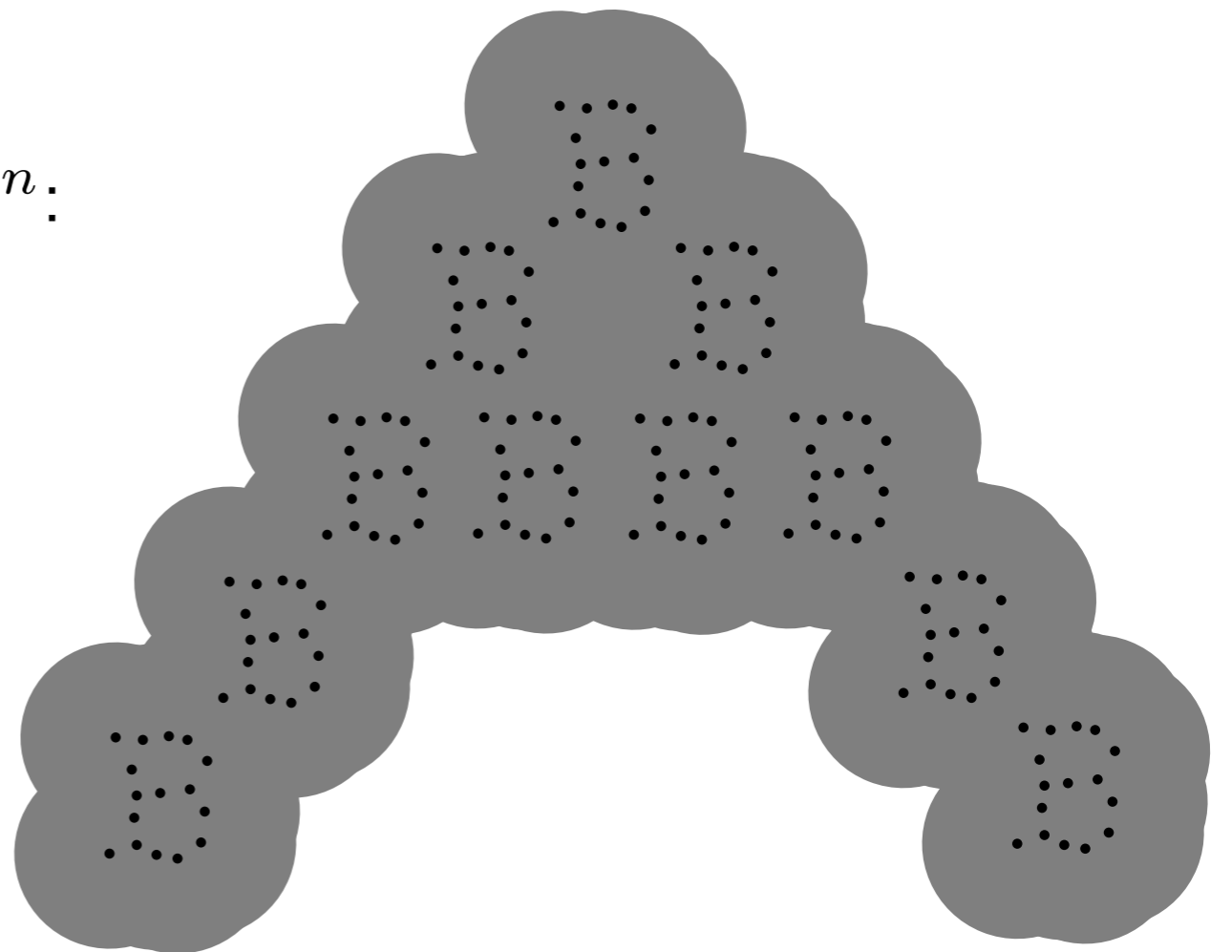
Filtration: functor $(P, \leq) \rightarrow \text{Top}$

- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$

Example: offsets filtration of $X \subseteq \mathbb{R}^n$:

$$f : \begin{cases} \mathbb{R}^n \rightarrow P = \mathbb{R} \\ y \mapsto \min_{x \in X} \|y - x\|_2 \end{cases}$$

$$\begin{aligned} F(t) &:= f^{-1}((-\infty, t]) \\ &= \bigcup_{x \in X} B(x, t) \end{aligned}$$



Filtrations

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$)

Filtration: functor $(P, \leq) \rightarrow \text{Top}$

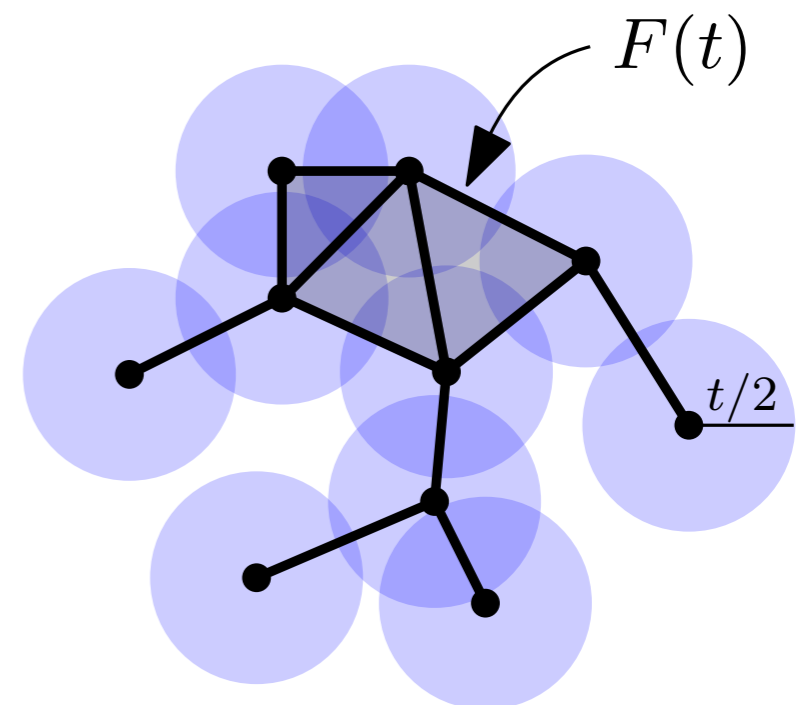
- ▶ typically, $F(t) := f^{-1}((-\infty, t])$ for some P -valued function f
- ▶ $F(t) \subseteq F(u)$ for all $t \leq u \in P$
- ▶ for computational purposes, take $F: P \rightarrow \text{Simp}$

Example: Vietoris-Rips filtration of $X \subseteq \mathbb{R}^n$:

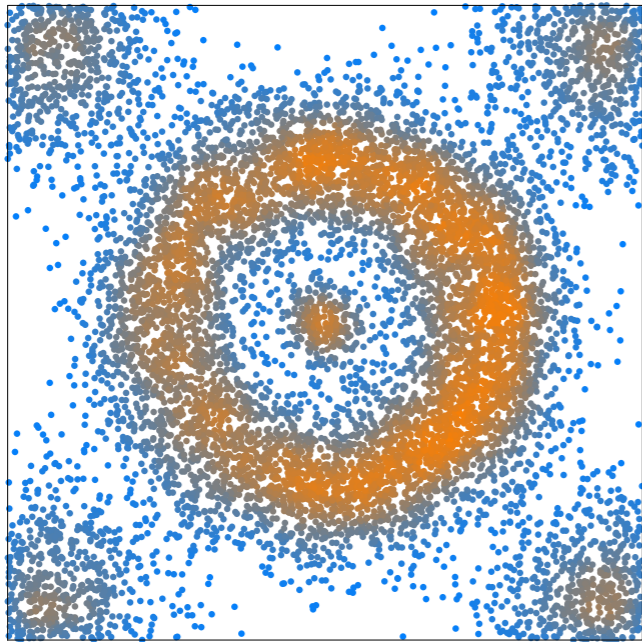
$$f : \begin{cases} 2^X \rightarrow P = \mathbb{R} \\ \{x_0, \dots, x_k\} \mapsto \max_{0 \leq i < j \leq k} \|x_i - x_j\|_2 \end{cases}$$

$F(t) =$ flag complex of intersection graph

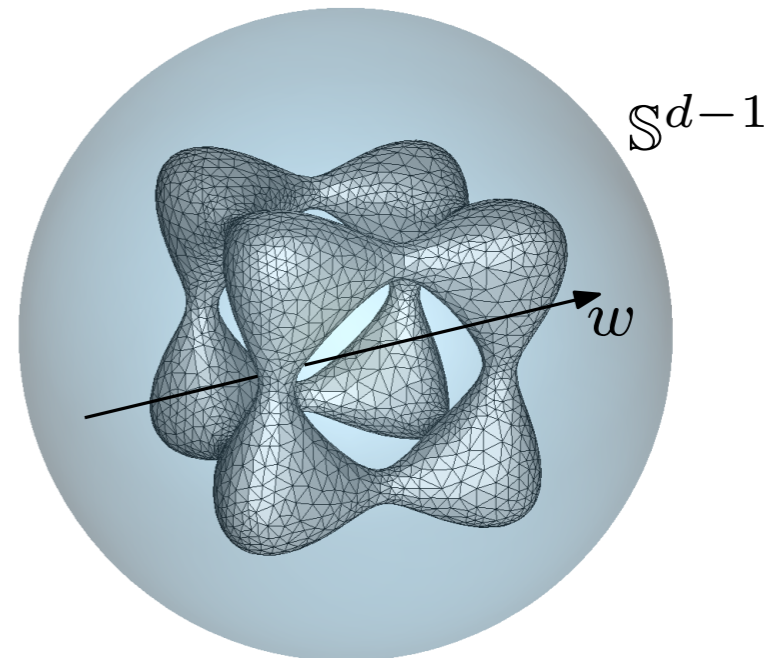
$$\text{of } \bigcup_{x \in X} B(x, t/2)$$



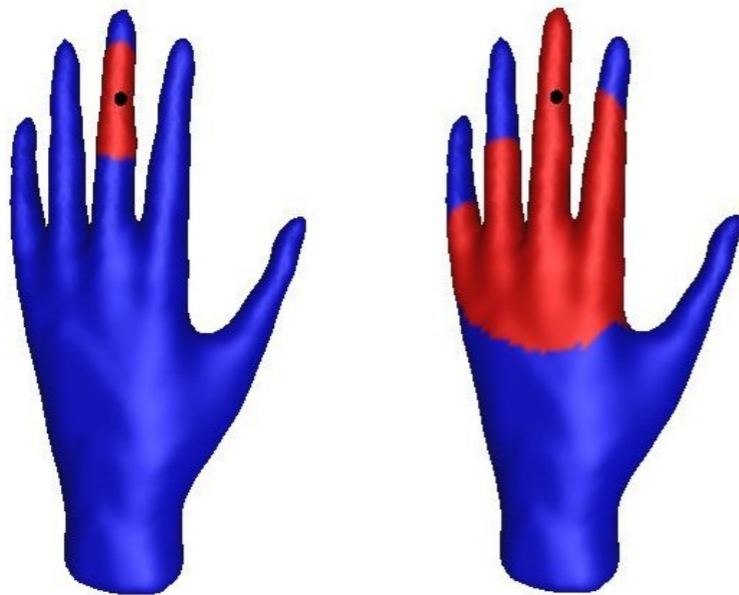
Other examples (any combination of the following)



density estimators



projections



single-source distances

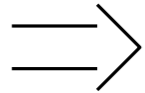
others:

- non-linear projections
- curvature measures
- PDE solutions (heat, wave)
- etc.

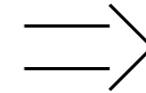
Topological Data Analysis pipeline (again)



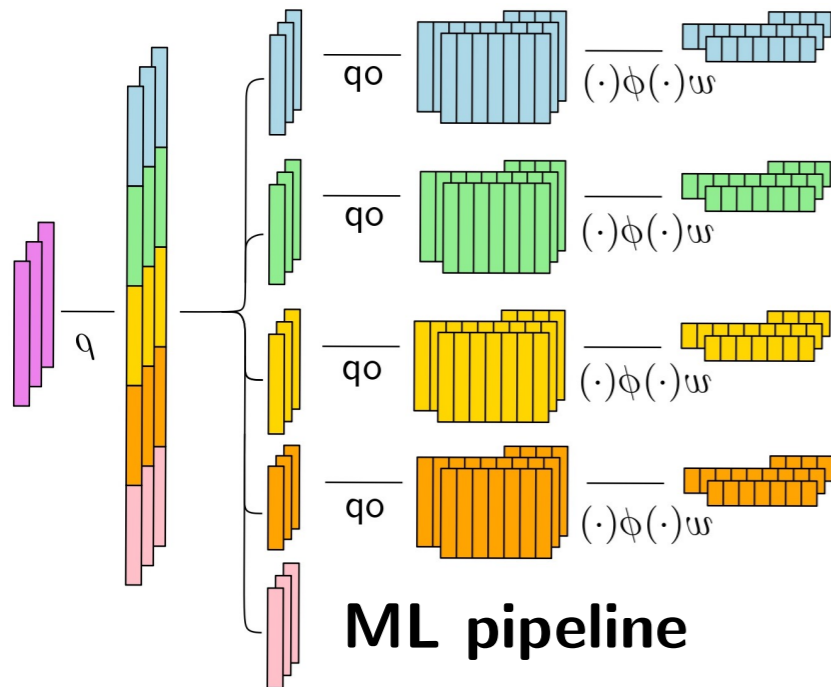
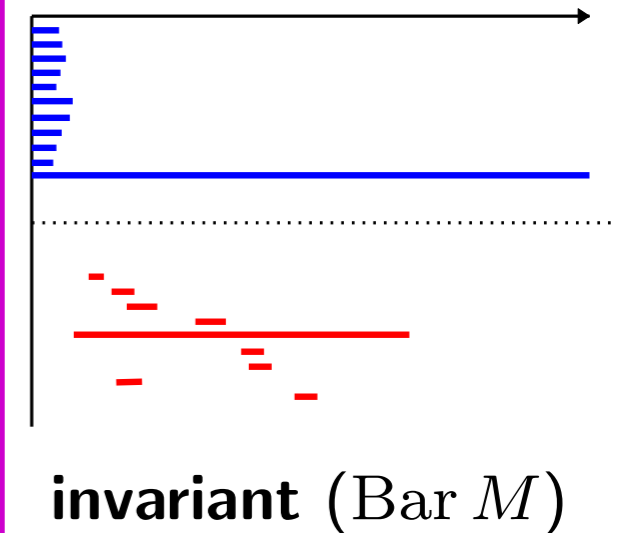
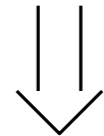
data



filtration $F : P \subseteq \mathbb{R}^d \rightarrow \text{Top}$
(sublevel sets of distance to data)

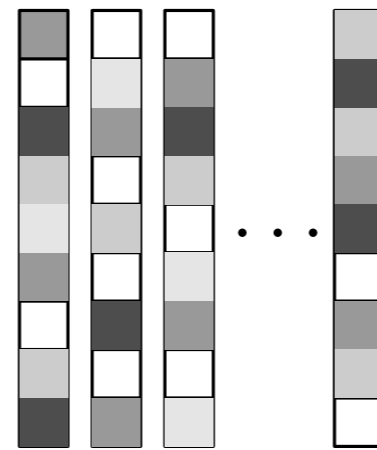
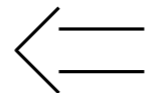


persistence module
 $M = HF : P \rightarrow \text{vect}_k$
(singular homology)

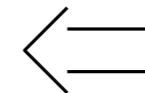


ML pipeline

(e.g. neural net.)



features (vectors)



Persistence modules

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$), k a field

Persistence module: functor $(P, \leq) \rightarrow \text{vect}_k$ (pointwise finite-dimensional, or pfd)

Persistence modules

(P, \leq) a poset (usually $(\mathbb{Z}^d, \leq_{\Pi})$ or $(\mathbb{R}^d, \leq_{\Pi})$), \mathbf{k} a field

Persistence module: functor $(P, \leq) \rightarrow \text{vect}_{\mathbf{k}}$ (pointwise finite-dimensional, or pfd)

Interval: $I \subseteq P$ that is:

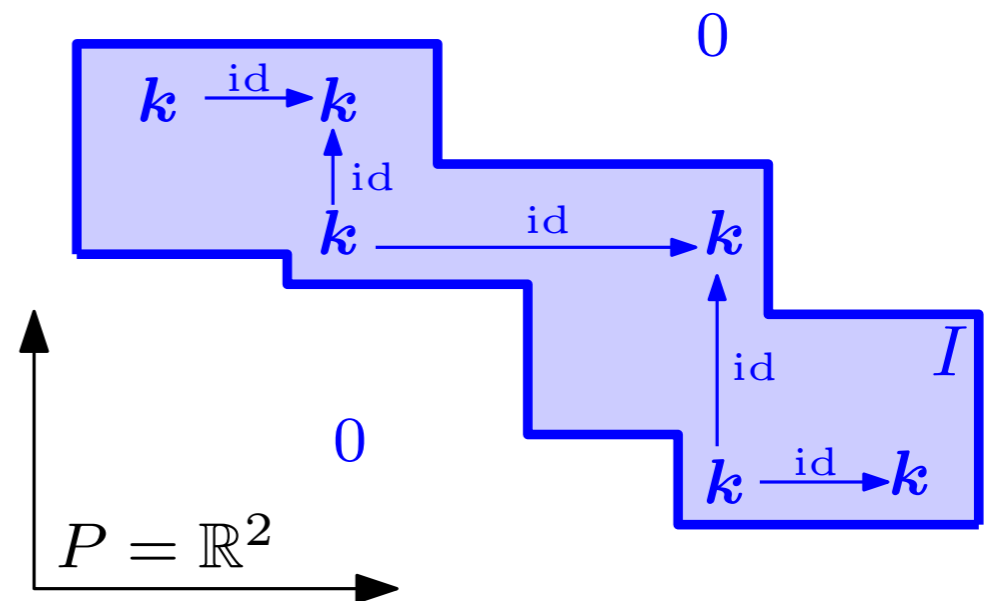
- convex ($s, t \in I \implies u \in I \forall s \leq u \leq t$)
- connected ($s, t \in I \implies \exists \{u_i\}_{i=0}^r \subseteq I$ s.t. $s = u_0 \leq u_1 \geq \dots \geq u_r = t$)

Interval module: indicator module \mathbf{k}_I of an interval $I \subseteq P$

$$\mathbf{k}_I(t) = \begin{cases} \mathbf{k} & \text{if } t \in I \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{k}_I(s \leq t) = \begin{cases} \text{id}_{\mathbf{k}} & \text{if } s, t \in I \\ 0 & \text{otherwise} \end{cases}$$

note: $\text{End}(\mathbf{k}_I) \simeq \mathbf{k}$



Persistence modules

Interval modules are **described by their support**:

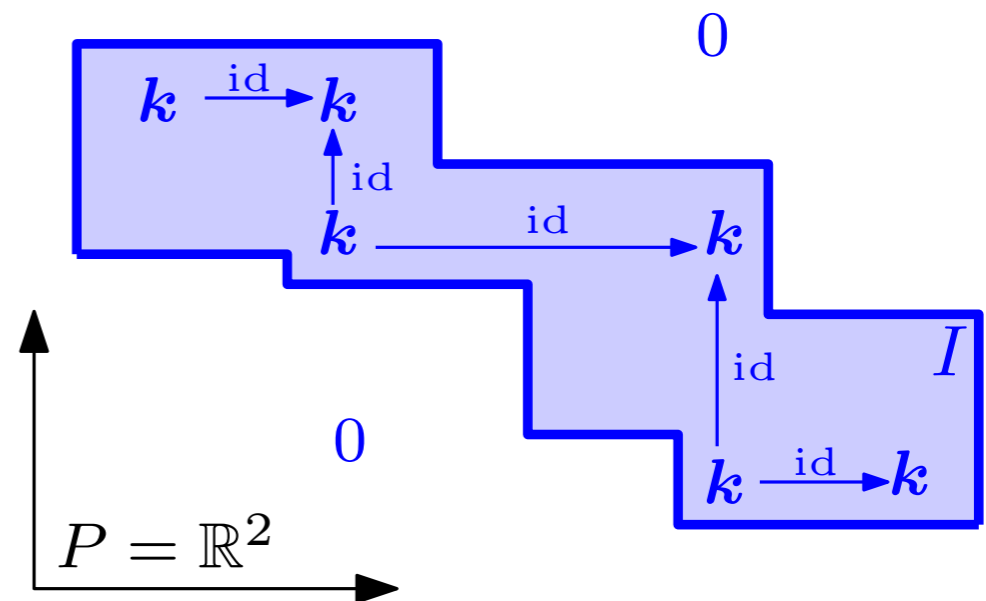
- ▶ complete geometric descriptor
- ▶ efficient to encode (small / simple dictionary)
- ▶ readily interpretable (for data exploration)
- ▶ easy to vectorize (for Machine Learning)
- ▶ enjoy stability properties (for statistics)

Interval module: indicator module \mathbf{k}_I of an interval $I \subseteq P$

$$\mathbf{k}_I(t) = \begin{cases} \mathbf{k} & \text{if } t \in I \\ 0 & \text{otherwise} \end{cases}$$

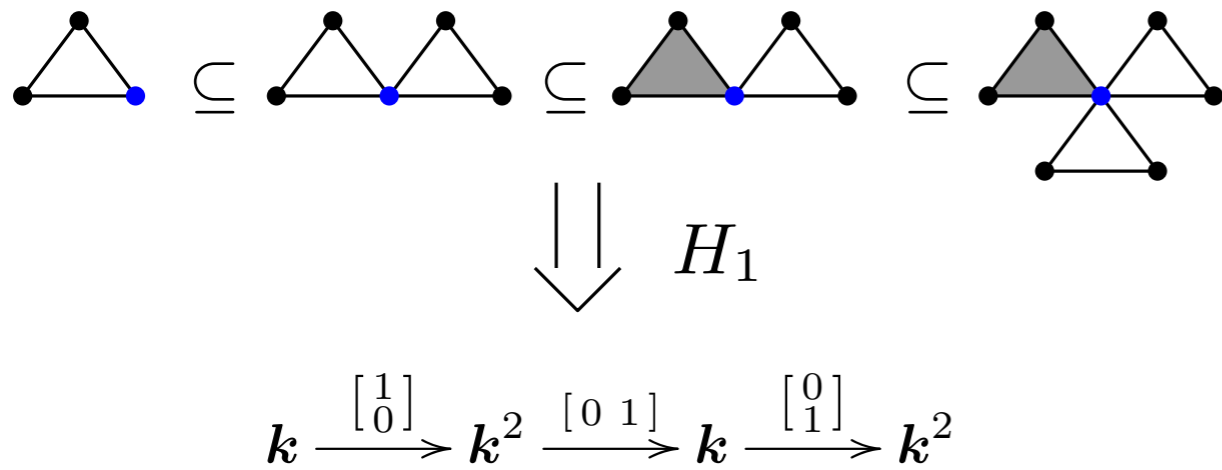
$$\mathbf{k}_I(s \leq t) = \begin{cases} \text{id}_{\mathbf{k}} & \text{if } s, t \in I \\ 0 & \text{otherwise} \end{cases}$$

note: $\text{End}(\mathbf{k}_I) \simeq \mathbf{k}$

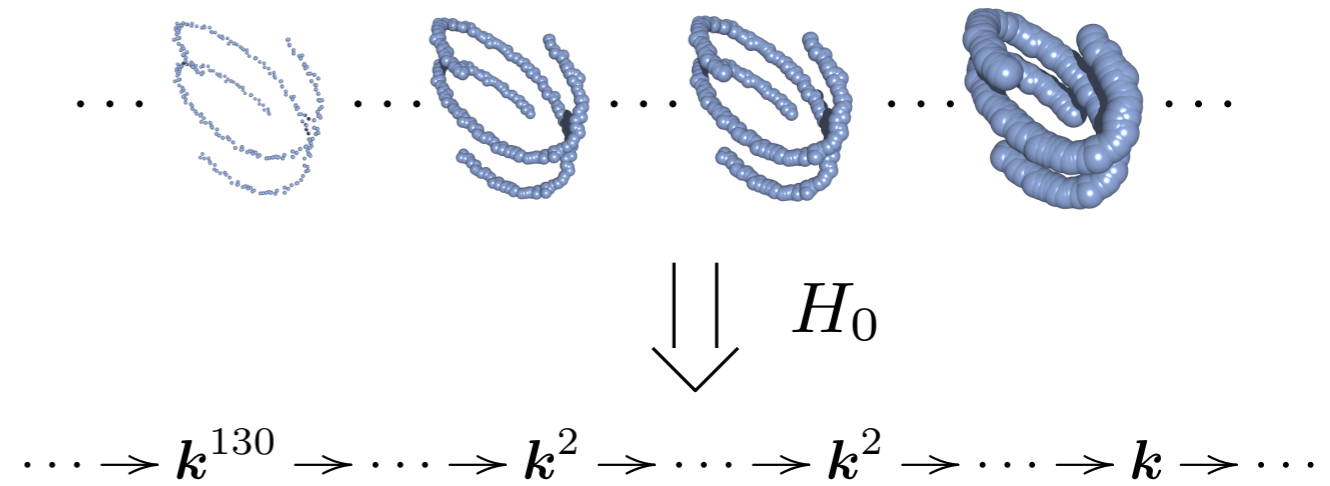


1-parameter persistence modules

discrete setting: $M : \llbracket 1, n \rrbracket \rightarrow \text{vect}_k$



continuous setting: $M : \mathbb{R} \rightarrow \text{vect}_k$

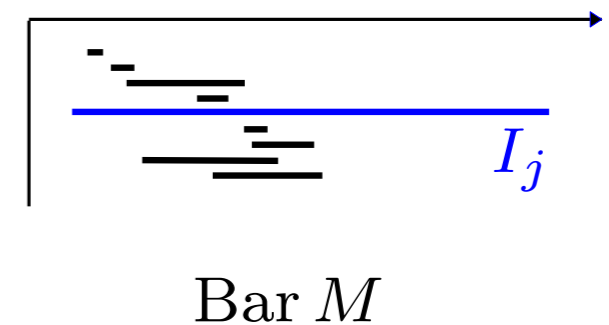


Thm [Gabriel][Auslander][Ringel][Webb][Crawley-Boevey]

For any $P \subseteq \mathbb{R}$ and any $M : (P, \leq) \rightarrow \text{vect}_k$:

$$M \simeq \bigoplus_{j \in J} k_{I_j}$$

where each $\text{End}(k_{I_j})$ is local

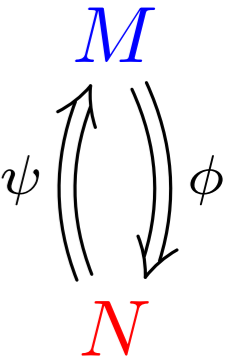


Metric viewpoint: interleaving distance

Given $M, N : (\mathbb{R}, \leq) \rightarrow \text{vect}_k$,

• **morphism:** natural transformation $M \Rightarrow N$

• **isomorphism:** pair of morphisms $M \xRightarrow{\phi} N$ and $N \xRightarrow{\psi} M$ such that:

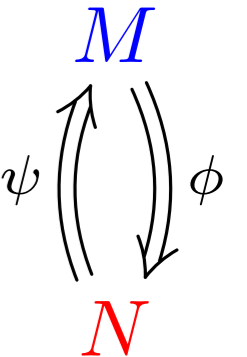


Metric viewpoint: interleaving distance

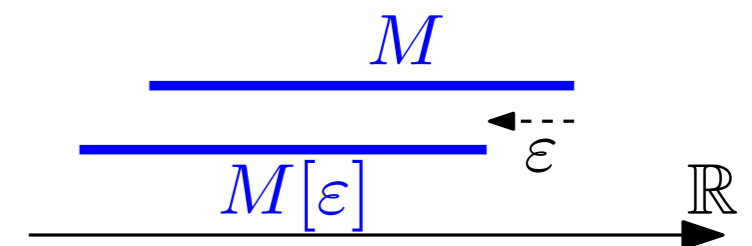
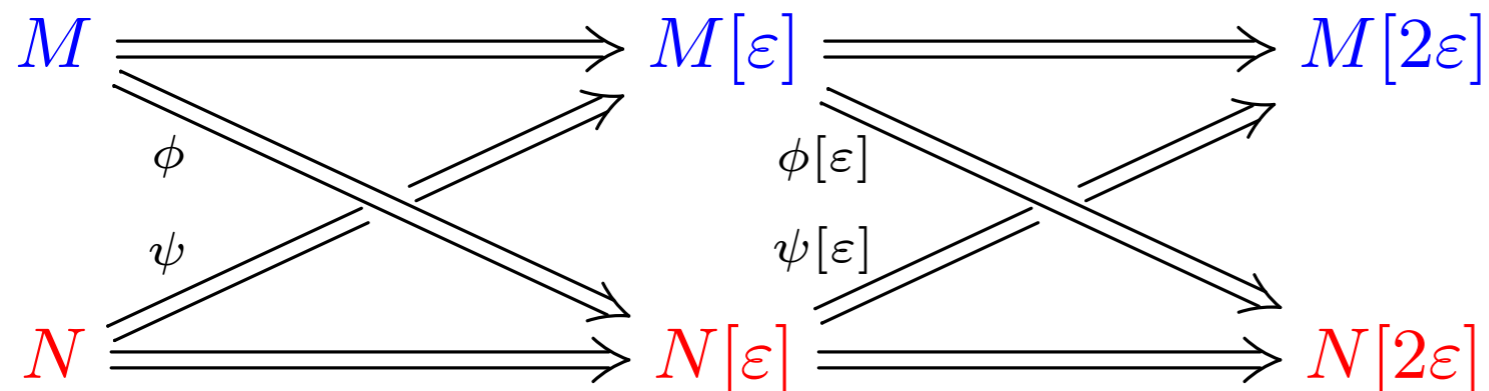
Given $M, N : (\mathbb{R}, \leq) \rightarrow \text{vect}_k$,

- **morphism:** natural transformation $M \Rightarrow N$

- **isomorphism:** pair of morphisms $M \xrightarrow{\phi} N$ and $N \xrightarrow{\psi} M$ such that:



- ε -**isomorphism:** pair of morphisms $M \xrightarrow{\phi} N[\varepsilon]$ and $N \xrightarrow{\psi} M[\varepsilon]$ such that:



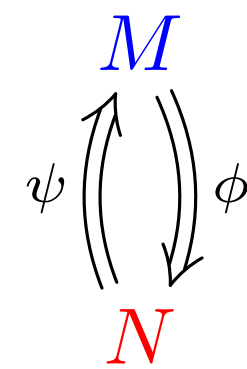
where $M[\varepsilon](t) := M(t + \varepsilon)$ and $\phi[\varepsilon](t) := \phi(t + \varepsilon)$

- **interleaving distance:** $d_i(M, N) := \inf\{\varepsilon \mid M, N \text{ } \varepsilon\text{-isomorphic}\}$

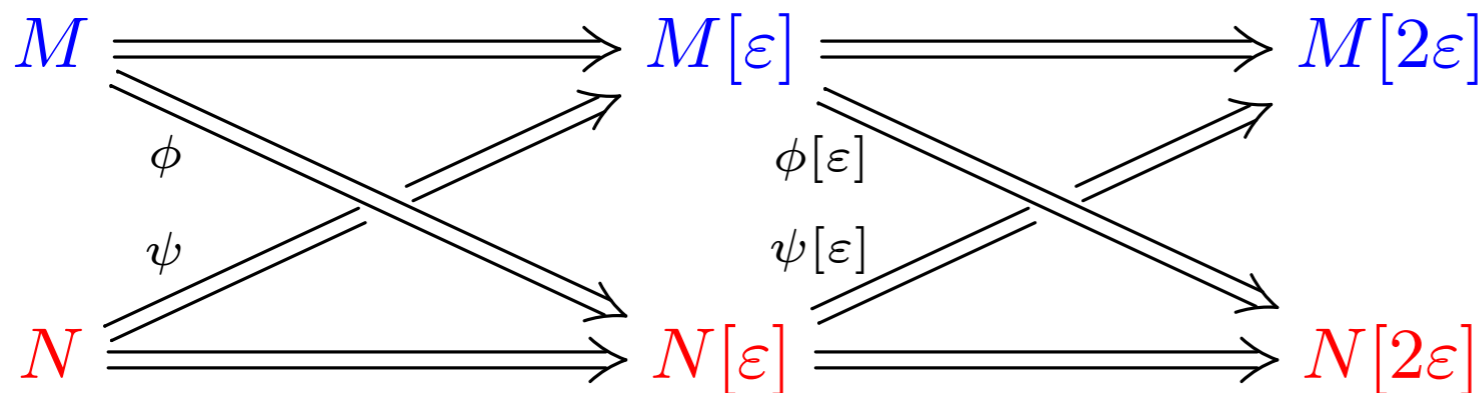
Metric viewpoint: interleaving distance

Given $M, N : (\mathbb{R}, \leq) \rightarrow \text{vect}_k$,

- **morphism:** natural transformation $M \Rightarrow N$

- **isomorphism:** pair of morphisms $M \xrightarrow{\phi} N$ and $N \xrightarrow{\psi} M$ such that: 

- ε -**isomorphism:** pair of morphisms $M \xrightarrow{\phi} N[\varepsilon]$ and $N \xrightarrow{\psi} M[\varepsilon]$ such that:



Prop: $\forall f, g : X \rightarrow \mathbb{R}$,
 $d_i(HF, HG) \leq \|f - g\|_\infty$

Thm: [Lesnick]
 d stable as above $\Rightarrow d \leq d_i$

where $M[\varepsilon](t) := M(t + \varepsilon)$ and $\phi[\varepsilon](t) := \phi(t + \varepsilon)$

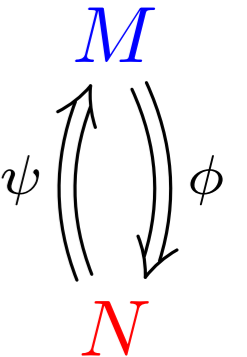
- **interleaving distance:** $d_i(M, N) := \inf\{\varepsilon \mid M, N \text{ } \varepsilon\text{-isomorphic}\}$

Metric viewpoint: bottleneck distance

Given $M = \bigoplus_{a \in A} M_a, N = \bigoplus_{b \in B} N_b : (\mathbb{R}, \leq) \rightarrow \text{vect}_k,$

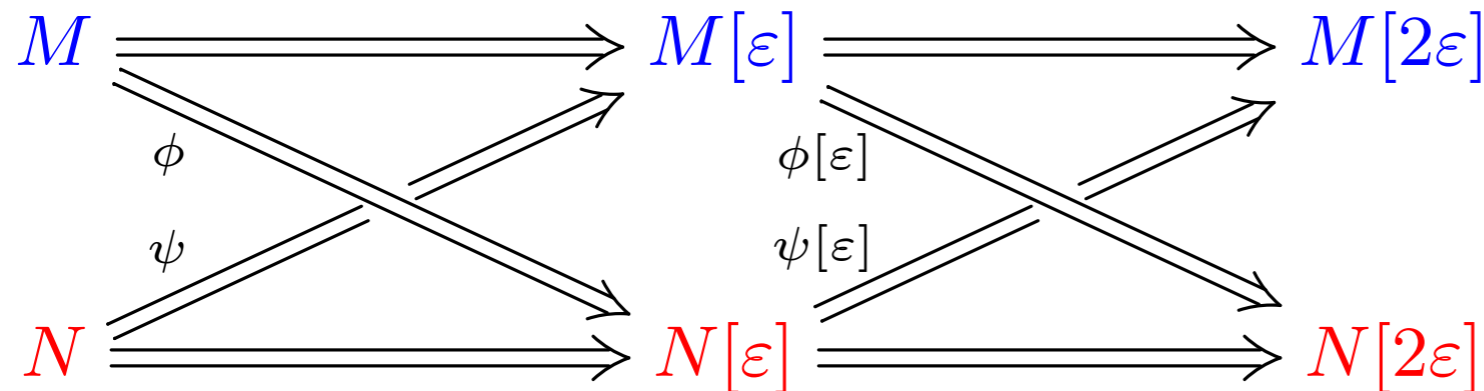
- **morphism:** natural transformation $M \Rightarrow N$

- **isomorphism:** pair of morphisms $M \xrightarrow{\phi} N$ and $N \xrightarrow{\psi} M$ such that:

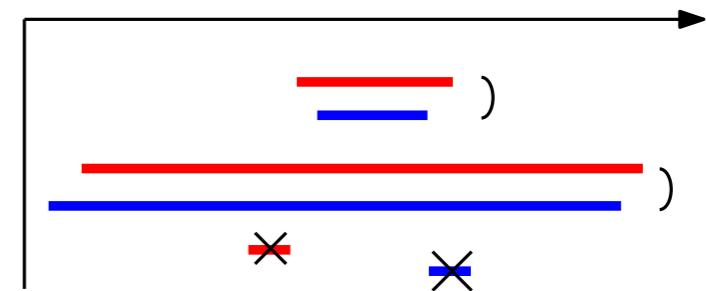


bottleneck

- **ε -isomorphism:** pair of morphisms $M \xrightarrow{\phi} N[\varepsilon]$ and $N \xrightarrow{\psi} M[\varepsilon]$ such that:



and ϕ, ψ factor through the decompositions of M and N



where $M[\varepsilon](t) := M(t + \varepsilon)$ and $\phi[\varepsilon](t) := \phi(t + \varepsilon)$

bottleneck

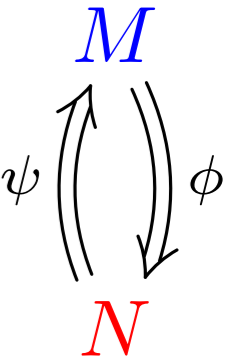
- **bottleneck distance:** $d_b(M, N) := \inf\{\varepsilon \mid M, N \text{ } \varepsilon\text{-isomorphic}\}$

Metric viewpoint: isometry

Given $M = \bigoplus_{a \in A} M_a, N = \bigoplus_{b \in B} N_b : (\mathbb{R}, \leq) \rightarrow \text{vect}_k,$

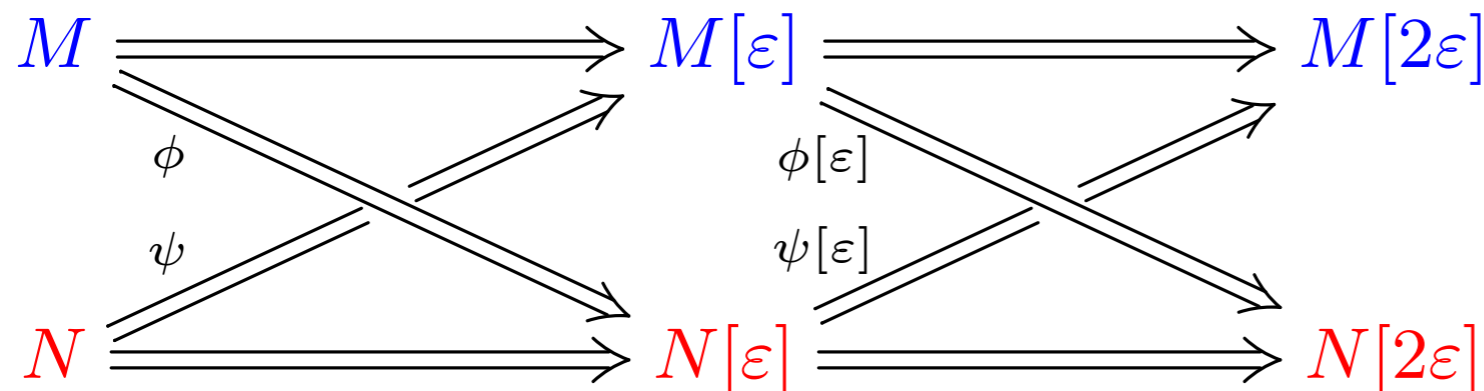
- **morphism:** natural transformation $M \Rightarrow N$

- **isomorphism:** pair of morphisms $M \xrightarrow{\phi} N$ and $N \xrightarrow{\psi} M$ such that:



bottleneck

- ε -**isomorphism:** pair of morphisms $M \xrightarrow{\phi} N[\varepsilon]$ and $N \xrightarrow{\psi} M[\varepsilon]$ such that:



and ϕ, ψ factor through the decompositions of M and N

Thm: $d_i = d_b$

where $M[\varepsilon](t) := M(t + \varepsilon)$ and $\phi[\varepsilon](t) := \phi(t + \varepsilon)$

bottleneck

- **bottleneck distance:** $d_b(M, N) := \inf\{\varepsilon \mid M, N \text{ } \varepsilon\text{-isomorphic}\}$

Wrap-up on 1-parameter persistence modules

Structure theorems

- ▶ complete classification of pfd persistence modules via their barcodes
- ▶ efficient algorithms for barcode computation

Isometry theorem

- ▶ barcodes as complete metric invariants for persistence modules
 - combinatorial algorithms for distance computation
- ▶ space of barcodes as a space of measures
 - bounds on intrinsic curvature
 - toolbox for statistics
 - vectorizations and kernels for ML

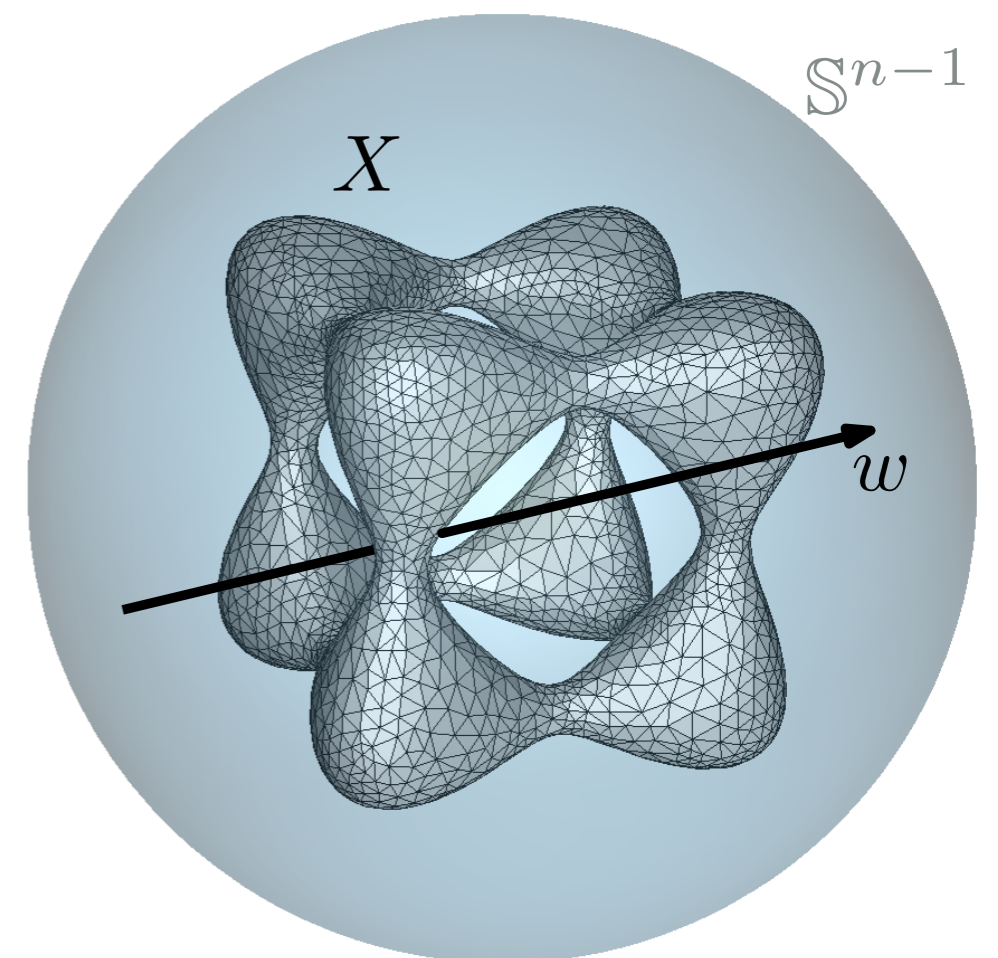
Multi-parameter persistence... what for?

- study joint variables (e.g. treatment efficacy vs. risk)
- increase feature sensitivity (by enhanced feature aggregation)

Thm: [Boyer, Curry, Mukherjee, Turner]
[Ghrist, Levanger, Mai]

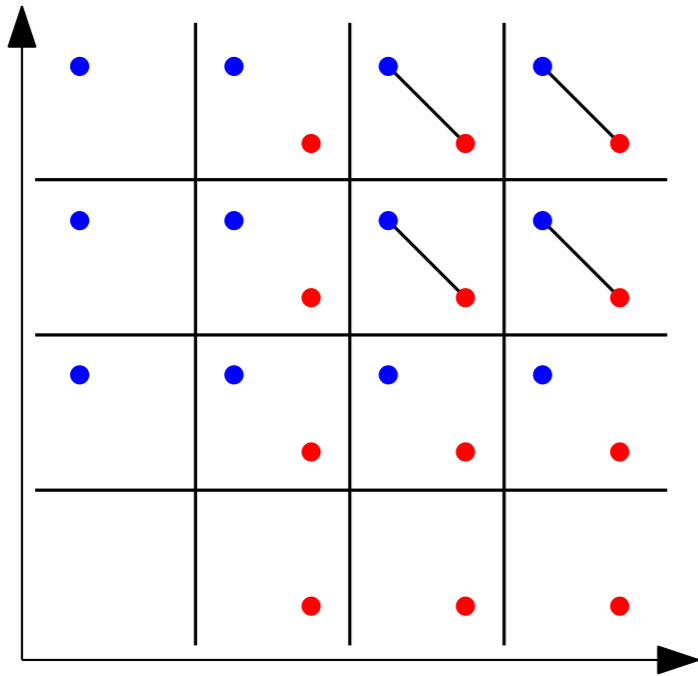
The map $X \mapsto \{\text{Bar} \langle \cdot, w \rangle|_X\}_{w \in \mathbb{S}^{n-1}}$ is injective on the class of compact subanalytic sets $X \subset \mathbb{R}^n$.

Q: can we reduce to finitely many directions using multi-parameter persistence?

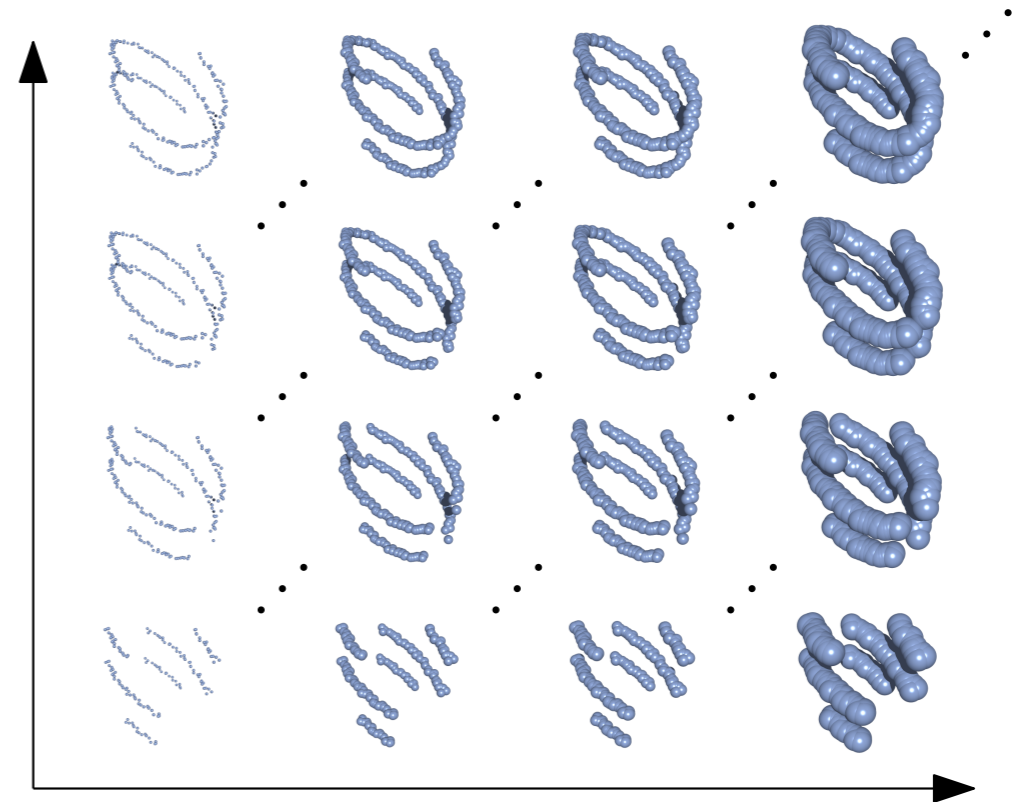


Multi-parameter persistence modules

discrete setting: $M : \llbracket 1, n \rrbracket^d \rightarrow \text{vect}_k$



continuous setting: $M : \mathbb{R}^d \rightarrow \text{vect}_k$



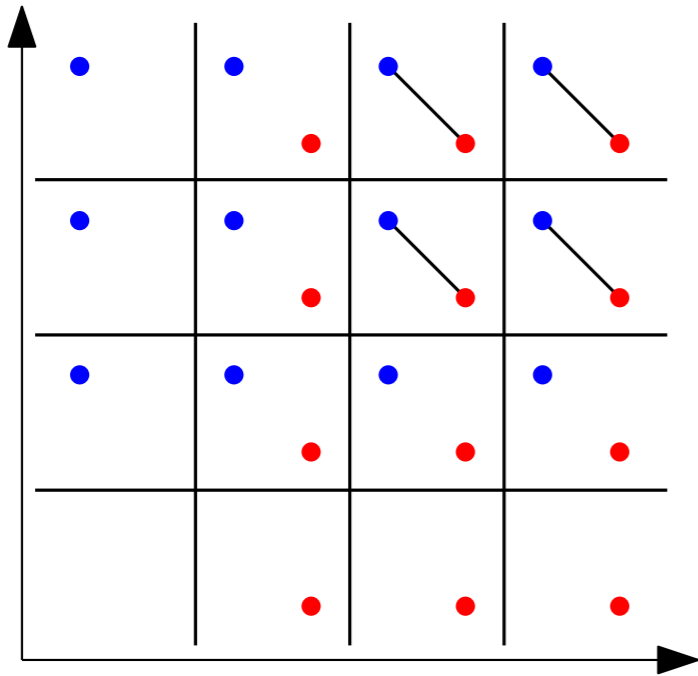
Thm [Botnan, Crawley-Boevey]

For any poset (P, \leq) and functor $M : (P, \leq) \rightarrow \text{vect}_k$:

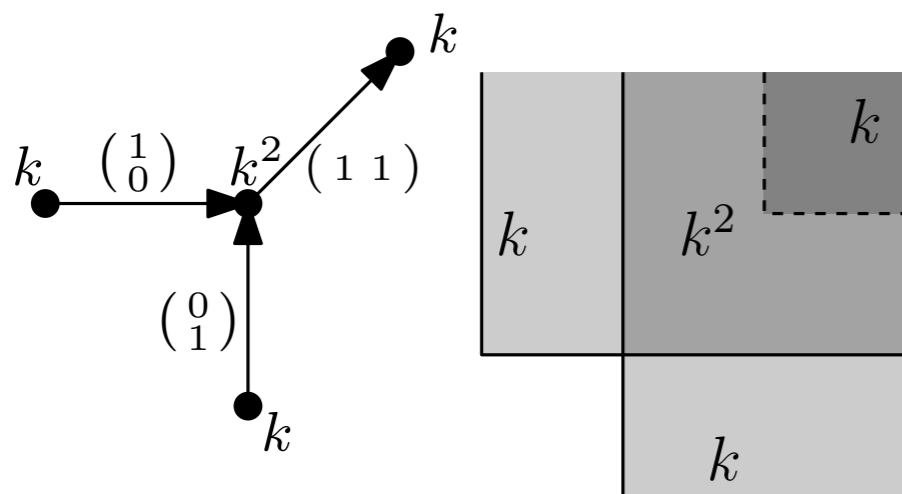
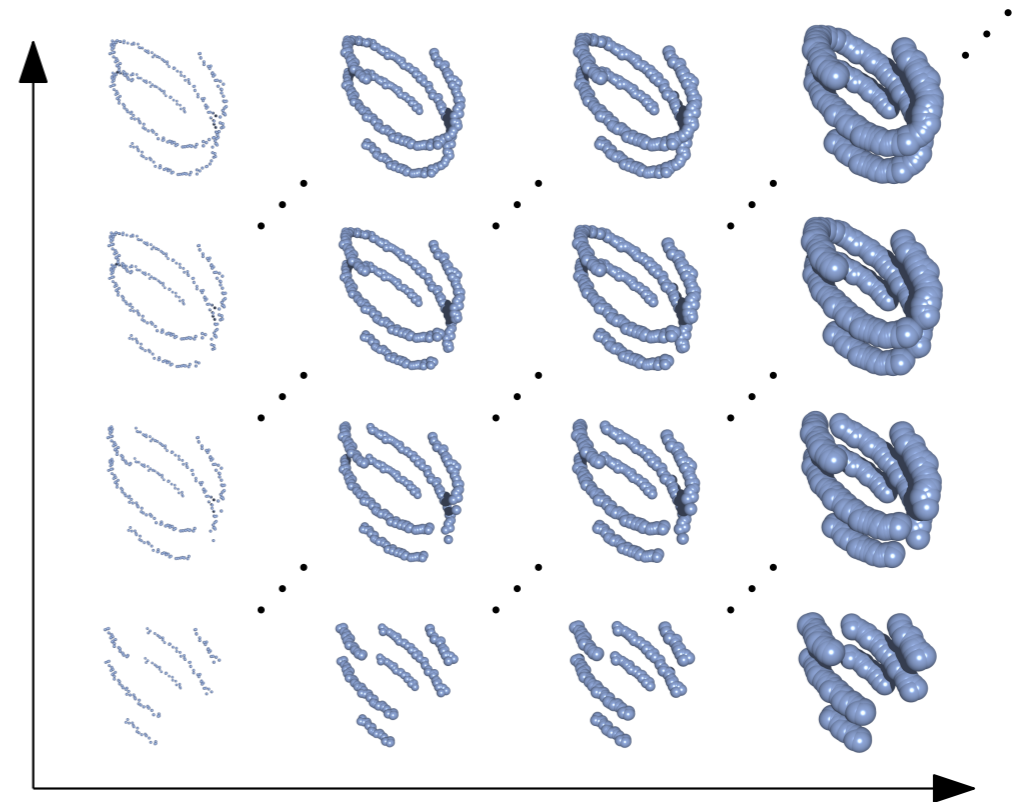
$$M \simeq \bigoplus_{j \in J} M_j \quad \text{where each } \text{End}(M_j) \text{ is local}$$

Multi-parameter persistence modules

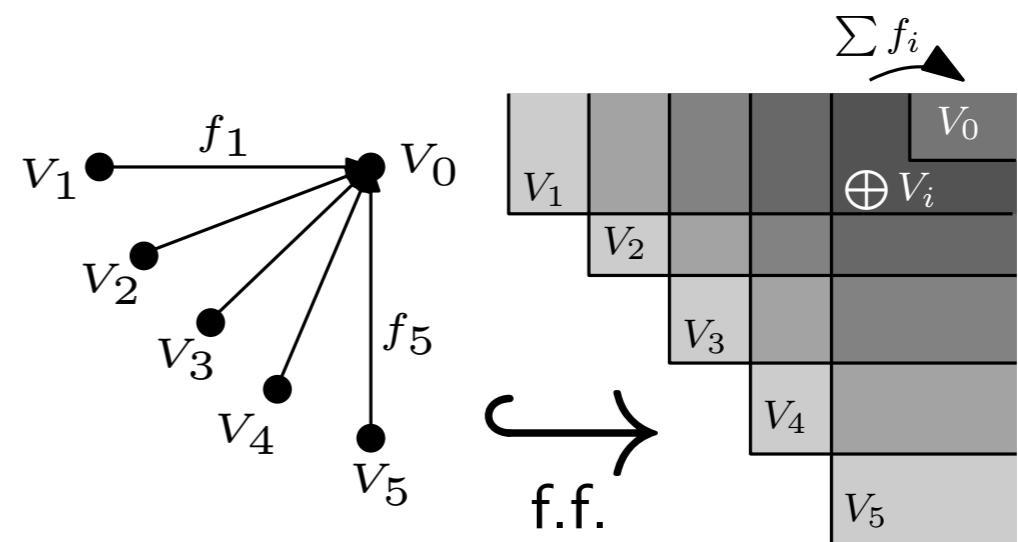
discrete setting: $M : \llbracket 1, n \rrbracket^d \rightarrow \text{vect}_k$



continuous setting: $M : \mathbb{R}^d \rightarrow \text{vect}_k$



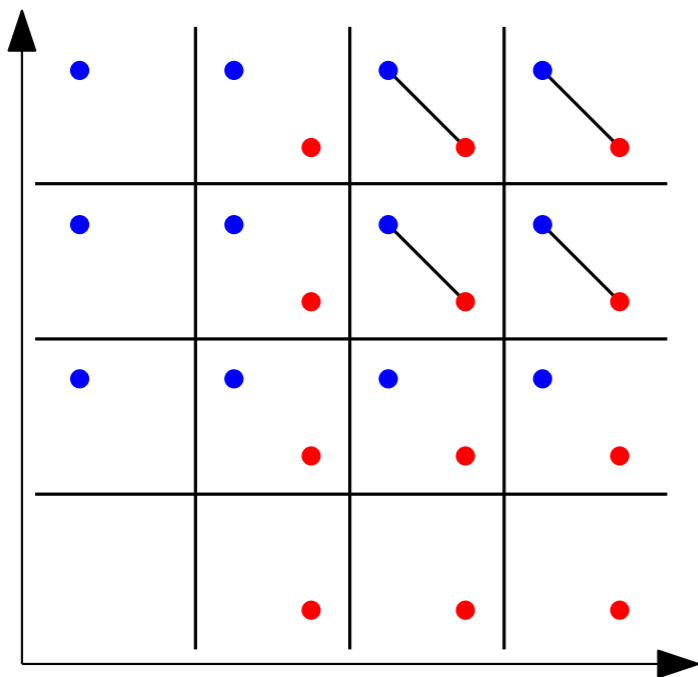
non-thin summands



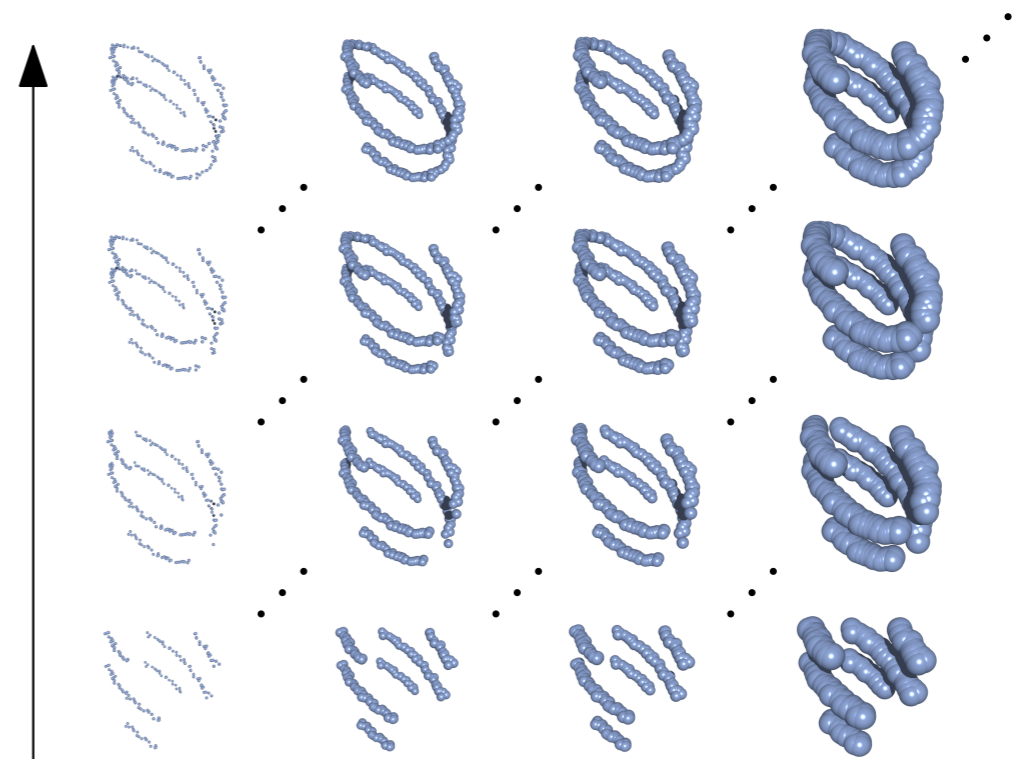
wild-type

Multi-parameter persistence modules

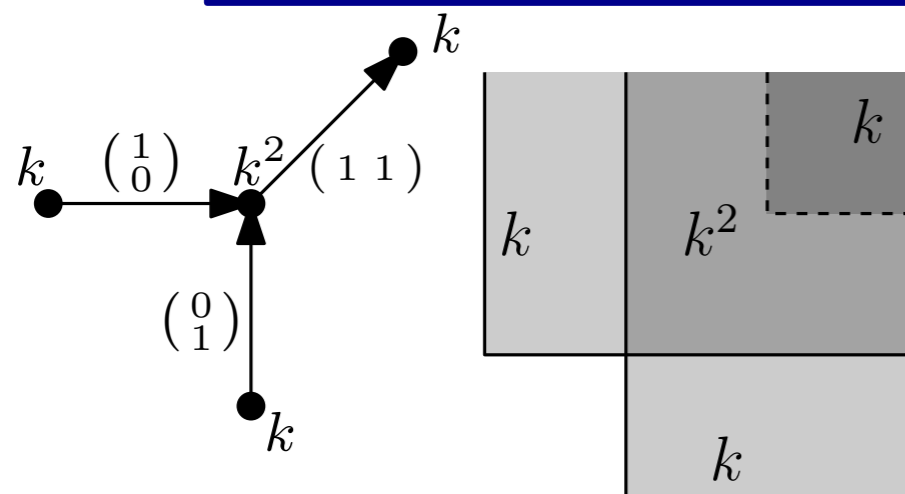
discrete setting: $M : \llbracket 1, n \rrbracket^d \rightarrow \text{vect}_k$



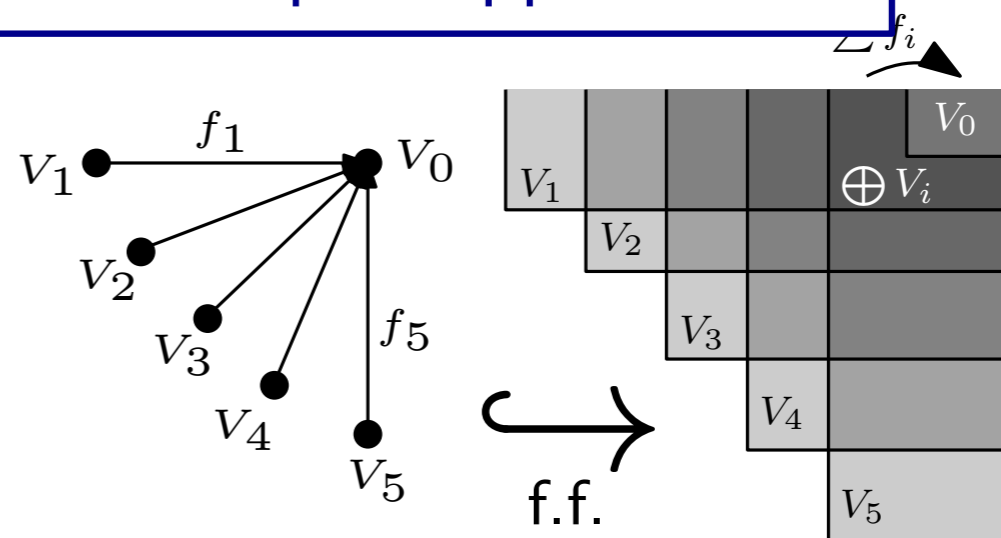
continuous setting: $M : \mathbb{R}^d \rightarrow \text{vect}_k$



Q: do non-thin indecomposables show up in applications?

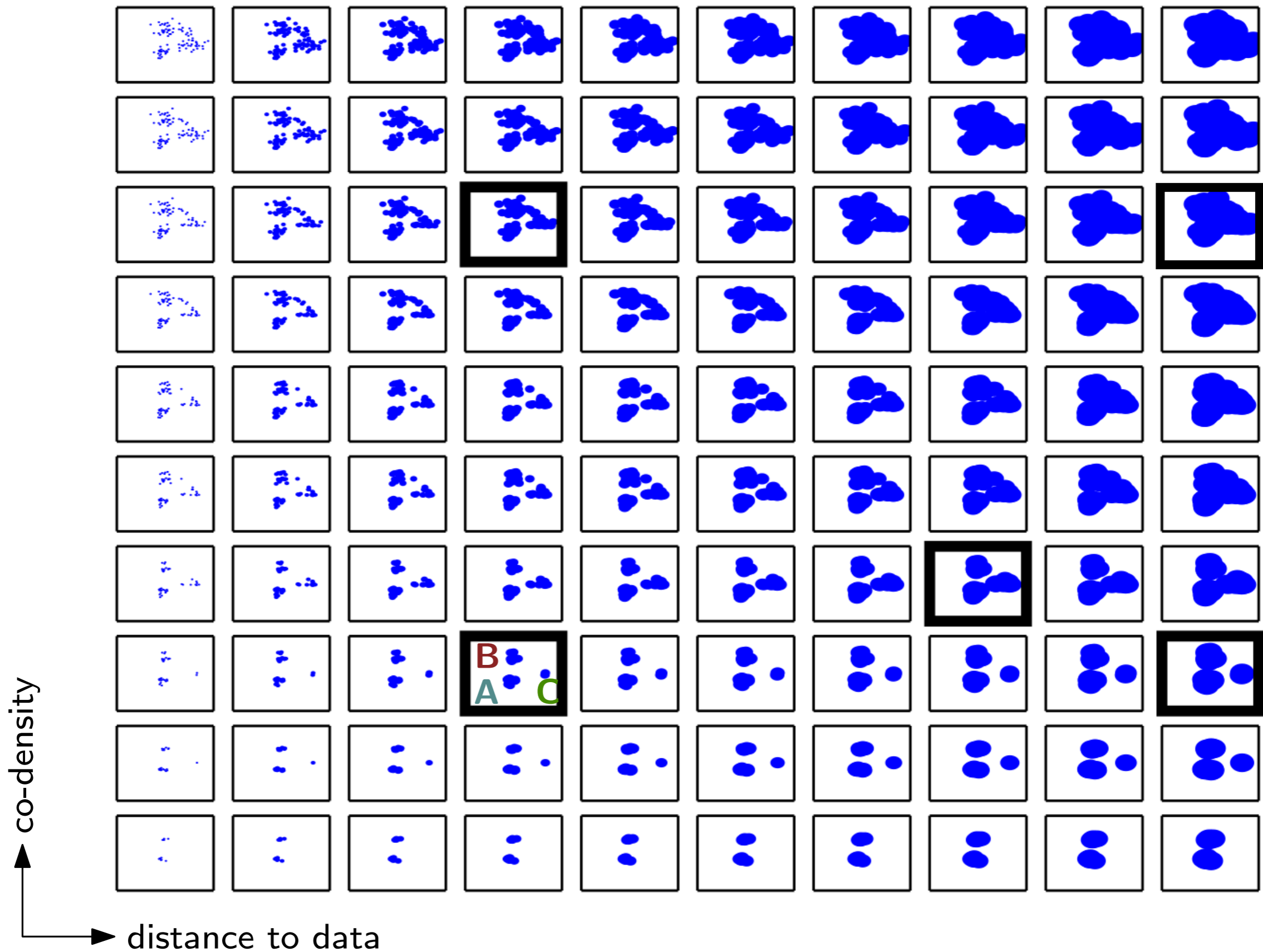


non-thin summands

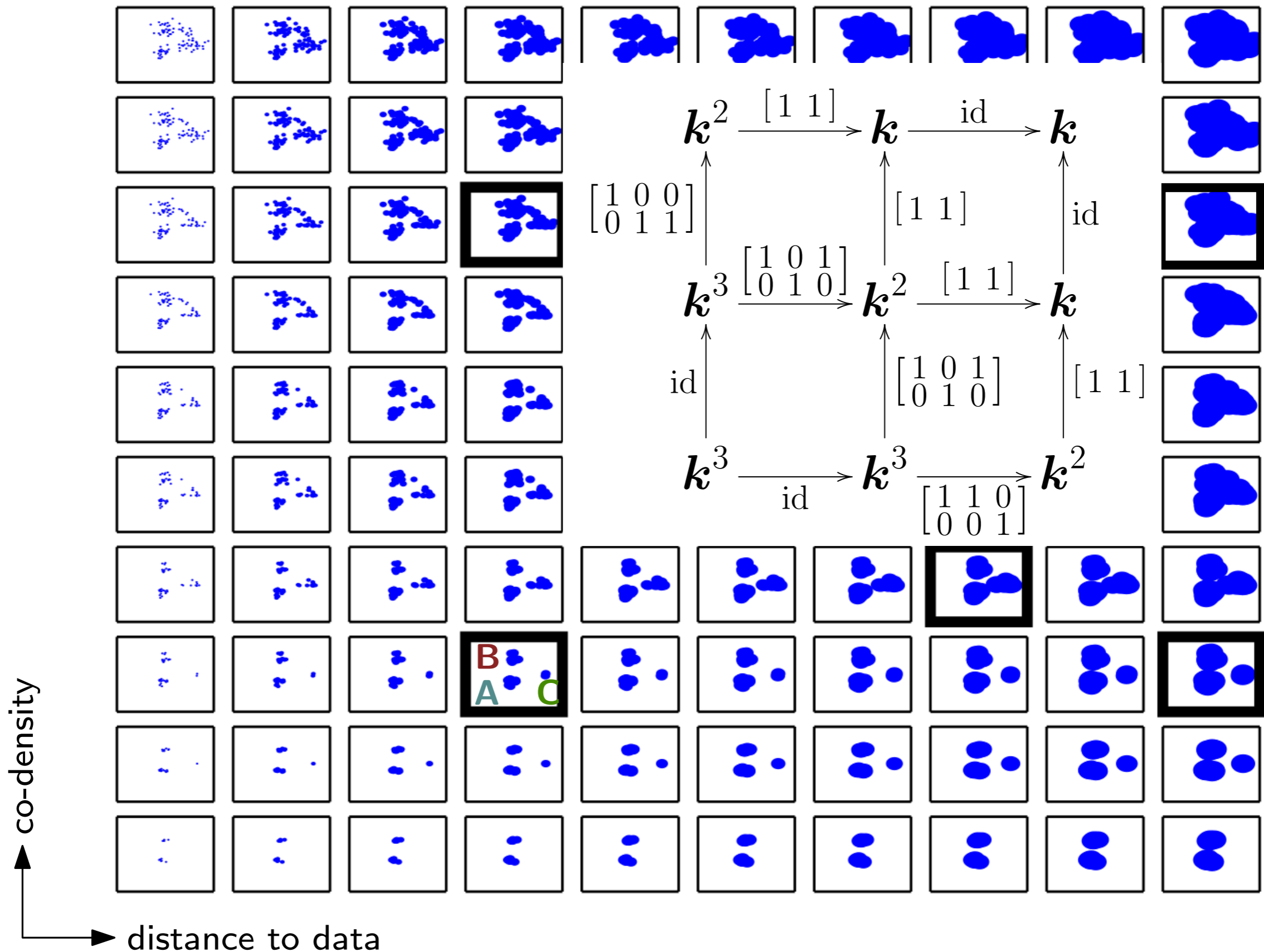


wild-type

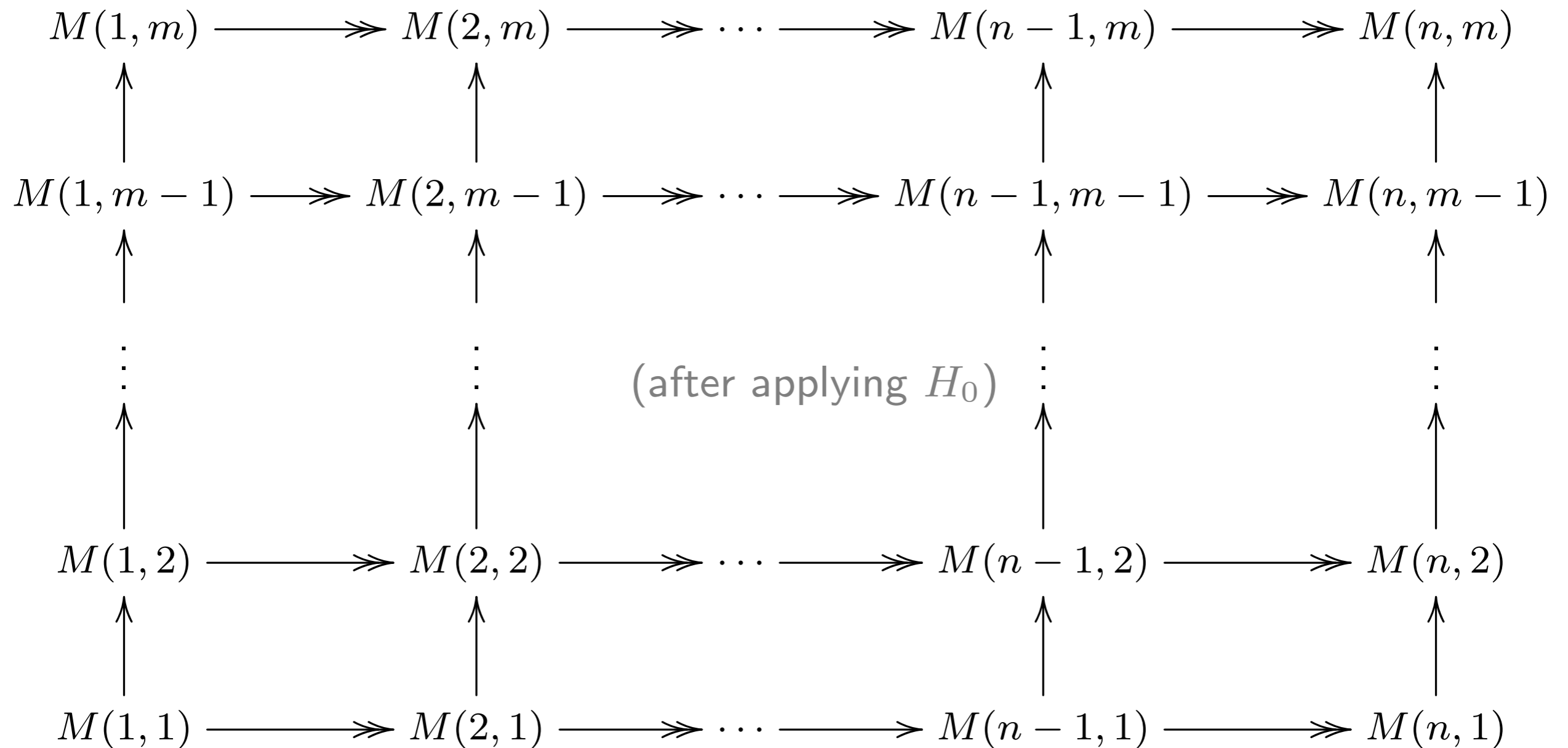
Example: two-parameter clustering



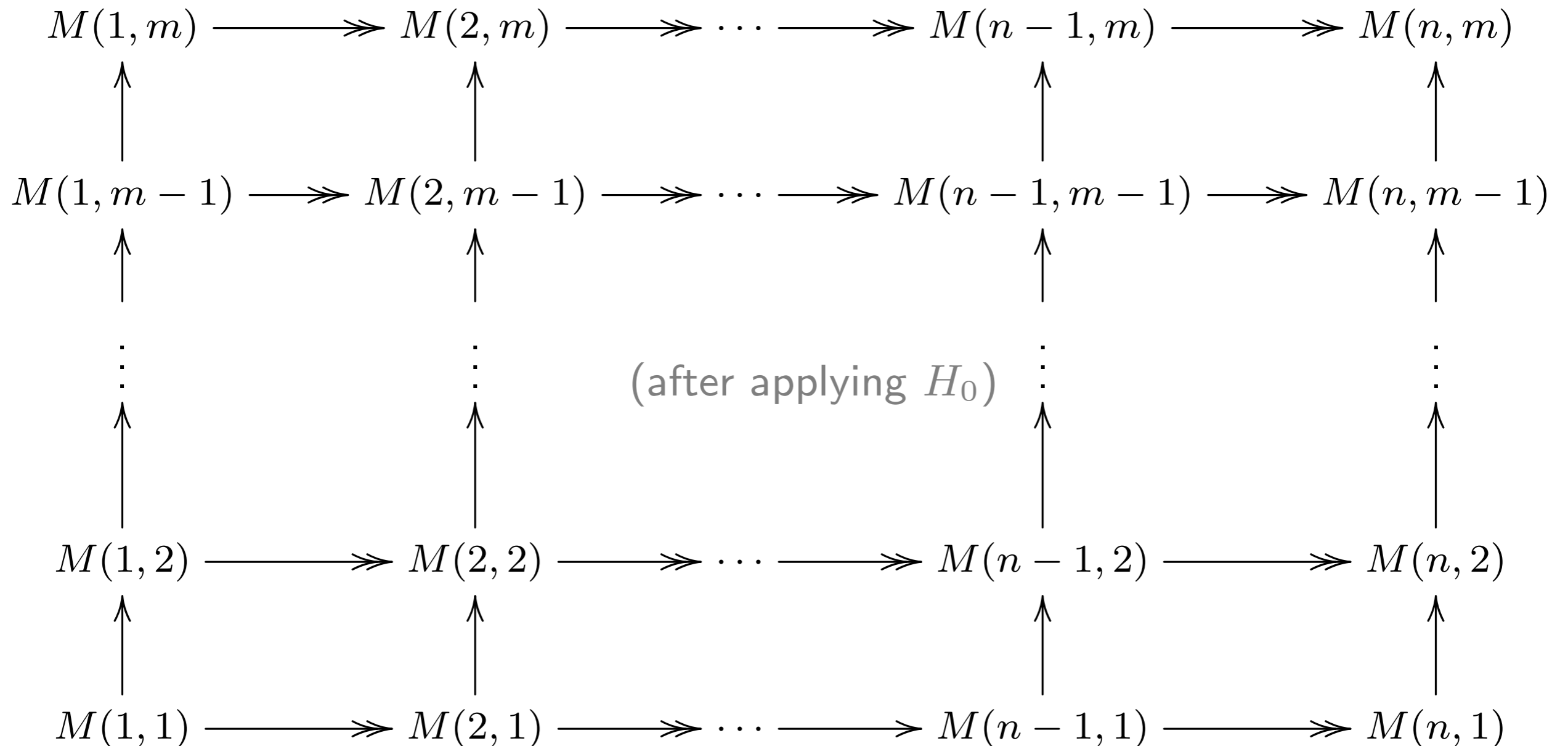
Example: two-parameter clustering



Example: two-parameter clustering



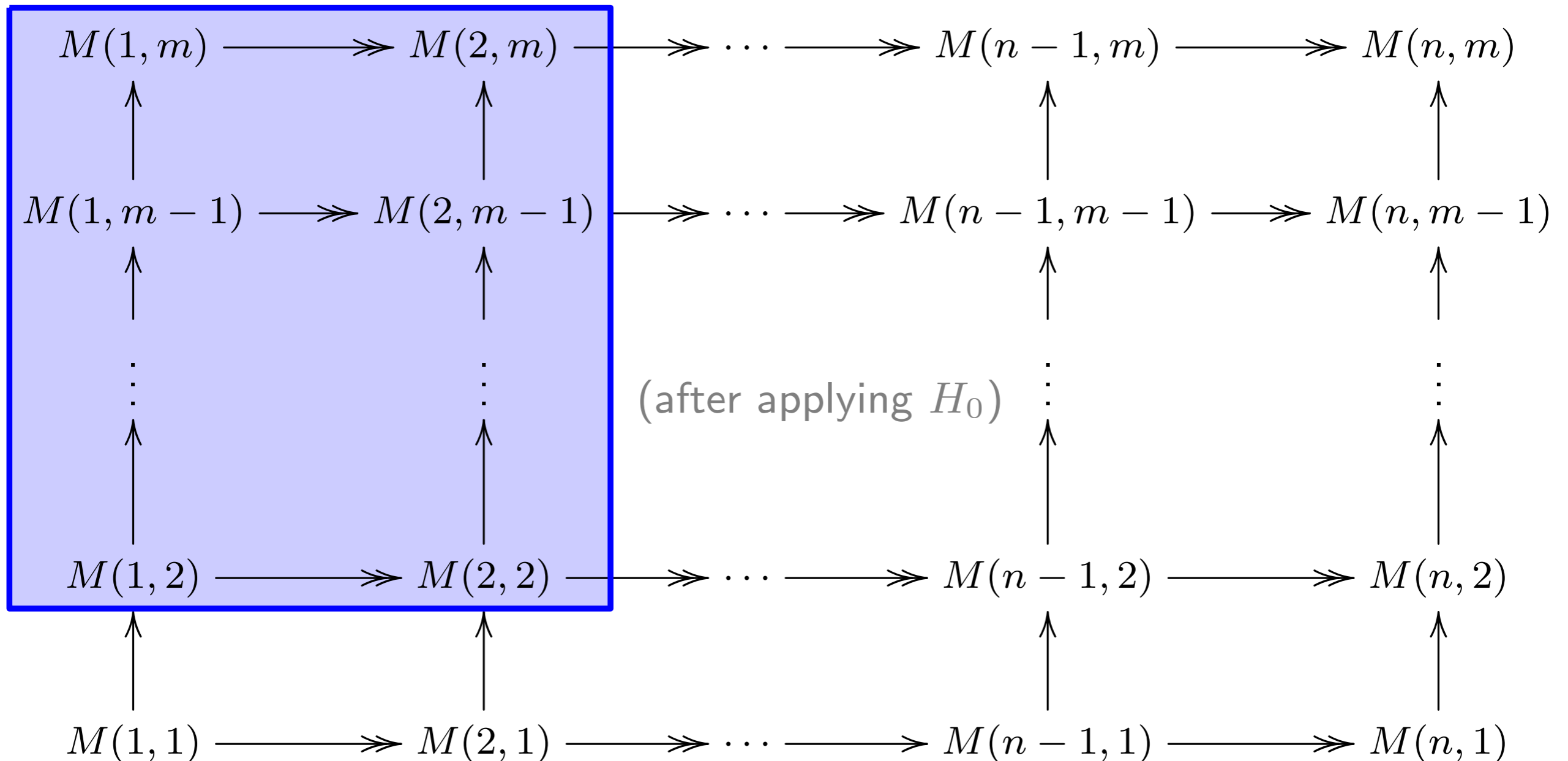
Example: two-parameter clustering



Thm: [Bauer, Botnan, Oppermann, Steen]

$$\frac{\text{Fun}^{e,*}(\llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket, \text{vect}_{\mathbf{k}})}{\text{Fun}^{e,m}(\llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket, \text{vect}_{\mathbf{k}})} \simeq \text{Fun}(\llbracket 1, n \rrbracket \times \llbracket 1, m-1 \rrbracket, \text{vect}_{\mathbf{k}})$$

Example: two-parameter clustering



Thm: [Bauer, Botnan, Oppermann, Steen]

$$\text{Fun}^{e,*}(\llbracket 1, n \rrbracket \times \llbracket 1, m \rrbracket, \text{vect}_{\mathbf{k}}) \simeq \text{Fun}(\llbracket 1, n \rrbracket \times \llbracket 1, m-1 \rrbracket, \text{vect}_{\mathbf{k}})$$

modules decompose into
summands $\mathbf{k}_{\llbracket 1, i \rrbracket \times \llbracket j, m \rrbracket}$

Incomplete invariants

Bottomline: look for incomplete invariants of persistence modules that are:

- ▶ as strong as possible (μ stronger than ν if $\mu(M) = \mu(N) \Rightarrow \nu(M) = \nu(N)$)
- ▶ manageable to compute (polynomial time in the input filtration size)
- ▶ stable w.r.t. perturbations of the modules in the interleaving distance d_i
- ▶ interpretable in terms of the module's structure

Incomplete invariants

Bottomline: look for incomplete invariants of persistence modules that are:

- ▶ as strong as possible (μ stronger than ν if $\mu(M) = \mu(N) \Rightarrow \nu(M) = \nu(N)$)
- ▶ manageable to compute (polynomial time in the input filtration size)
- ▶ stable w.r.t. perturbations of the modules in the interleaving distance d_i
- ▶ interpretable in terms of the module's structure

Examples:

- ▶ dimension vector / Hilbert function
- ▶ rank invariant
- ▶ global rank function / generalized rank invariant
- ▶ graded Betti numbers
- ▶ ...

Incomplete invariants

Bottomline: look for incomplete invariants of persistence modules that are:

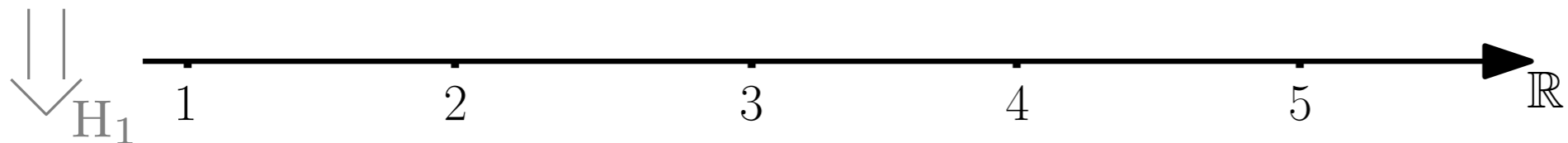
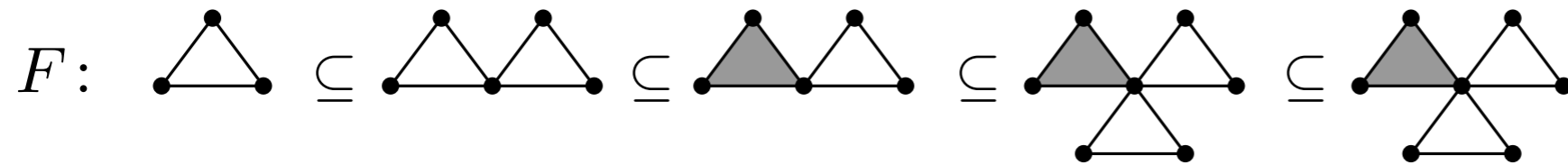
- ▶ as strong as possible (μ stronger than ν if $\mu(M) = \mu(N) \Rightarrow \nu(M) = \nu(N)$)
- ▶ manageable to compute (polynomial time in the input filtration size)
- ▶ stable w.r.t. perturbations of the modules in the interleaving distance d_i
- ▶ interpretable in terms of the module's structure

Examples:

- ▶ dimension vector / Hilbert function
- ▶ rank invariant
- ▶ global rank function / generalized rank invariant
- ▶ graded Betti numbers
- ▶ ...

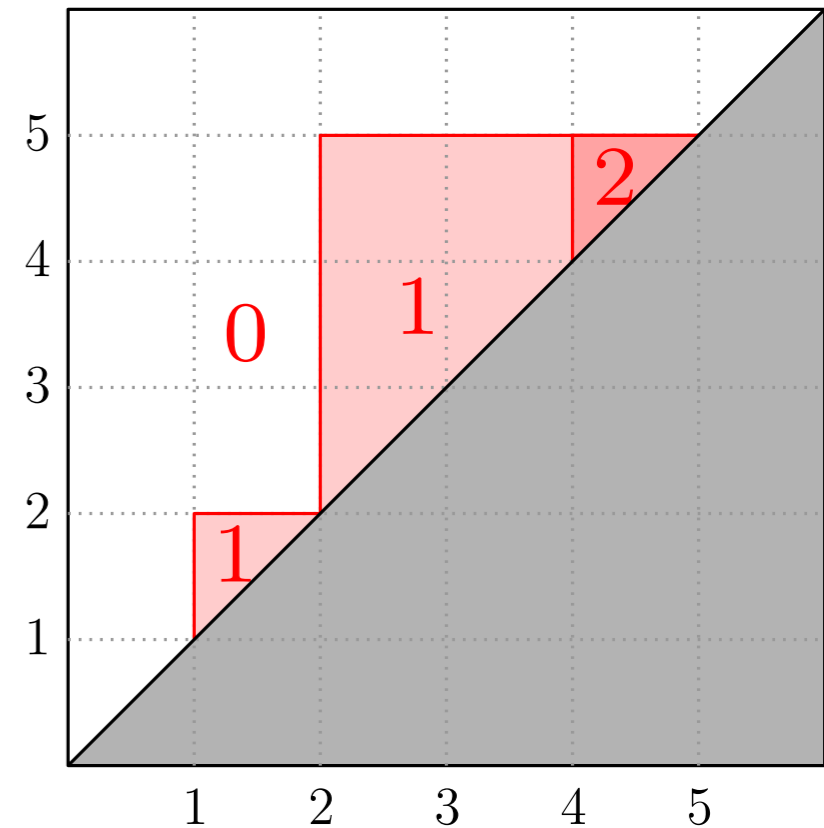
The rank invariant of 1-parameter modules

$$P = [1, 5] \subseteq (\mathbb{R}, \leq)$$



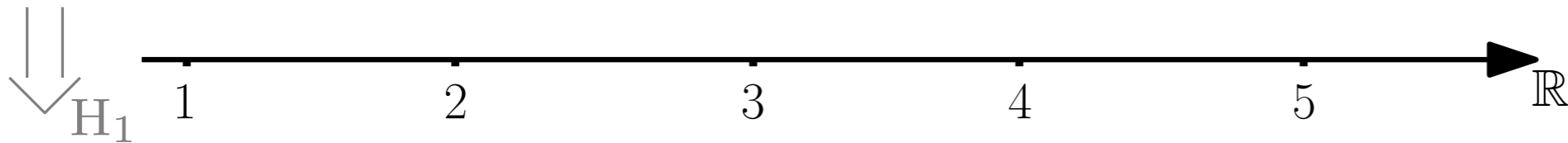
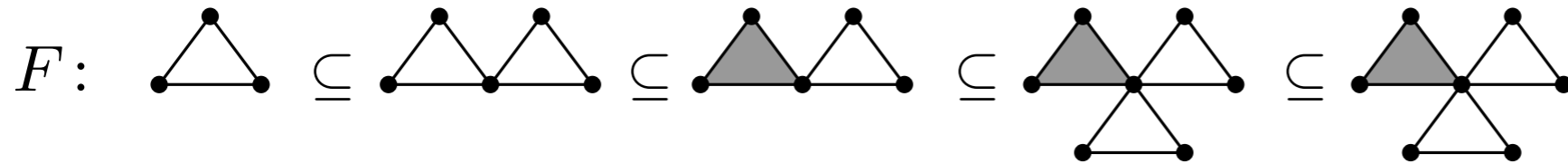
$$M: \mathbf{k} \xrightarrow{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 0 & 1 \end{pmatrix}} \mathbf{k} \xrightarrow{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \mathbf{k}^2$$

$$\mathbf{Rk} M: (s \leq t) \mapsto \text{rank} [M_s \rightarrow M_t]$$



The rank invariant of 1-parameter modules

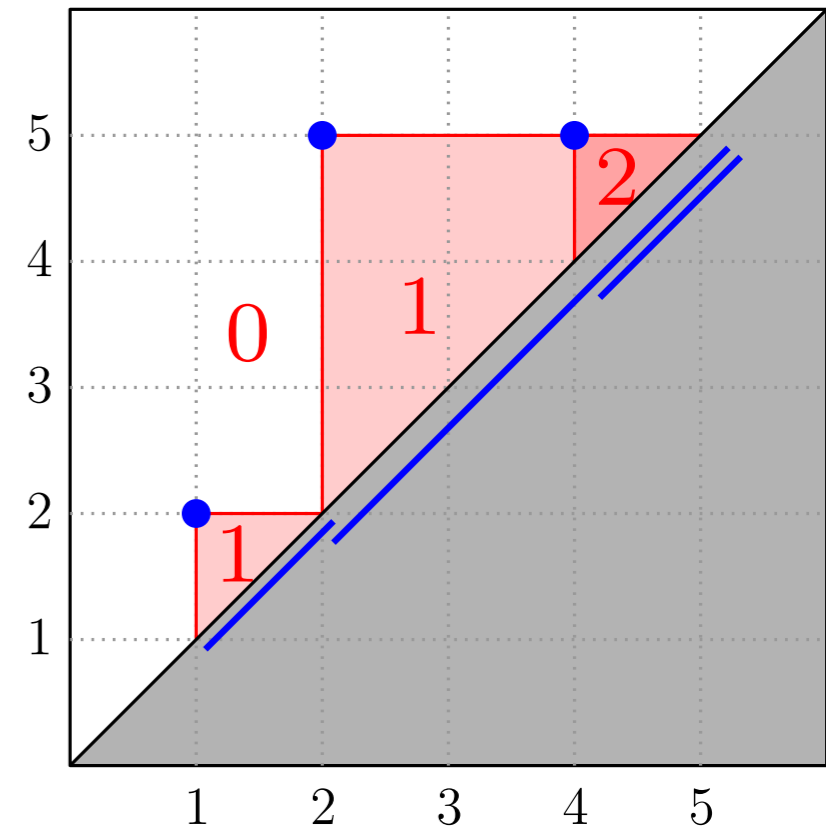
$$P = [1, 5] \subseteq (\mathbb{R}, \leq)$$



$$M: \mathbf{k} \xrightarrow{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 0 & 1 \end{pmatrix}} \mathbf{k} \xrightarrow{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \mathbf{k}^2$$

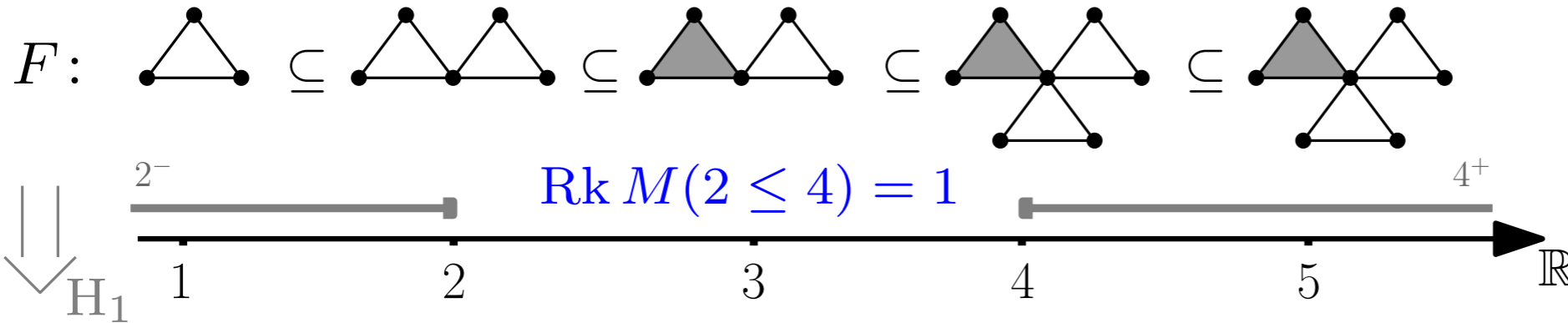
$$\text{Rk } M = \sum_{I \in \text{Bar } M} \text{Rk } \mathbf{k}_I = \text{Rk} \left(\bigoplus_{I \in \text{Bar } M} \mathbf{k}_I \right)$$

(rank invariant \mathbb{N} -decomposes on interval rank functions)




The rank invariant of 1-parameter modules

$$P = [1, 5] \subseteq (\mathbb{R}, \leq)$$

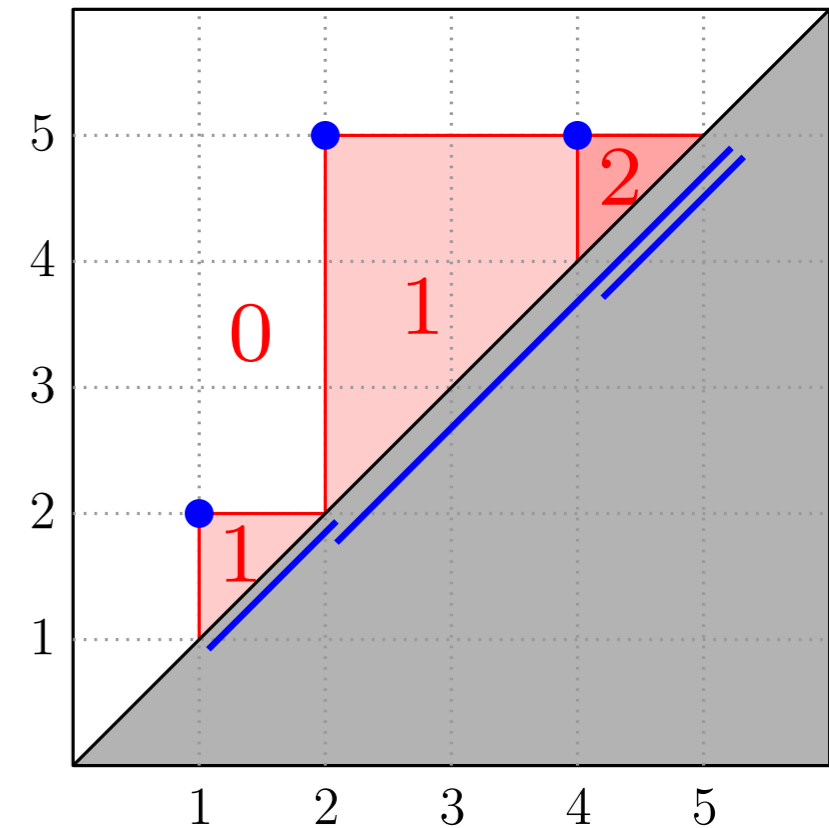


$$M: \mathbf{k} \xrightarrow{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 0 & 1 \end{pmatrix}} \mathbf{k} \xrightarrow{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \mathbf{k}^2$$



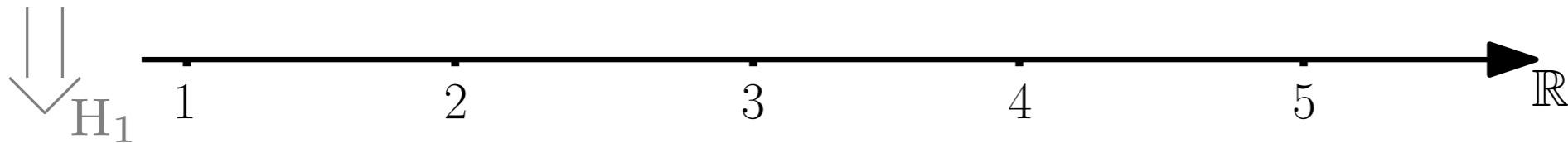
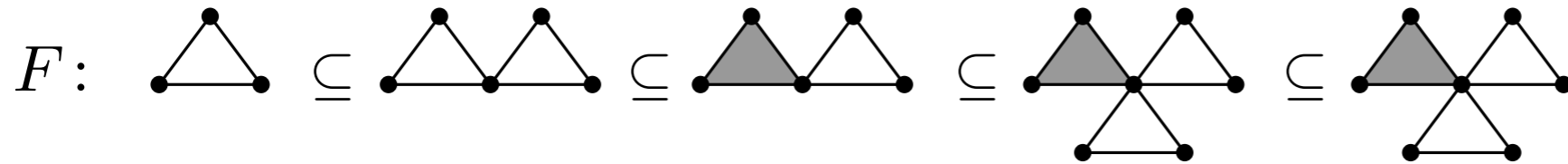
$$\text{Rk } M = \sum_{I \in \text{Bar } M} \text{Rk } \mathbf{k}_I = \text{Rk} \left(\bigoplus_{I \in \text{Bar } M} \mathbf{k}_I \right)$$

(rank invariant \mathbb{N} -decomposes on interval rank functions)



The rank invariant of 1-parameter modules

$$P = [1, 5] \subseteq (\mathbb{R}, \leq)$$



$$M: \mathbf{k} \xrightarrow{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 0 & 1 \end{pmatrix}} \mathbf{k} \xrightarrow{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \mathbf{k}^2$$

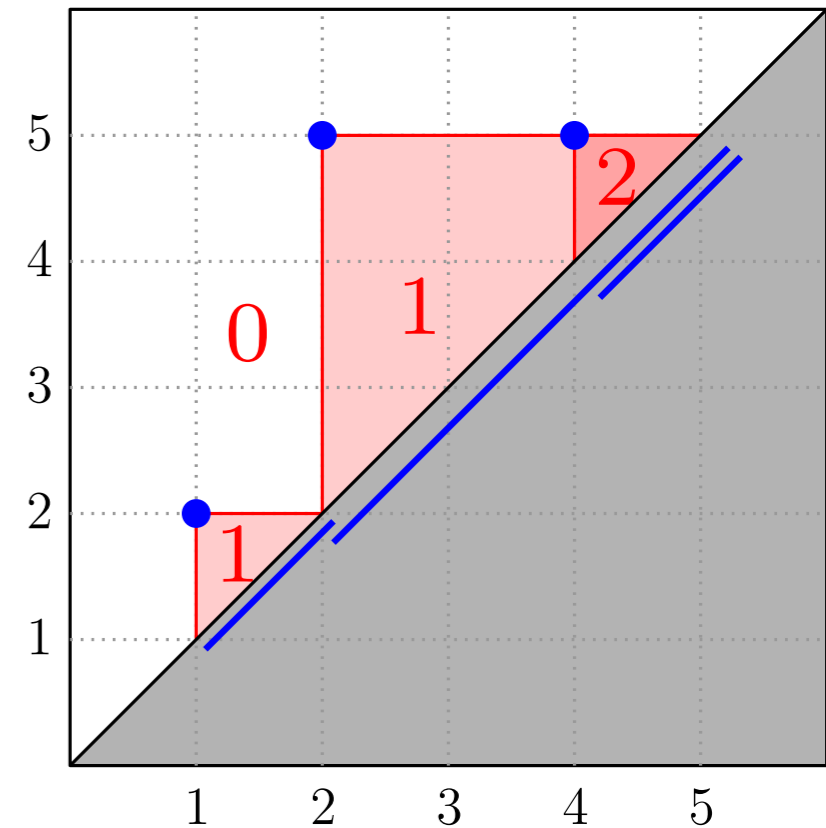
Bar M :



$$\text{Rk } M = \sum_{I \in \text{Bar } M} \text{Rk } \mathbf{k}_I = \text{Rk} \left(\bigoplus_{I \in \text{Bar } M} \mathbf{k}_I \right)$$

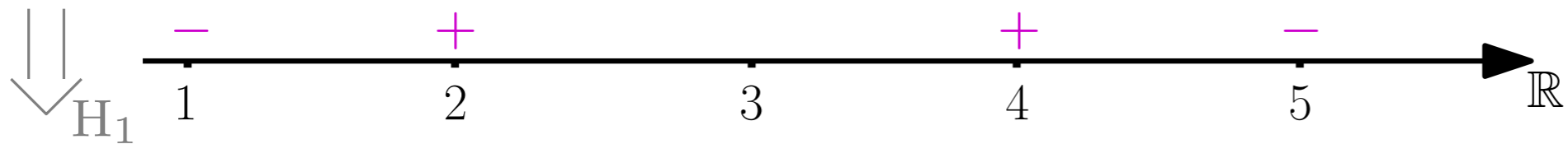
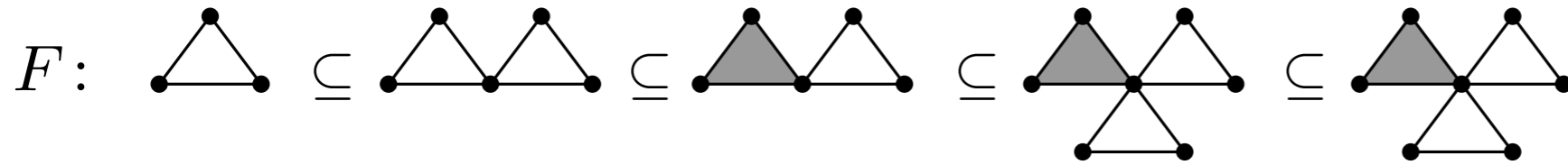
(unique decomposition, terms match with summands of M)

$$M \simeq \bigoplus_{I \in \text{Bar } M} \mathbf{k}_I \quad (\Rightarrow \text{rank invariant is complete})$$



The rank invariant of 1-parameter modules

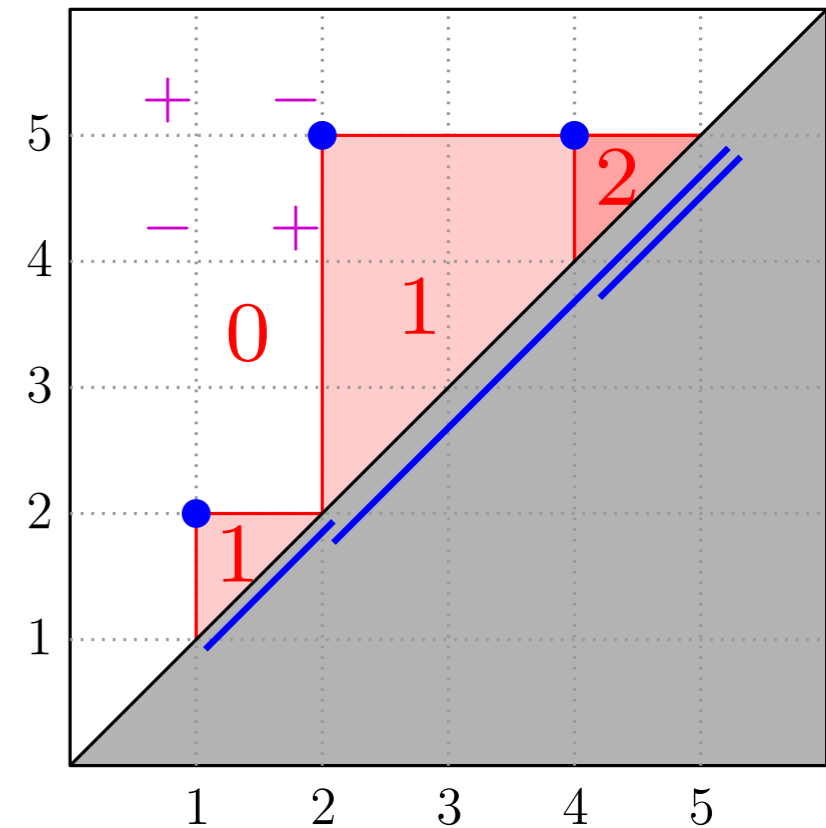
$$P = \llbracket 1, 5 \rrbracket \subseteq (\mathbb{R}, \leq)$$



$$M: \mathbf{k} \xrightarrow{\begin{pmatrix} 1 \\ 0 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 0 & 1 \end{pmatrix}} \mathbf{k} \xrightarrow{\begin{pmatrix} 0 \\ 1 \end{pmatrix}} \mathbf{k}^2 \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} \mathbf{k}^2$$

$$\text{Rk } M = \sum_{I \in \text{Bar } M} \text{Rk } \mathbf{k}_I = \text{Rk} \left(\bigoplus_{I \in \text{Bar } M} \mathbf{k}_I \right)$$

$$\begin{aligned} \text{mult}_{\llbracket i, j \rrbracket} \text{Bar } M &= \text{Rk } M(i, j) - \text{Rk } M(i-1, j) \\ &\quad - \text{Rk } M(i, j+1) + \text{Rk } M(i-1, j+1) \end{aligned}$$



The rank invariant of multi-parameter modules

$$P = \llbracket 1, 3 \rrbracket^2 \subseteq (\mathbb{R}^2, \leq)$$

$$\text{Rk } M : (s \leq t) \mapsto \text{rank} [M_s \rightarrow M_t]$$

$$\text{Rk} \left(\begin{array}{c} \begin{array}{ccccc} \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow \text{id} & & \uparrow \text{id} & & \uparrow \text{id} \\ \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow & & \uparrow \text{id} & & \uparrow \text{id} \\ 0 & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \end{array} \\ \oplus \\ \begin{array}{ccccc} 0 & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow & & \uparrow \text{id} & & \uparrow \text{id} \\ 0 & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \xrightarrow{\text{id}} & 0 & \xrightarrow{\text{id}} & 0 \end{array} \end{array} \right) = \text{Rk} \left(\begin{array}{c} \begin{array}{ccccc} 0 & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow & & \uparrow \text{id} & & \uparrow \text{id} \\ 0 & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow & & \uparrow \text{id} & & \uparrow \text{id} \\ 0 & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \end{array} \\ \oplus \\ \begin{array}{ccccc} \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow \text{id} & & \uparrow \text{id} & & \uparrow \text{id} \\ \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ \uparrow & & \uparrow & & \uparrow \\ 0 & \xrightarrow{\text{id}} & 0 & \xrightarrow{\text{id}} & 0 \end{array} \end{array} \right)$$

(rank invariant is not complete)

The rank invariant of multi-parameter modules

$$P = \llbracket 1, 3 \rrbracket^2 \subseteq (\mathbb{R}^2, \leq)$$

$$\text{Rk} \left(\begin{array}{ccccc} & & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \longrightarrow & 0 \\ & & \uparrow & & \uparrow & & \uparrow \\ & & \mathbf{k} & \xrightarrow{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} & \mathbf{k}^2 & \xrightarrow{\begin{bmatrix} 1 & 1 \end{bmatrix}} & \mathbf{k} \\ & & \uparrow & & \uparrow & & \uparrow \\ & & 0 & \longrightarrow & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \\ & & & & \uparrow & & \uparrow \\ & & & & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \end{array} \right) = \text{Rk} \left(\begin{array}{c} \text{?} \\ \text{?} \end{array} \right)$$

(rank invariant does not \mathbb{N} -decompose on interval rank functions)

The rank invariant of multi-parameter modules

$$P = \llbracket 1, 3 \rrbracket^2 \subseteq (\mathbb{R}^2, \leq)$$

$$\text{Rk} \left(\begin{array}{ccccc} \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \longrightarrow & 0 \\ \uparrow \text{id} & & \uparrow [1 \ 0] & & \uparrow \\ \mathbf{k} & \xrightarrow{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} & \mathbf{k}^2 & \xrightarrow{[1 \ 1]} & \mathbf{k} \\ \uparrow & & \uparrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} & & \uparrow \text{id} \\ 0 & \longrightarrow & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \end{array} \right) = \text{Rk} \left(\begin{array}{ccc} \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \end{array} \oplus \begin{array}{ccc} \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \end{array} \oplus \begin{array}{ccc} \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \end{array} \right)$$

$$- \text{Rk} \left(\begin{array}{ccc} \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet & \longrightarrow & \bullet \end{array} \right)$$

(rank invariant \mathbb{Z} -decomposes on interval rank functions)

The rank invariant of multi-parameter modules

$$P = \llbracket 1, 3 \rrbracket^2 \subseteq (\mathbb{R}^2, \leq)$$

$$\text{Rk} \left(\begin{array}{ccccc} \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \longrightarrow & 0 \\ \uparrow \text{id} & & \uparrow [1 \ 0] & & \uparrow \\ \mathbf{k} & \xrightarrow{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} & \mathbf{k}^2 & \xrightarrow{[1 \ 1]} & \mathbf{k} \\ \uparrow & & \uparrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} & & \uparrow \text{id} \\ 0 & \longrightarrow & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \end{array} \right) = \text{Rk} \left(\begin{array}{ccc} \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} & \oplus & \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} \\ \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} & \oplus & \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} \end{array} \right)$$

(rank invariant \mathbb{Z} -decomposes on interval rank functions)

The rank invariant of multi-parameter modules

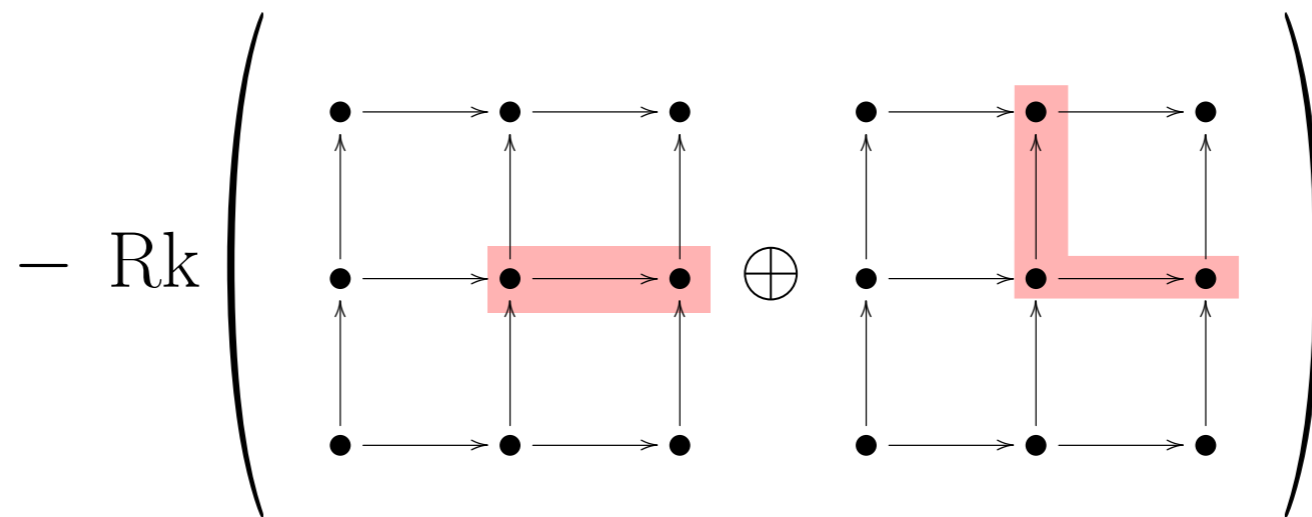
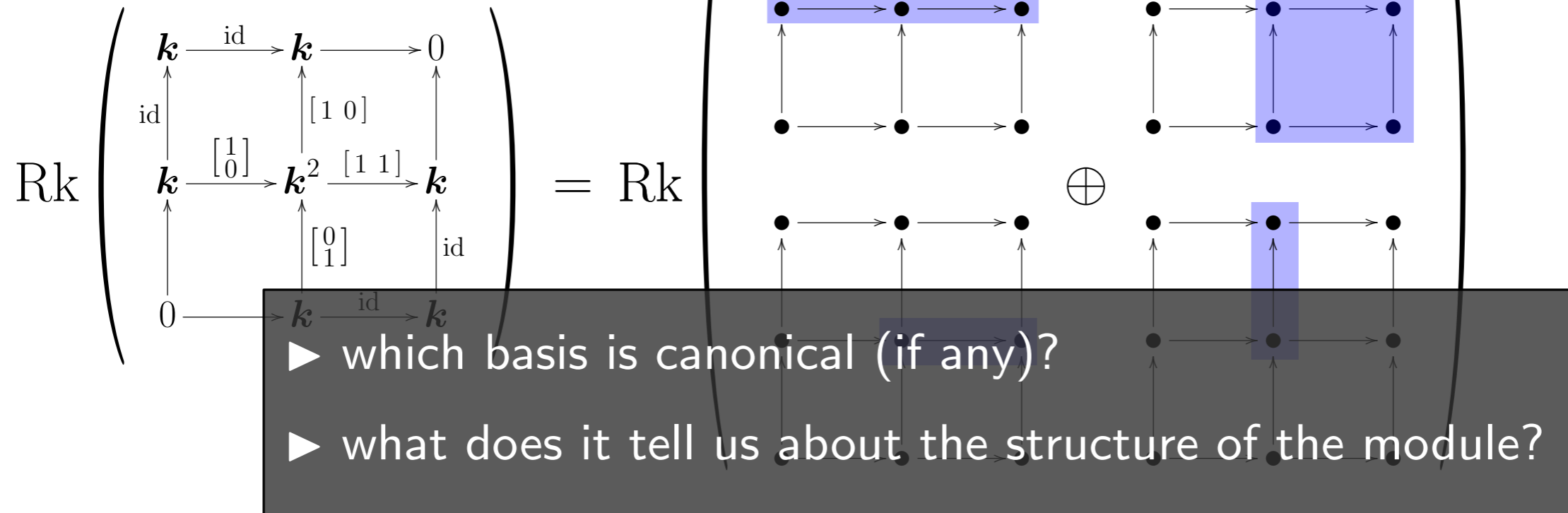
$$P = \llbracket 1, 3 \rrbracket^2 \subseteq (\mathbb{R}^2, \leq)$$

$$\text{Rk} \left(\begin{array}{ccccc} \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} & \longrightarrow & 0 \\ \uparrow \text{id} & & \uparrow [1 \ 0] & & \uparrow \\ \mathbf{k} & \xrightarrow{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} & \mathbf{k}^2 & \xrightarrow{[1 \ 1]} & \mathbf{k} \\ \uparrow & & \uparrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} & & \uparrow \text{id} \\ 0 & \longrightarrow & \mathbf{k} & \xrightarrow{\text{id}} & \mathbf{k} \end{array} \right) = \text{Rk} \left(\begin{array}{ccc} \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} & \oplus & \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} \\ \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} & \oplus & \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} \\ \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} & \oplus & \begin{array}{ccc} \bullet & \longrightarrow & \bullet \\ \uparrow & & \uparrow \\ \bullet & \longrightarrow & \bullet \end{array} \end{array} \right)$$

(rank invariant \mathbb{Z} -decomposes on interval rank functions)

The rank invariant of multi-parameter modules

$$P = \llbracket 1, 3 \rrbracket^2 \subseteq (\mathbb{R}^2, \leq)$$



(rank invariant \mathbb{Z} -decomposes on interval rank functions)

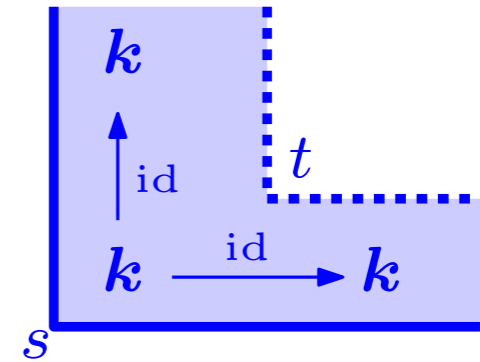
Canonical basis for the rank invariant

Hook / Upset: for $s < t \in P \cup \{\infty\}$,

$$\langle s, t \rangle = \{u \in P \mid s \leq u \not\leq t\}$$

$$s^+ = \langle s, \infty \rangle$$

$$\begin{array}{c} \uparrow \\ P = \mathbb{R}^2 \\ \rightarrow \end{array}$$



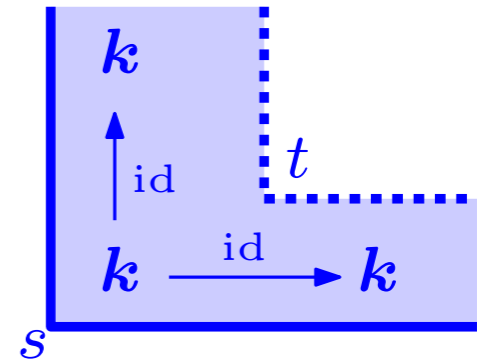
Canonical basis for the rank invariant

Hook / Upset: for $s < t \in P \cup \{\infty\}$,

$$\langle s, t \rangle = \{u \in P \mid s \leq u \not\leq t\}$$

$$s^+ = \langle s, \infty \rangle$$

$$\uparrow P = \mathbb{R}^2$$



Theorem ([Botnan, Oppermann, O.]):

If P is finite or an upper semi-lattice, then

$$\{\text{Rk } \mathbf{k}_{\langle i, j \rangle} \mid i < j \in P \cup \{\infty\}\}$$

generates uniquely

$$\{\text{Rk } M \mid M: P \rightarrow \text{vect}_{\mathbf{k}} \text{ finitely presentable (fp)}\}$$

via projective resolutions relative to the rank-exact structure \mathcal{E}_{Rk} .

The rank-exact structure

Let \mathcal{E}_{Rk} be the collection of short *rank-exact* sequences of pfd persistence modules:

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0 \quad \text{s.t.} \quad \text{Rk } B = \text{Rk } A + \text{Rk } C$$

The rank-exact structure

Let \mathcal{E}_{Rk} be the collection of short *rank-exact* sequences of pfd persistence modules:

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0 \quad \text{s.t.} \quad \text{Rk } B = \text{Rk } A + \text{Rk } C$$

Proposition: (follows in part from [Auslander, Solberg])

- \mathcal{E}_{Rk} defines the structure of an exact category on $\text{vect}_{\mathbf{k}}^P$
- the indecomposable projectives relative to \mathcal{E}_{Rk} are the hook modules
- P finite \implies $\left\{ \begin{array}{l} \mathcal{E}_{\text{Rk}} \text{ has enough projectives} \\ \text{every } M \in \text{vect}_{\mathbf{k}}^P \text{ has a finite projective resolution} \end{array} \right.$
- P upper semi-lattice \implies $\left\{ \begin{array}{l} \mathcal{E}_{\text{Rk}}^{\text{fp}} \text{ has enough projectives} \\ \text{every fp } M \in \text{Vect}_{\mathbf{k}}^P \text{ has a finite projective resolution} \end{array} \right.$

The rank-exact structure

Let \mathcal{E}_{Rk} be the collection of short *rank-exact* sequences of pfd persistence modules:

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0 \quad \text{s.t.} \quad \text{Rk } B = \text{Rk } A + \text{Rk } C$$
$$[B] = [A] + [C]$$

Corollary:

- P finite \implies the $[\mathbf{k}_{\langle i, j \rangle}]$ for $i < j \in P \cup \{\infty\}$ generate $K_0(\mathcal{E}_{\text{Rk}})$
- P upper semi-lattice \implies the $[\mathbf{k}_{\langle i, j \rangle}]$ for $i < j \in P \cup \{\infty\}$ generate $K_0(\mathcal{E}_{\text{Rk}}^{\text{fp}})$

Given a finite rank-exact projective resolution M_\bullet of M :

$$[M] = \sum_{i=0}^{\infty} (-1)^i [M_i]$$

The rank-exact structure

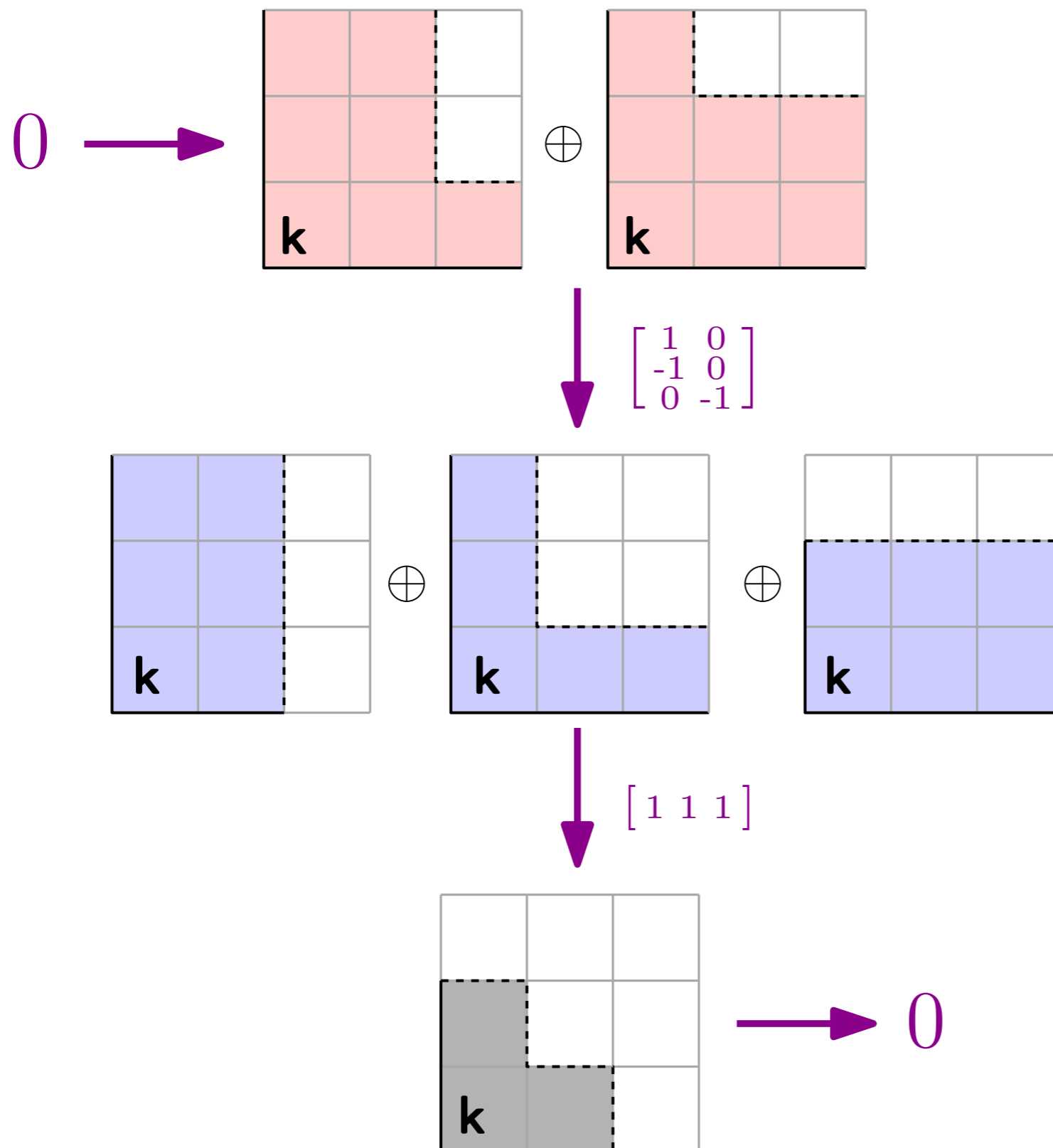
Let \mathcal{E}_{Rk} be the collection of short *rank-exact* sequences of pfd persistence modules:

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0 \quad \text{s.t.} \quad \text{Rk } B = \text{Rk } A + \text{Rk } C$$
$$[B] = [A] + [C]$$

Corollary:

- P finite \implies the $[\mathbf{k}_{\langle i, j \rangle}]$ for $i < j \in P \cup \{\infty\}$ generate $K_0(\mathcal{E}_{\text{Rk}})$
- P upper semi-lattice \implies the $[\mathbf{k}_{\langle i, j \rangle}]$ for $i < j \in P \cup \{\infty\}$ generate $K_0(\mathcal{E}_{\text{Rk}}^{\text{fp}})$
- In both cases, the $[\mathbf{k}_{\langle i, j \rangle}]$ actually form a basis, and Rk defines a monomorphism of abelian groups $K_0(\mathcal{E}) \rightarrow \mathbb{Z}^{\text{Seg}(P)}$

The rank-exact structure



The rank-exact structure

$$-Rk \left(\begin{array}{c} \begin{array}{|c|c|c|} \hline \color{red} \square & \color{red} \square & \square \\ \hline \color{red} \square & \color{red} \square & \square \\ \hline \color{red} \square & \color{red} \square & \square \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \color{red} \square & \square & \square \\ \hline \color{red} \square & \color{red} \square & \color{red} \square \\ \hline \color{red} \square & \color{red} \square & \color{red} \square \\ \hline \end{array} \end{array} \right)$$

$$+Rk \left(\begin{array}{c} \begin{array}{|c|c|c|} \hline \color{blue} \square & \color{blue} \square & \square \\ \hline \color{blue} \square & \color{blue} \square & \square \\ \hline \color{blue} \square & \color{blue} \square & \square \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \color{blue} \square & \square & \square \\ \hline \color{blue} \square & \square & \square \\ \hline \color{blue} \square & \color{blue} \square & \color{blue} \square \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \color{blue} \square & \color{blue} \square & \color{blue} \square \\ \hline \color{blue} \square & \color{blue} \square & \color{blue} \square \\ \hline \end{array} \end{array} \right)$$

$$= Rk \left(\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \color{gray} \square & \color{gray} \square & \square \\ \hline \color{gray} \square & \color{gray} \square & \square \\ \hline \end{array} \right)$$

The rank-exact structure

$$-Rk \left(\begin{array}{c} \begin{array}{|c|c|c|} \hline \color{red}{\square} & \color{red}{\square} & \square \\ \hline \color{red}{\square} & \color{red}{\square} & \square \\ \hline \color{red}{\square} & \color{red}{\square} & \square \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \color{red}{\square} & \square & \square \\ \hline \color{red}{\square} & \color{red}{\square} & \color{red}{\square} \\ \hline \color{red}{\square} & \color{red}{\square} & \color{red}{\square} \\ \hline \end{array} \end{array} \right)$$

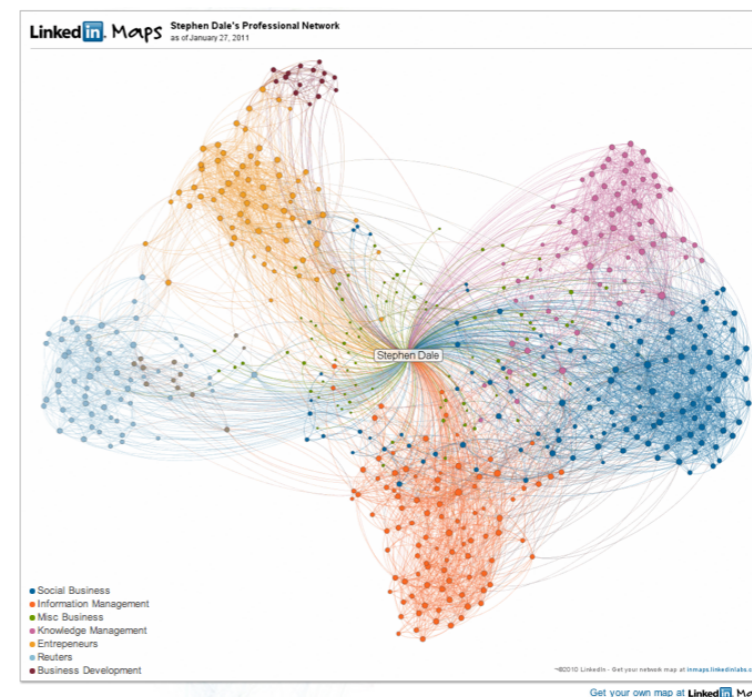
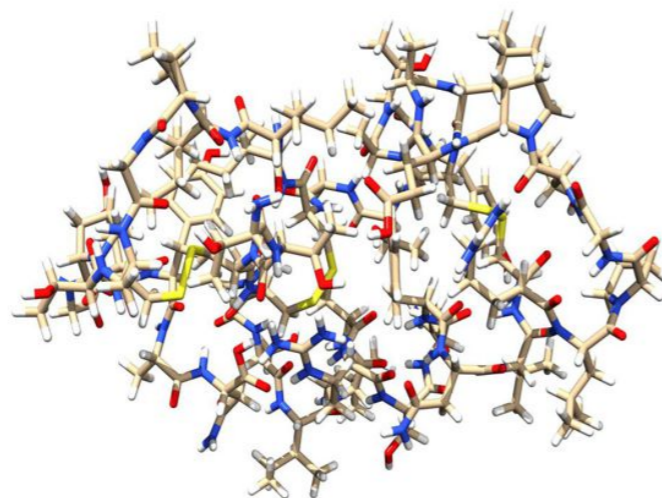
identical terms cancel out in minimal rank decomposition

$$+Rk \left(\begin{array}{c} \begin{array}{|c|c|c|} \hline \color{blue}{\square} & \color{blue}{\square} & \square \\ \hline \color{blue}{\square} & \color{blue}{\square} & \square \\ \hline \color{blue}{\square} & \color{blue}{\square} & \square \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \color{blue}{\square} & \square & \square \\ \hline \color{blue}{\square} & \square & \square \\ \hline \color{blue}{\square} & \color{blue}{\square} & \color{blue}{\square} \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \color{blue}{\square} & \color{blue}{\square} & \color{blue}{\square} \\ \hline \color{blue}{\square} & \color{blue}{\square} & \color{blue}{\square} \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \color{blue}{\square} & \square & \square \\ \hline \color{blue}{\square} & \color{blue}{\square} & \color{blue}{\square} \\ \hline \color{blue}{\square} & \color{blue}{\square} & \color{blue}{\square} \\ \hline \end{array} \end{array} \right)$$

$$= Rk \left(\begin{array}{c} \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \color{gray}{\square} & \square & \square \\ \hline \color{gray}{\square} & \color{gray}{\square} & \square \\ \hline \end{array} \oplus \begin{array}{|c|c|c|} \hline \color{gray}{\square} & \square & \square \\ \hline \color{gray}{\square} & \color{gray}{\square} & \color{gray}{\square} \\ \hline \color{gray}{\square} & \color{gray}{\square} & \color{gray}{\square} \\ \hline \end{array} \end{array} \right)$$

Application to graph classification

- ▶ various classification tasks on molecules and social network graphs
- ▶ 5-fold evaluation (80% train, 20% test)



Dataset	COX2	DHFR	IMDB-B	IMDB-M	MUTAG	PROTEINS
1-d barcode	76.0(4.1)	70.9(3.1)	54.0(1.9)	36.3(1.1)	79.2(7.7)	65.4(2.7)
MP-Kernel	79.9(1.8)	81.7(1.9)	68.2(1.2)	46.9(2.6)	86.1(5.2)	67.5(3.1)
MP-Landscapes	79.0(3.3)	79.5(2.3)	71.2(2.0)	46.2(2.3)	84.0(6.8)	65.8(3.3)
MP-Images	77.9(2.7)	80.2(2.2)	71.1(2.1)	46.7(2.7)	85.6(7.3)	67.3(3.5)
GRIL	79.8(2.9)	77.6(2.5)	65.2(2.6)	NA	87.8(4.2)	70.9(3.1)
Rank inv.	78.2(1.7)	79.9(2.1)	73.0(4.5)	49.1(1.6)	87.2(5.8)	70.2(2.1)
Rank decomp.	78.4(0.7)	78.7(1.7)	75.1(3.4)	51.1(1.3)	89.9(4.3)	73.9(1.7)
Baseline (10-fold)	80.1	81.5	74.3	52.4	92.1	76.3

More on homological invariants for TDA

- ▶ Framework: dim-Hom vs. homological invariants [Blanchette, Brüstle, Hanson]

<https://arxiv.org/abs/2112.07632>

- ▶ Computation via Koszul complexes [Chacholski et al.]

<https://arxiv.org/abs/2209.05923>

- ▶ Hilbert function: decomposition and stability [O., Scoccola]

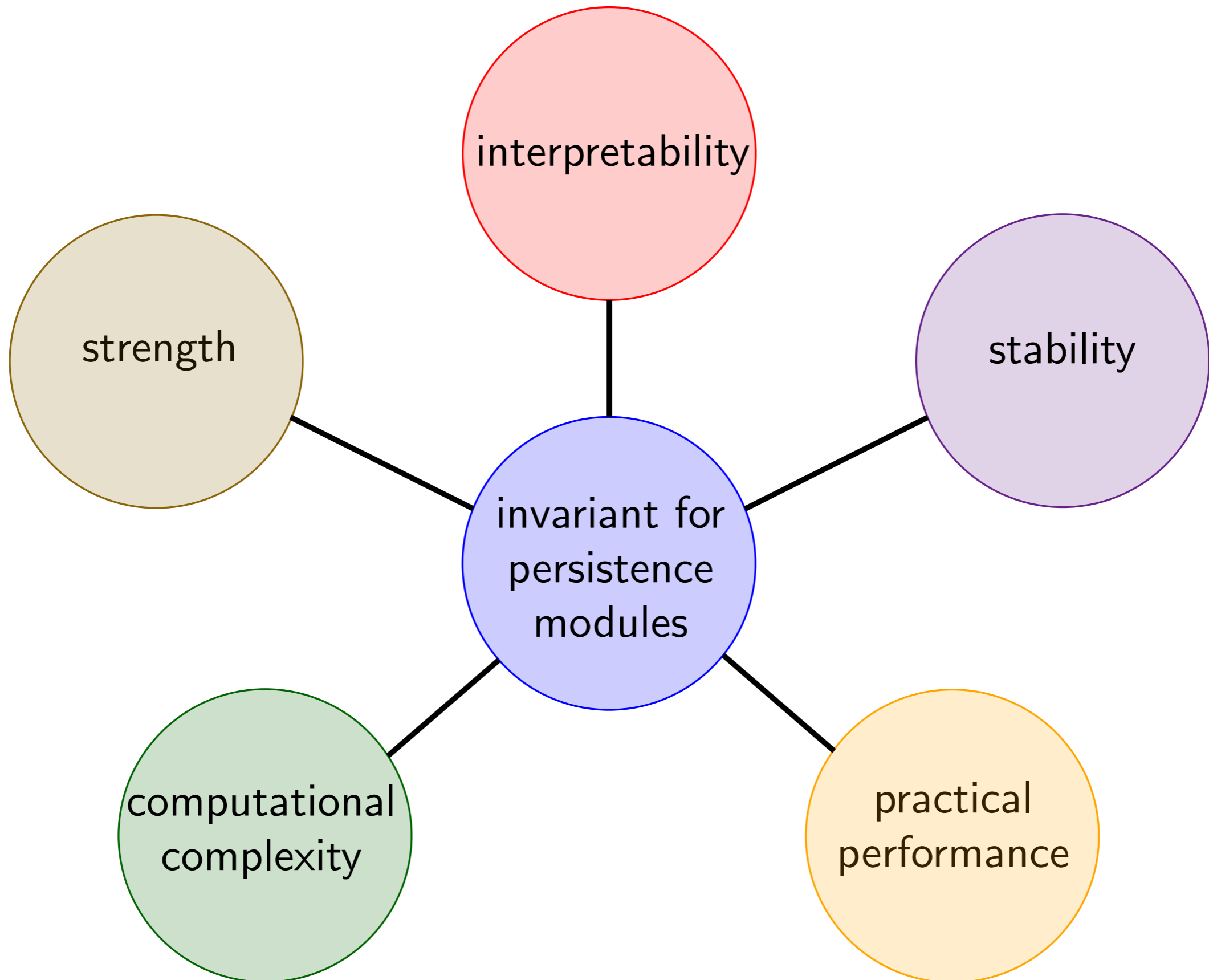
<https://arxiv.org/abs/2112.11901>

- ▶ Rank invariant: decomposition and stability [Botnan, Oppermann, O., Scoccola]

<https://arxiv.org/abs/2107.06800>

<https://arxiv.org/abs/2208.00300>

Take-home message



Homological invariants

Definition: An *additive invariant* is a map $\alpha: \text{Vect}_{\text{fp}}^{\mathbb{R}^d} \rightarrow A$, with A an Abelian group, such that $\alpha(M \oplus N) = \alpha(M) + \alpha(N)$ for all $M, N \in \text{Vect}_{\text{fp}}^{\mathbb{R}^d}$.

► Typical example (**dim-Hom** invariant): given a collection \mathcal{I} of intervals,

$$\alpha(-) := (\dim \text{Hom}(\mathbf{k}_I, -))_{I \in \mathcal{I}} \in \mathbb{Z}^{\mathcal{I}}$$

(\mathcal{I} = positive quadrants: Hilbert function)

(\mathcal{I} = hooks: $\dim \ker$ invariant, or equivalently rank invariant)

Homological invariants

Definition: An *additive invariant* is a map $\alpha: \text{Vect}_{\text{fp}}^{\mathbb{R}^d} \rightarrow A$, with A an Abelian group, such that $\alpha(M \oplus N) = \alpha(M) + \alpha(N)$ for all $M, N \in \text{Vect}_{\text{fp}}^{\mathbb{R}^d}$.

► Typical example (**dim-Hom** invariant): given a collection \mathcal{I} of intervals,

$$\alpha(-) := (\dim \text{Hom}(\mathbf{k}_I, -))_{I \in \mathcal{I}} \in \mathbb{Z}^{\mathcal{I}}$$

► Let \mathcal{E}_α be the collection of short exact sequences $0 \rightarrow L \rightarrow M \rightarrow N \rightarrow 0$ on which α is additive (use the shorthand $\mathcal{E}_\mathcal{I}$ when $\alpha = (\dim \text{Hom}(\mathbf{k}_I, -))_{I \in \mathcal{I}}$)

Proposition: (follows from [Auslander, Solberg])

If \mathcal{I} contains all the up-sets, then:

- $\mathcal{E}_\mathcal{I}$ forms an exact structure on $\text{Vect}_{\text{fp}}^{\mathbb{R}^d}$
- the indecomposable projectives relative to $\mathcal{E}_\mathcal{I}$ are the \mathbf{k}_I for $I \in \mathcal{I}$
- $\mathcal{E}_\mathcal{I}$ has enough projectives

► can do homological algebra with $\mathcal{E}_\mathcal{I}$

Homological invariants

Definition: An *additive invariant* is a map $\alpha: \text{Vect}_{\text{fp}}^{\mathbb{R}^d} \rightarrow A$, with A an Abelian group, such that $\alpha(M \oplus N) = \alpha(M) + \alpha(N)$ for all $M, N \in \text{Vect}_{\text{fp}}^{\mathbb{R}^d}$.

► Typical example (**dim-Hom** invariant): given a collection \mathcal{I} of intervals,

$$\alpha(-) := (\dim \text{Hom}(\mathbf{k}_I, -))_{I \in \mathcal{I}} \in \mathbb{Z}^{\mathcal{I}}$$

► Let \mathcal{E}_α be the collection of short exact sequences $0 \rightarrow L \rightarrow M \rightarrow N \rightarrow 0$ on which α is additive (use the shorthand $\mathcal{E}_\mathcal{I}$ when $\alpha = (\dim \text{Hom}(\mathbf{k}_I, -))_{I \in \mathcal{I}}$)

► for every fp M that has a finite minimal $\mathcal{E}_\mathcal{I}$ -resolution $M_n \rightarrow \cdots \rightarrow M_0 \twoheadrightarrow M$, we get a canonical decomposition:

$$\begin{aligned} (\dim \text{Hom}(\mathbf{k}_I, M))_{I \in \mathcal{I}} &= \sum_{j \in \mathbb{N}} (-1)^j (\dim \text{Hom}(M_j, M))_{I \in \mathcal{I}} \\ &= \sum_{j \in 2\mathbb{N}} (\dim \text{Hom}(M_j, M))_{I \in \mathcal{I}} - \sum_{j \in 2\mathbb{N}+1} (\dim \text{Hom}(M_j, M))_{I \in \mathcal{I}} \end{aligned}$$

► canonical signed barcode $(\beta_{2\mathbb{N}}^{\mathcal{I}}(M), \beta_{2\mathbb{N}+1}^{\mathcal{I}}(M))$ (with cancellable pairs)

Homological invariants

Question: size of the decompositions / length of the resolutions?

► case $\mathcal{I} = \{\text{up-sets}\}$: $\text{gldim} \left(\text{Vect}_{\text{fp}}^{\mathbb{R}^d} \right) = d$ (Hilbert's Syzygy theorem)

Homological invariants

Question: size of the decompositions / length of the resolutions?

► case $\mathcal{I} = \{\text{up-sets}\}$: $\text{gldim} \left(\text{Vect}_{\text{fp}}^{\mathbb{R}^d} \right) = d$ (Hilbert's Syzygy theorem)

► case $\mathcal{I} = \{\text{hooks}\}$:

$$\text{gldim}^{\mathcal{E}_{\mathcal{I}}} \left(\text{Vect}_{\text{fp}}^{\mathbb{R}^d} \right) = 2d - 2 \quad [\text{Botnan, Oppermann, O., Scoccola}]$$

Homological invariants

Question: size of the decompositions / length of the resolutions?

► case $\mathcal{I} = \{\text{up-sets}\}$: $\text{gldim} \left(\text{Vect}_{\text{fp}}^{\mathbb{R}^d} \right) = d$ (Hilbert's Syzygy theorem)

► case $\mathcal{I} = \{\text{hooks}\}$:

$$\text{gldim}^{\mathcal{E}_{\mathcal{I}}} \left(\text{Vect}_{\text{fp}}^{\mathbb{R}^d} \right) = 2d - 2 \quad [\text{Botnan, Oppermann, O., Scoccola}]$$

► more general cases: relate $\mathcal{E}_{\mathcal{I}}$ to the usual exact structure on some alternative module category (hypothesis: \mathbb{R}^d replaced by some finite poset P):

[Blanchette, Brüstle, Hanson] consider $\text{mod } \text{End}_{\mathbf{k}P} (T_{\mathcal{I}})^{\text{op}}$ where $T_{\mathcal{I}} = \bigoplus_{I \in \mathcal{I}} \mathbf{k}_I$

$$\text{add}(\{\mathbf{k}_I \mid I \in \mathcal{I}\}) \xrightarrow[\text{ff}]{\text{Hom}_{\mathbf{k}P}(T_{\mathcal{I}}, -)} \text{proj}(\text{mod } \text{End}_{\mathbf{k}P}(T_{\mathcal{I}})^{\text{op}})$$

$$\Rightarrow \text{gldim}^{\mathcal{E}_{\mathcal{I}}}(\text{mod } \mathbf{k}P) \leq \text{gldim}(\text{mod } \text{End}_{\mathbf{k}P}(T_{\mathcal{I}})^{\text{op}})$$