Plan of the course

1. ToMATo for colocalizing cell types

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

2. Rips persistence for marker gene correlations

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

3. Multi-persistence for immune cell arrangements

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

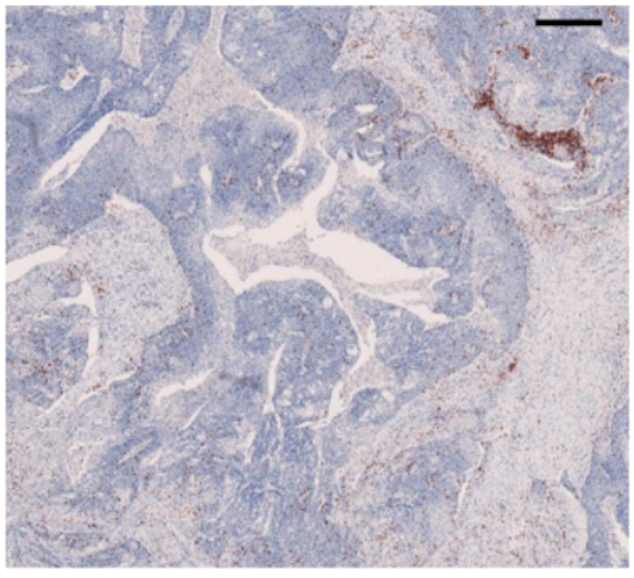
4. Future research directions

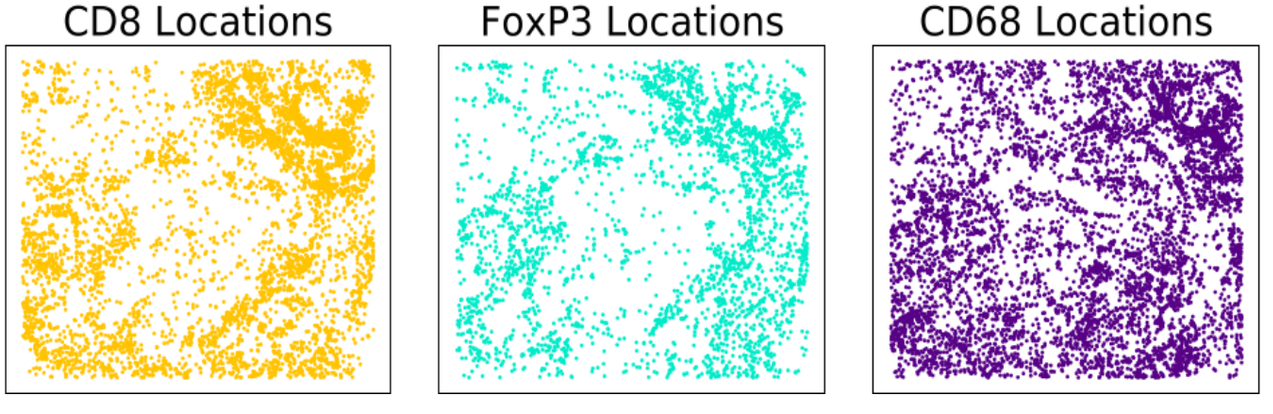
3. Multi-persistence for immune cell arrangements

3. Multi-persistence for immune cell arrangements

Q: How can one characterize the *geometric arrangement* of cells *of a given cell type* in a way that is robust to *cell type misclassifications*?

Indeed, cell types are often obtained by integration with scRNA-seq and prone to errors...

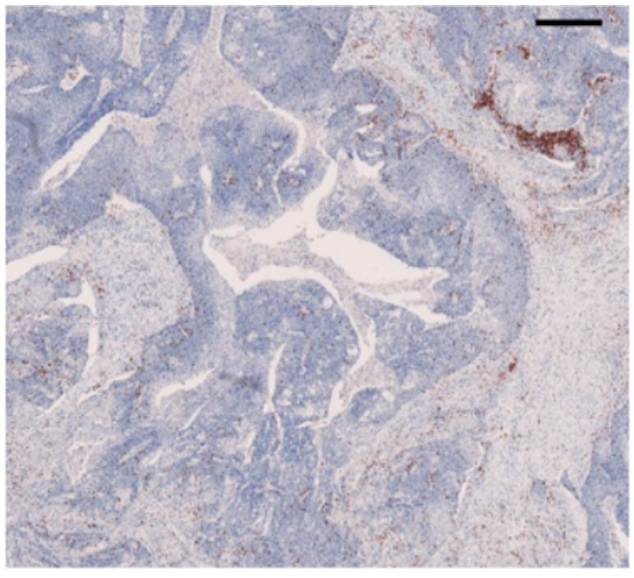




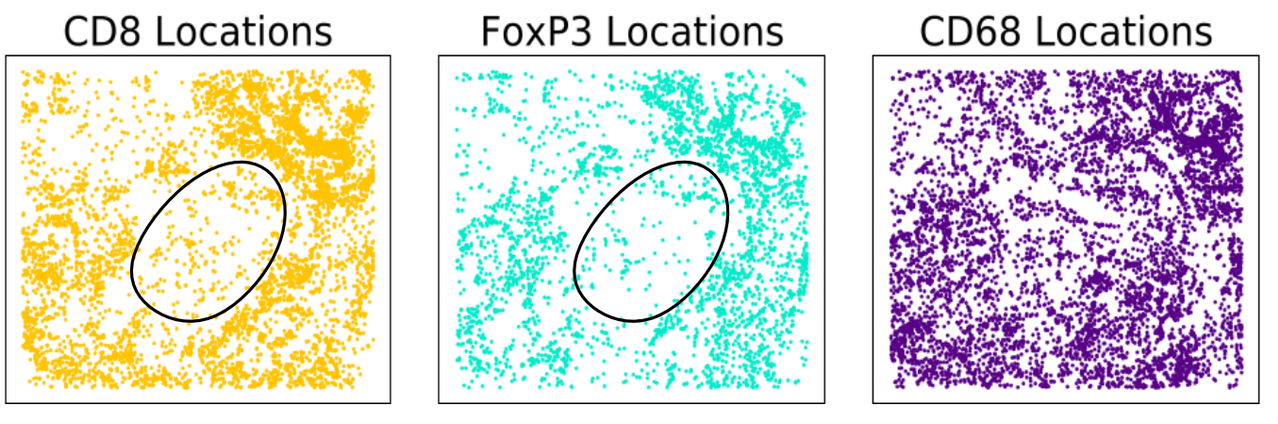
3. Multi-persistence for immune cell arrangements

Q: How can one characterize the *geometric arrangement* of cells *of a given cell type* in a way that is robust to *cell type misclassifications*?

Indeed, cell types are often obtained by integration with scRNA-seq and prone to errors...



Loop structure

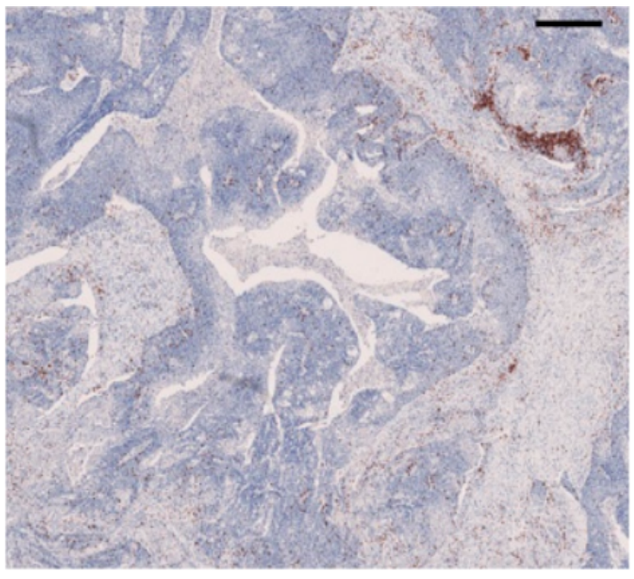


No loop structure

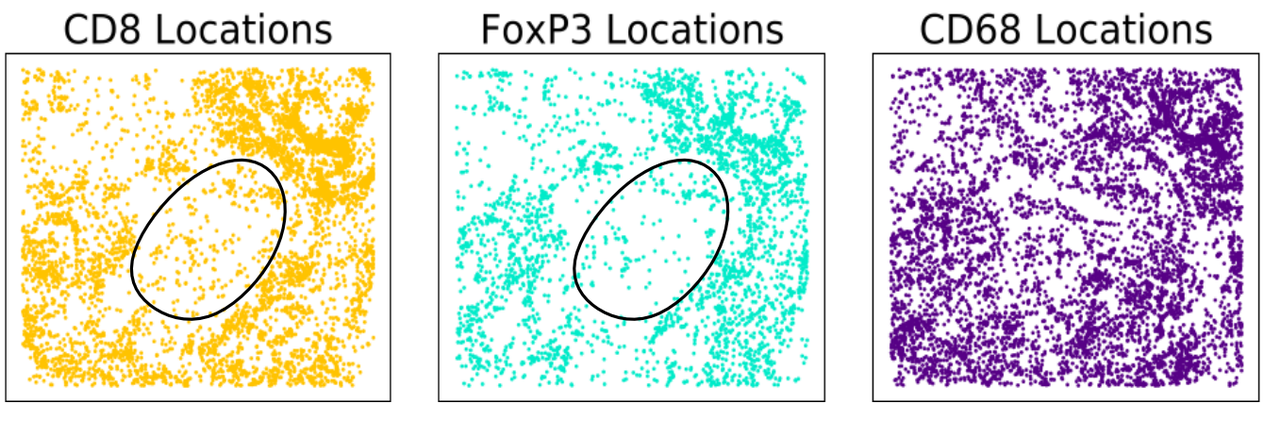
3. Multi-persistence for immune cell arrangements

Q: How can one characterize the *geometric arrangement* of cells *of a given cell type* in a way that is robust to *cell type misclassifications*?

Indeed, cell types are often obtained by integration with scRNA-seq and prone to errors...



Loop structure



No loop structure

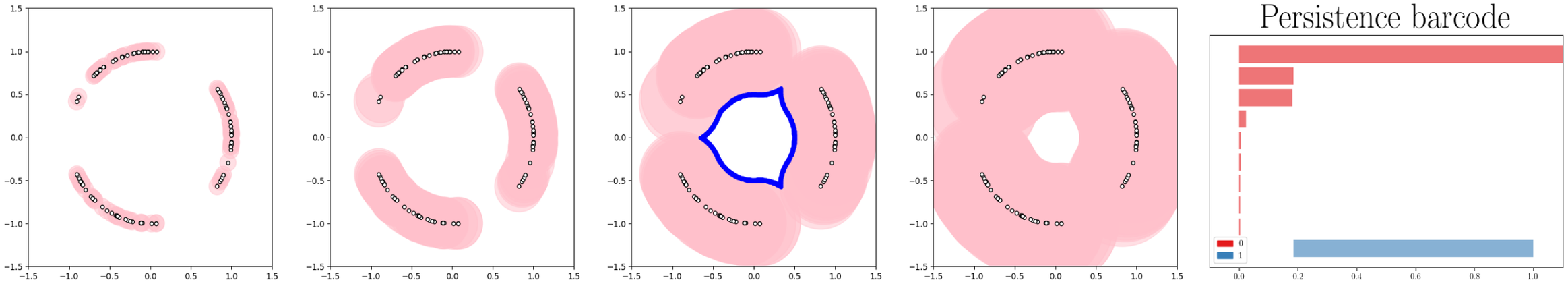
A: Combine *Rips filtration* with *density estimation* to create an outlier-robust analogue of persistence diagrams with 2-parameter PH.

Instability of Rips persistence

Rips persistence diagrams are known to be sensitive to outliers.

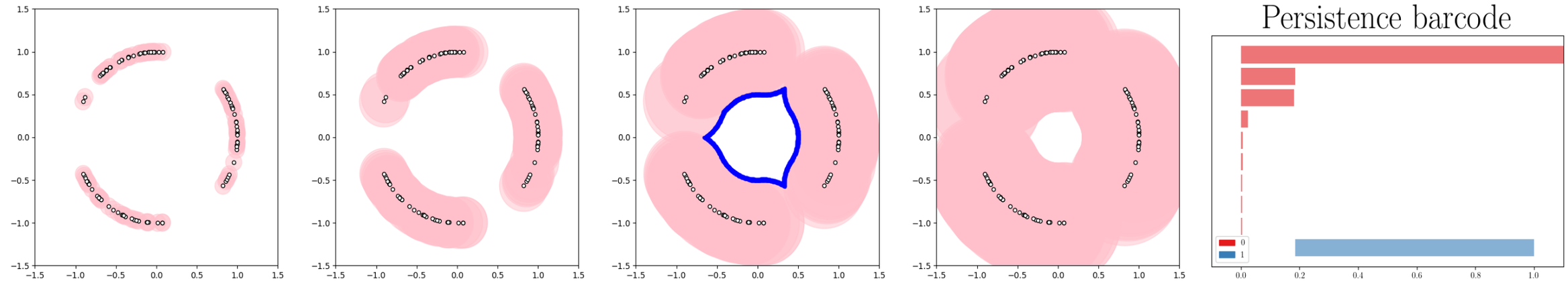
Instability of Rips persistence

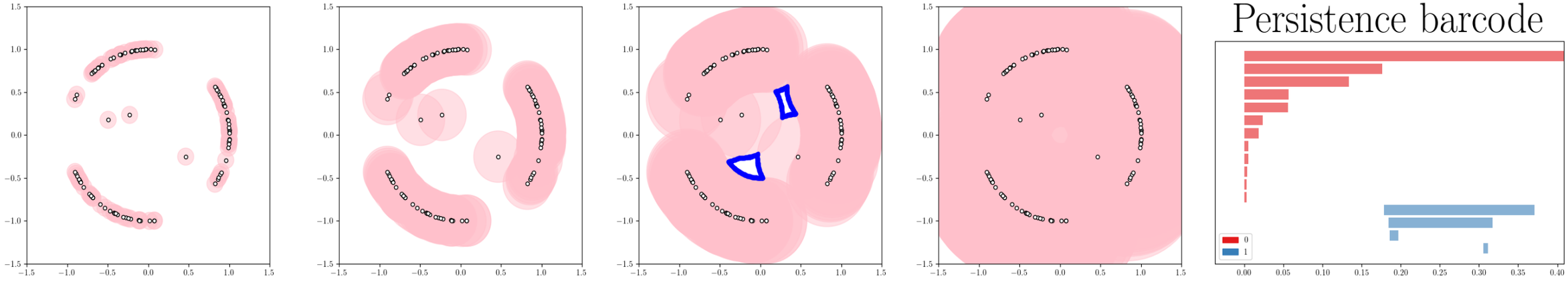
Rips persistence diagrams are known to be sensitive to outliers.



Instability of Rips persistence

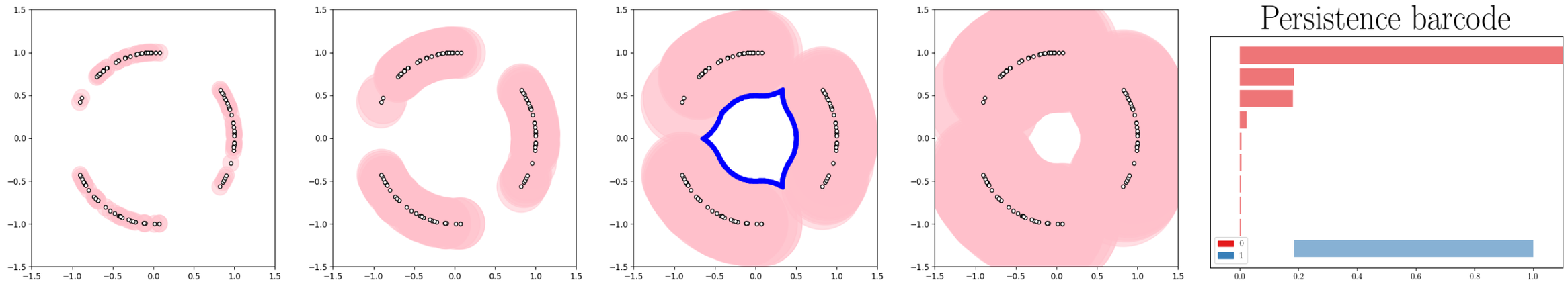
Rips persistence diagrams are known to be sensitive to outliers.

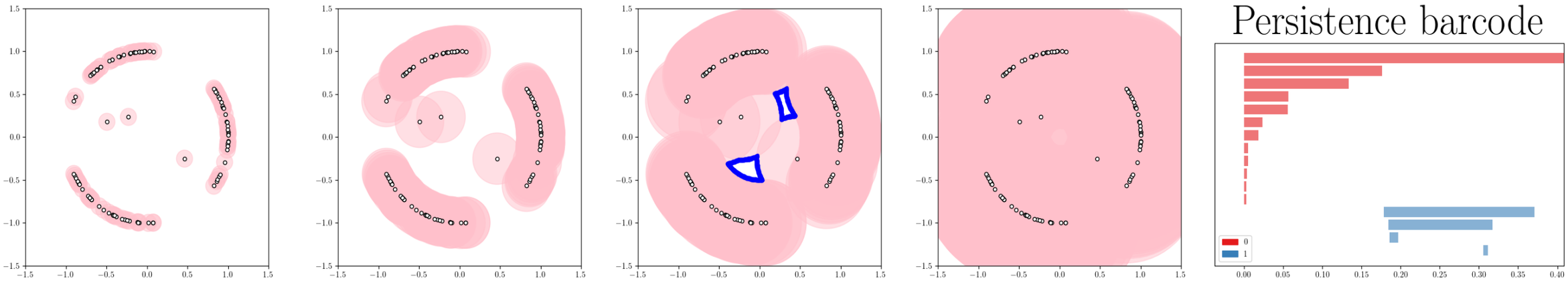




Instability of Rips persistence

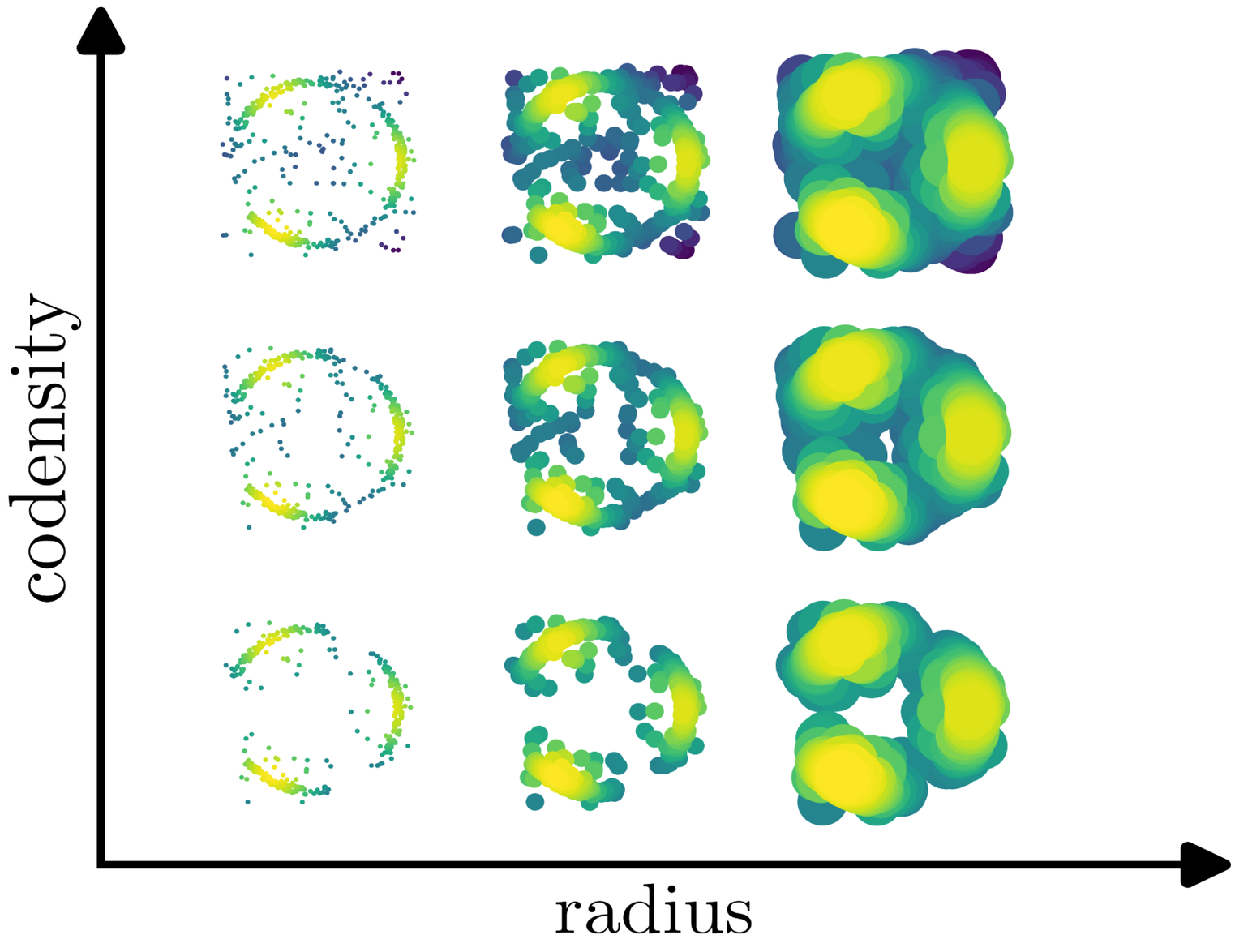
Rips persistence diagrams are known to be sensitive to outliers.





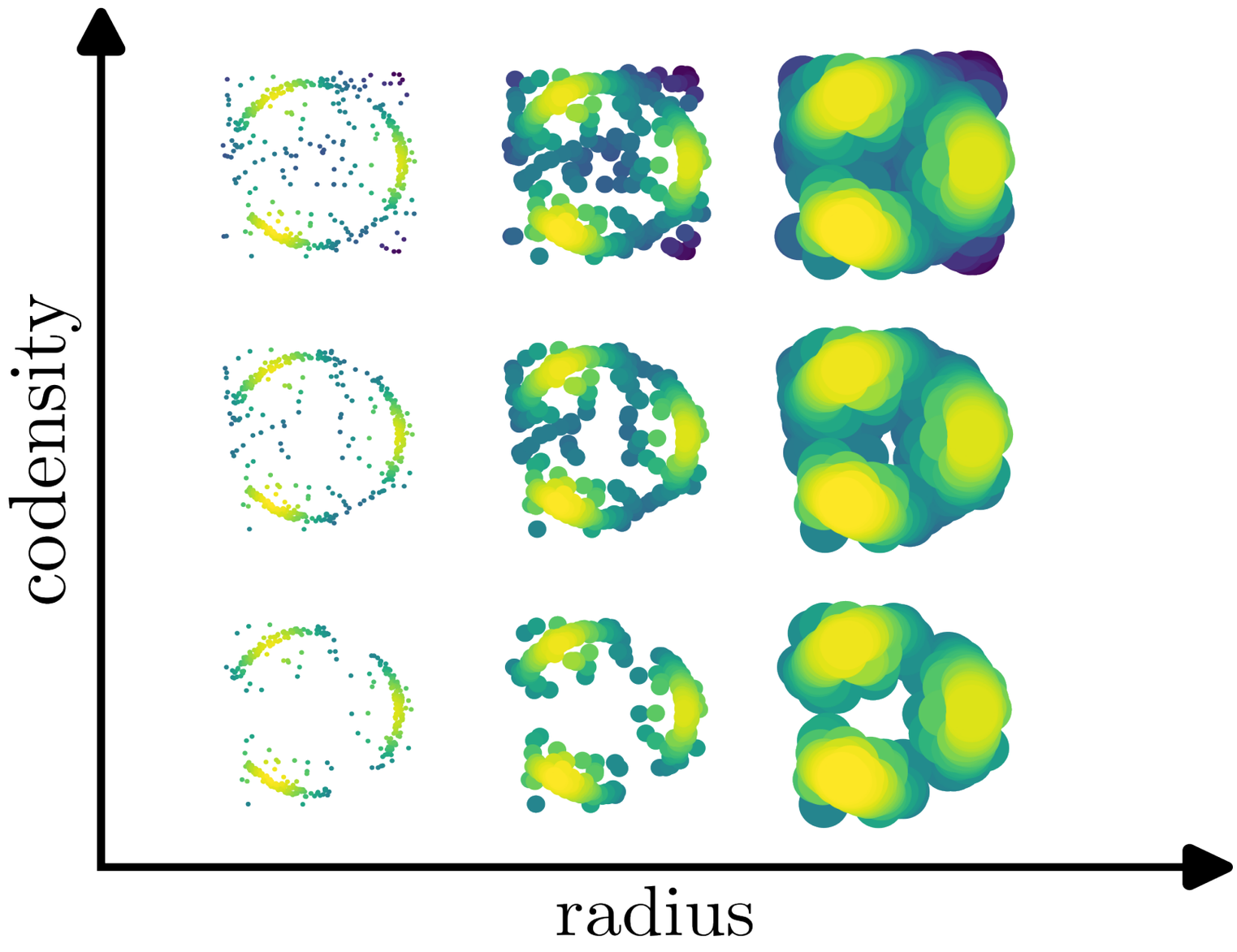
Hence, it seems natural to use density functions to prevent outliers from destroying topological structures.

2-parameter Persistent Homology



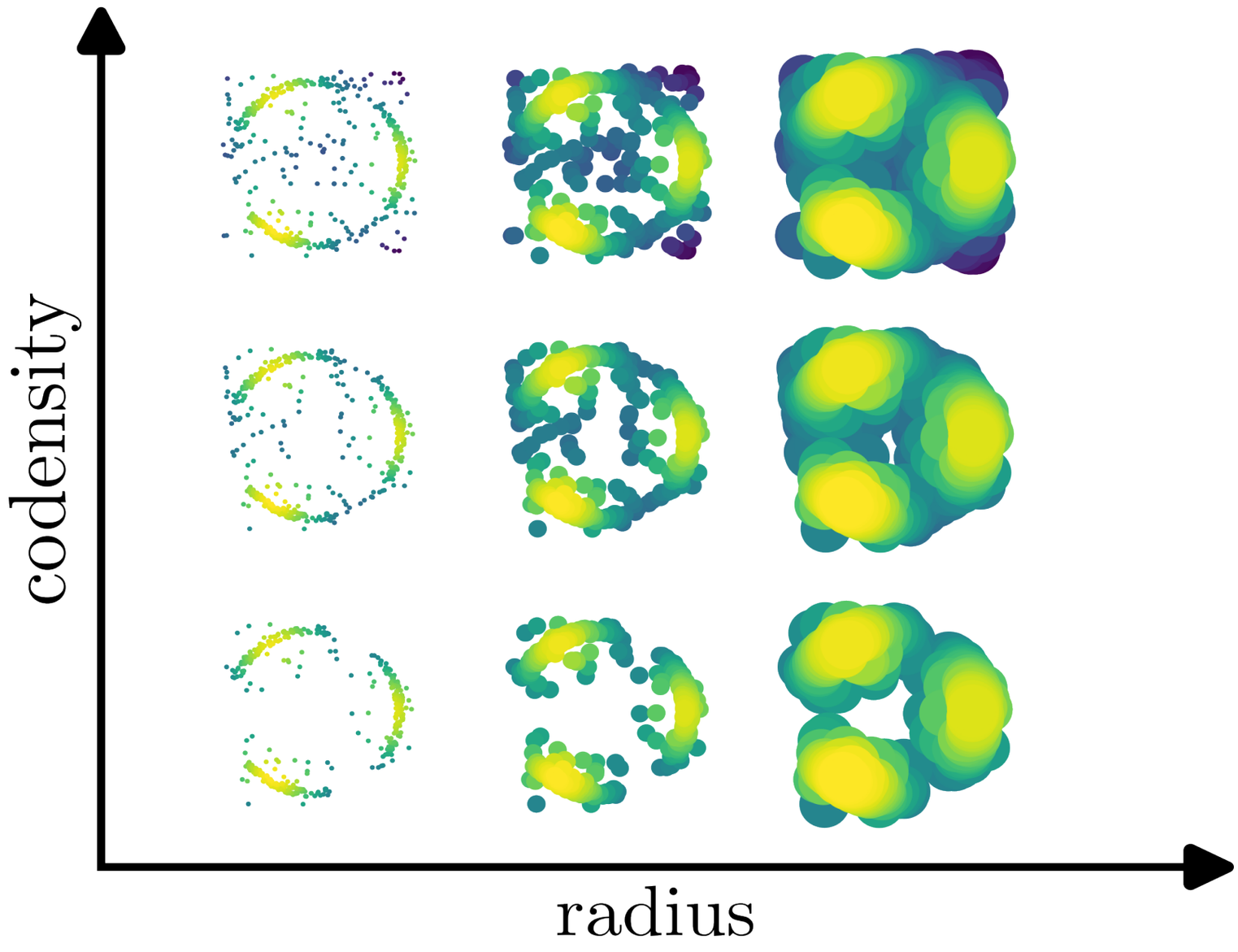
2-parameter Persistent Homology

Pb: persistence diagrams are not well-defined for more than 1 parameter.
(no decomposition theorem)



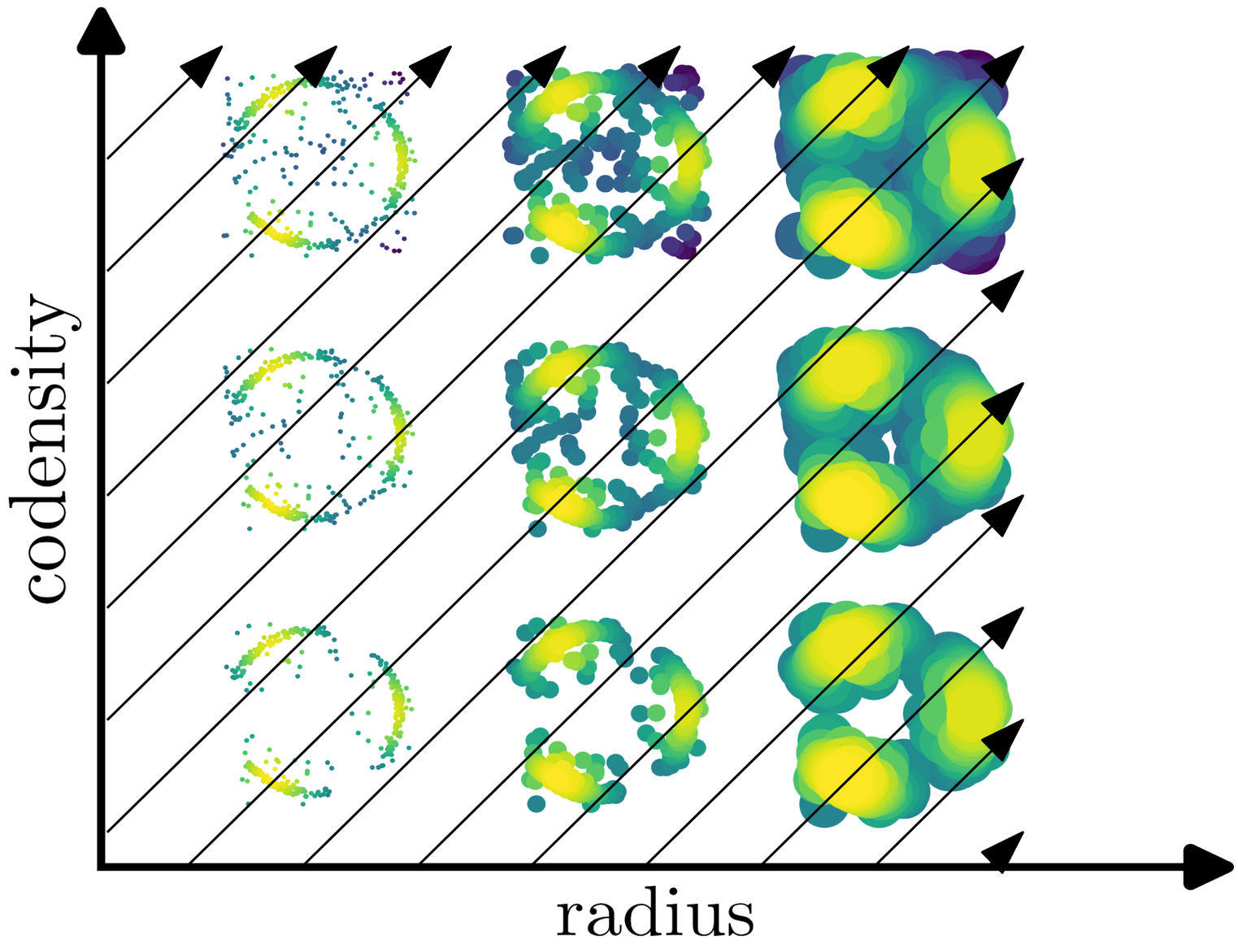
2-parameter Persistent Homology

A: Get back to the 1-parameter case by **slicing**.



2-parameter Persistent Homology

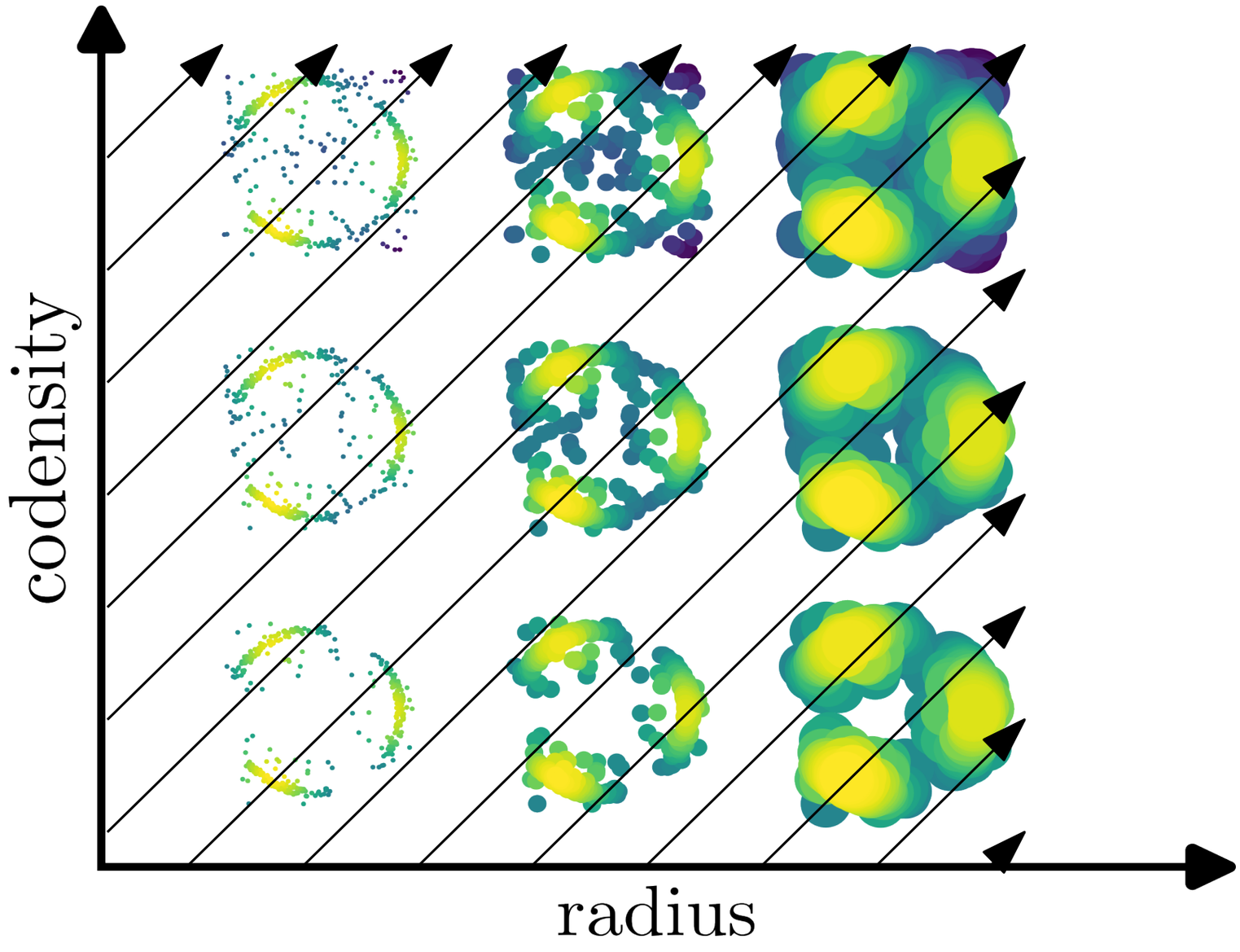
A: Get back to the 1-parameter case by **slicing**.



2-parameter Persistent Homology

A: Get back to the 1-parameter case by **slicing**.

Each line is a 1-parameter filtration obtained by linear combination, and has an associated persistence diagram. Hence one can represent 2-parameter PH by stacking the representations of the PDs from all lines.



2-parameter Persistent Homology

A: Get back to the 1-parameter case by **slicing**.

Each line is a 1-parameter filtration obtained by linear combination, and has an associated persistence diagram. Hence one can represent 2-parameter PH by stacking the representations of the PDs from all lines.

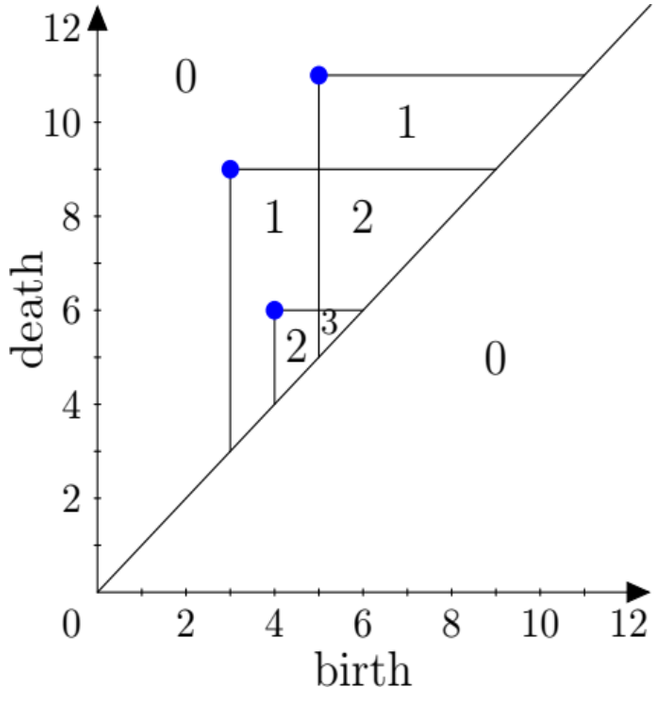
Ex: Persistence Landscapes.

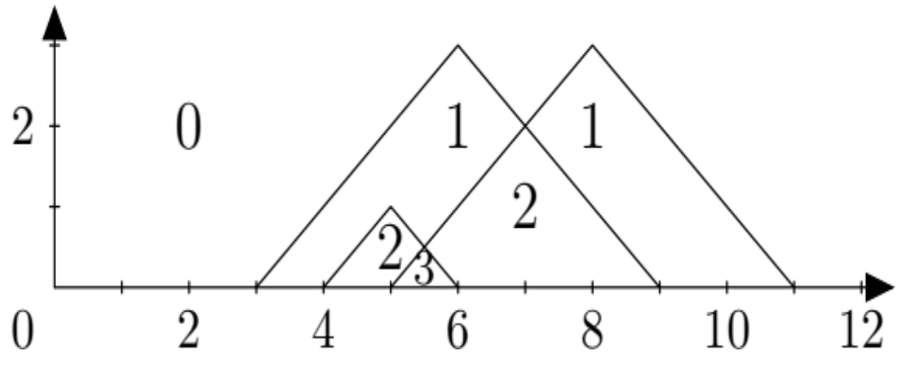
2-parameter Persistent Homology

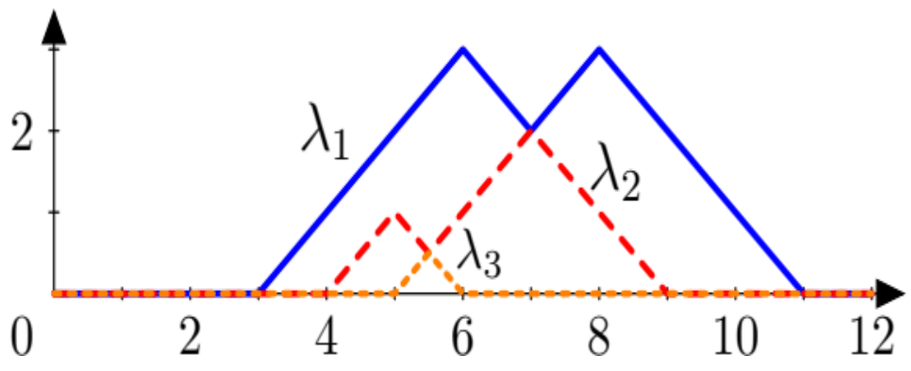
A: Get back to the 1-parameter case by **slicing**.

Each line is a 1-parameter filtration obtained by linear combination, and has an associated persistence diagram. Hence one can represent 2-parameter PH by stacking the representations of the PDs from all lines.

Ex: Persistence Landscapes.







Rotate PD
Compute rank function

Use boundaries of rank function

2-parameter Persistent Homology

A: Get back to the 1-parameter case by **slicing**.

Each line is a 1-parameter filtration obtained by linear combination, and has an associated persistence diagram. Hence one can represent 2-parameter PH by stacking the representations of the PDs from all lines.

Ex: Persistence Landscapes.

Ex: Multiparameter Persistence Landscapes.

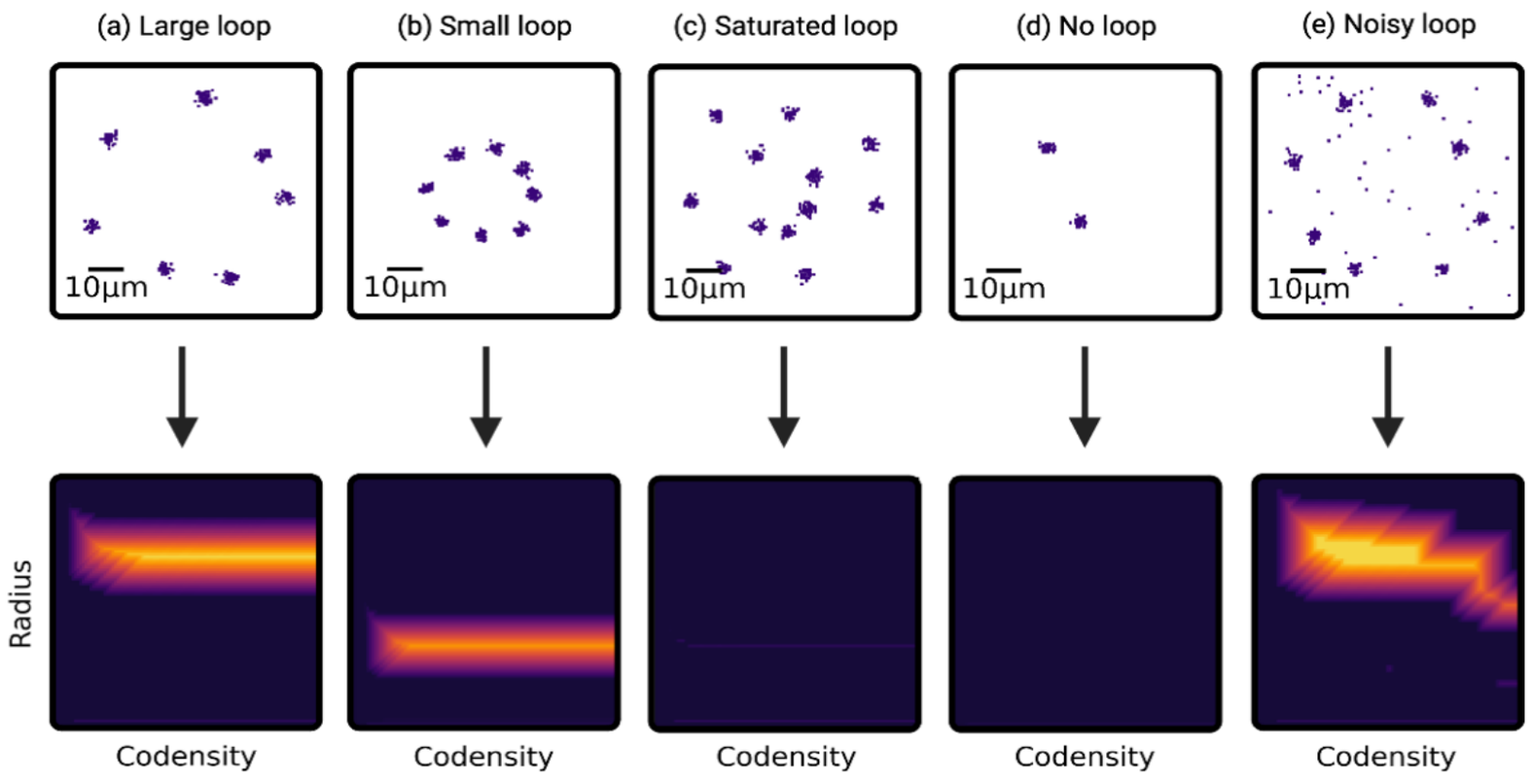
2-parameter Persistent Homology

A: Get back to the 1-parameter case by **slicing**.

Each line is a 1-parameter filtration obtained by linear combination, and has an associated persistence diagram. Hence one can represent 2-parameter PH by stacking the representations of the PDs from all lines.

Ex: Persistence Landscapes.

Ex: Multiparameter Persistence Landscapes.



n-parameter Persistent Homology

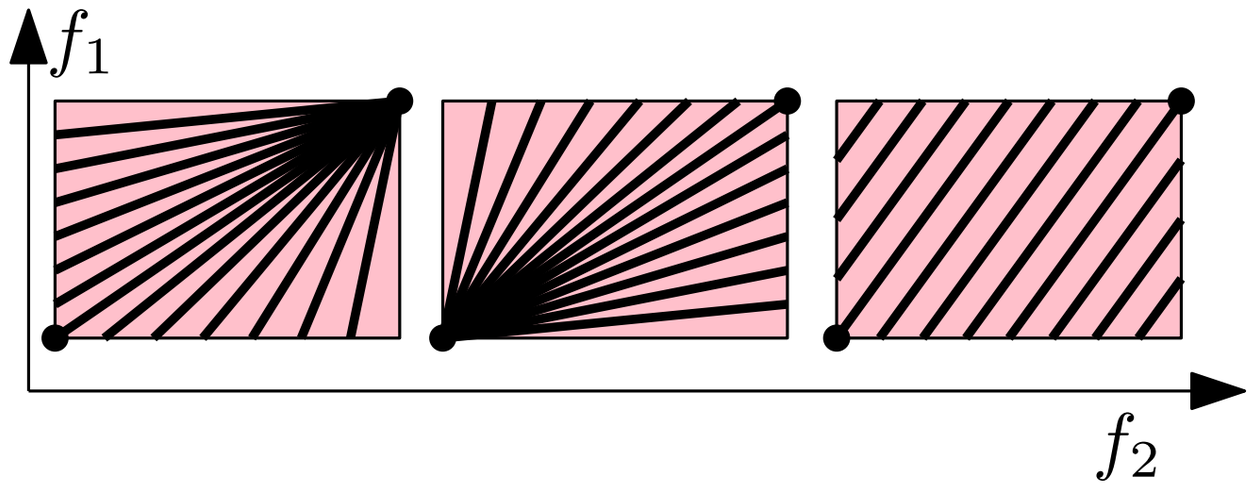
[Loiseaux et al. - 2022 - *Fast, stable and efficient approximation of multi-parameter persistence modules with MMA*]

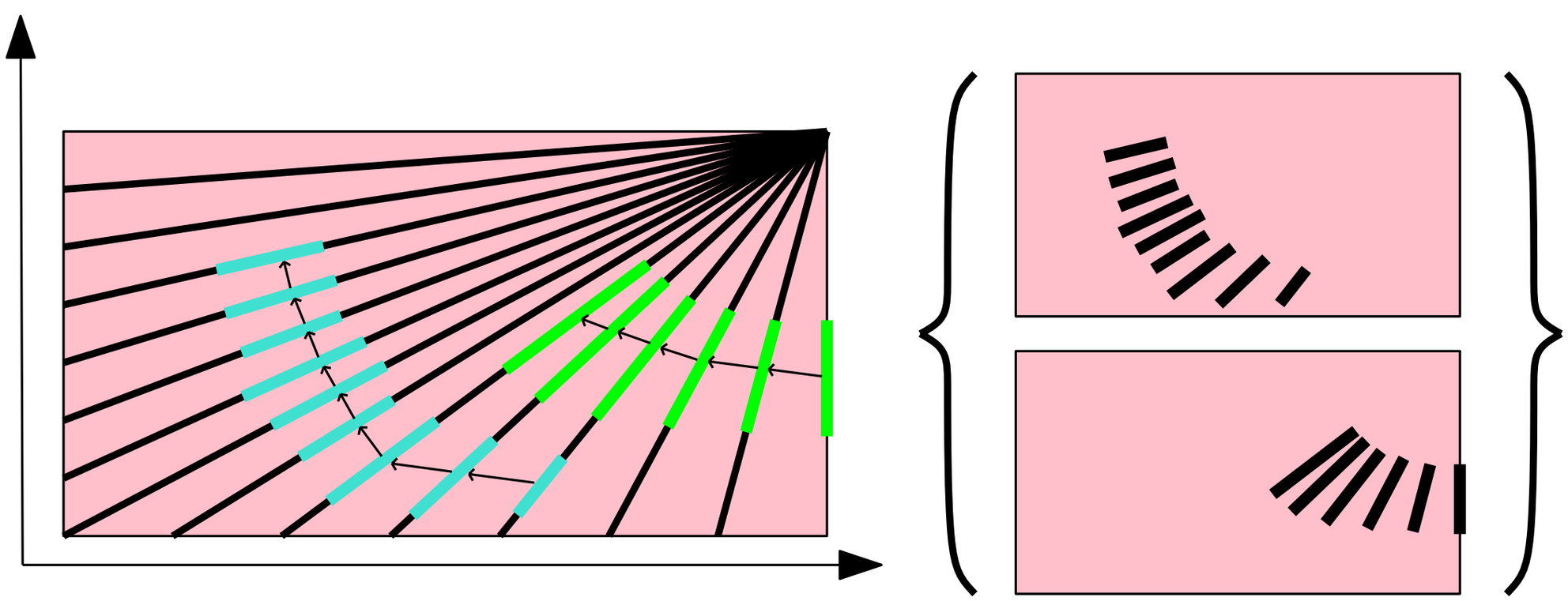
A: Get back to the 1-parameter case by **slicing + matching**.

n-parameter Persistent Homology

[Loiseaux et al. - 2022 - *Fast, stable and efficient approximation of multi-parameter persistence modules with MMA*]

A: Get back to the 1-parameter case by **slicing + matching**.



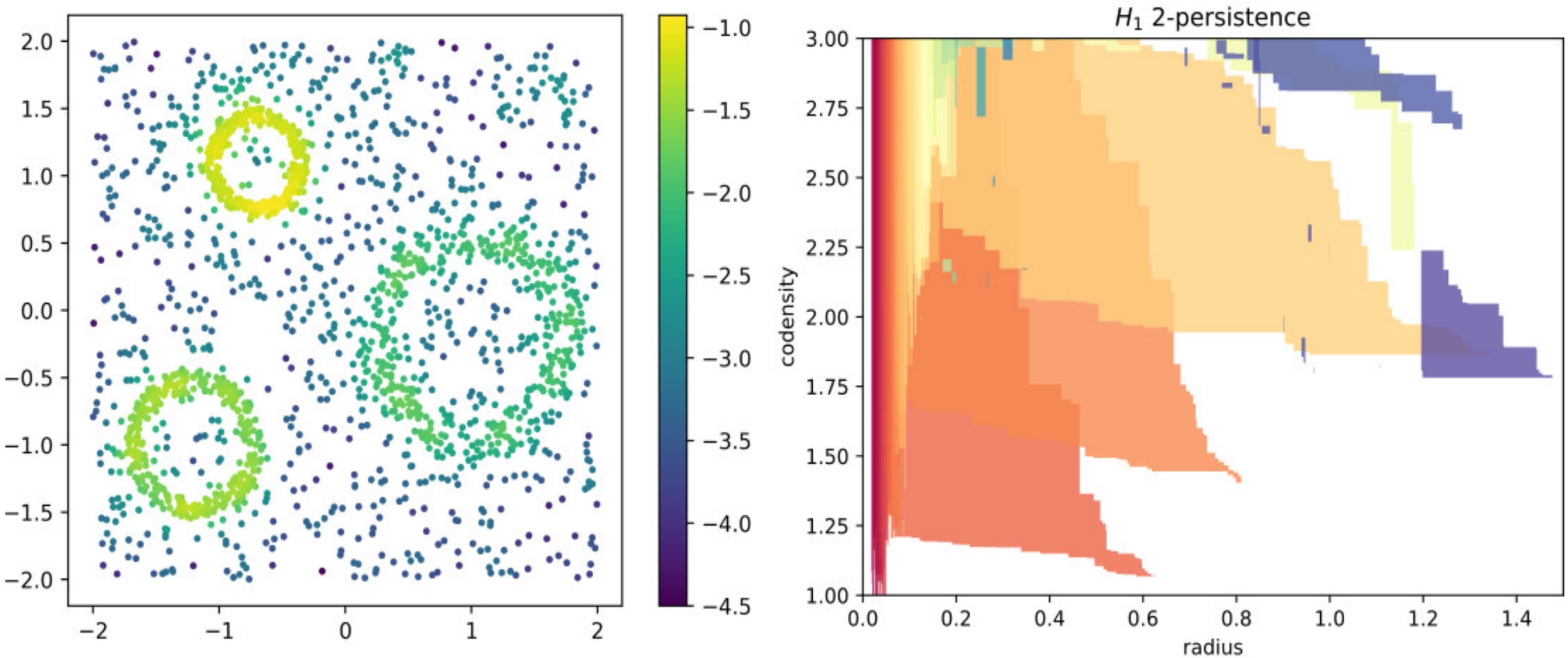


n-parameter Persistent Homology

[Loiseaux et al. - 2022 - *Fast, stable and efficient approximation of multi-parameter persistence modules with MMA*]

A: Get back to the 1-parameter case by **slicing + matching**.

Approximate persistence module



Thm: Let M be an *interval decomposable* multiparameter persistence module, and $\tilde{M}$ be its approximation with a family of lines at distance δ. Then:

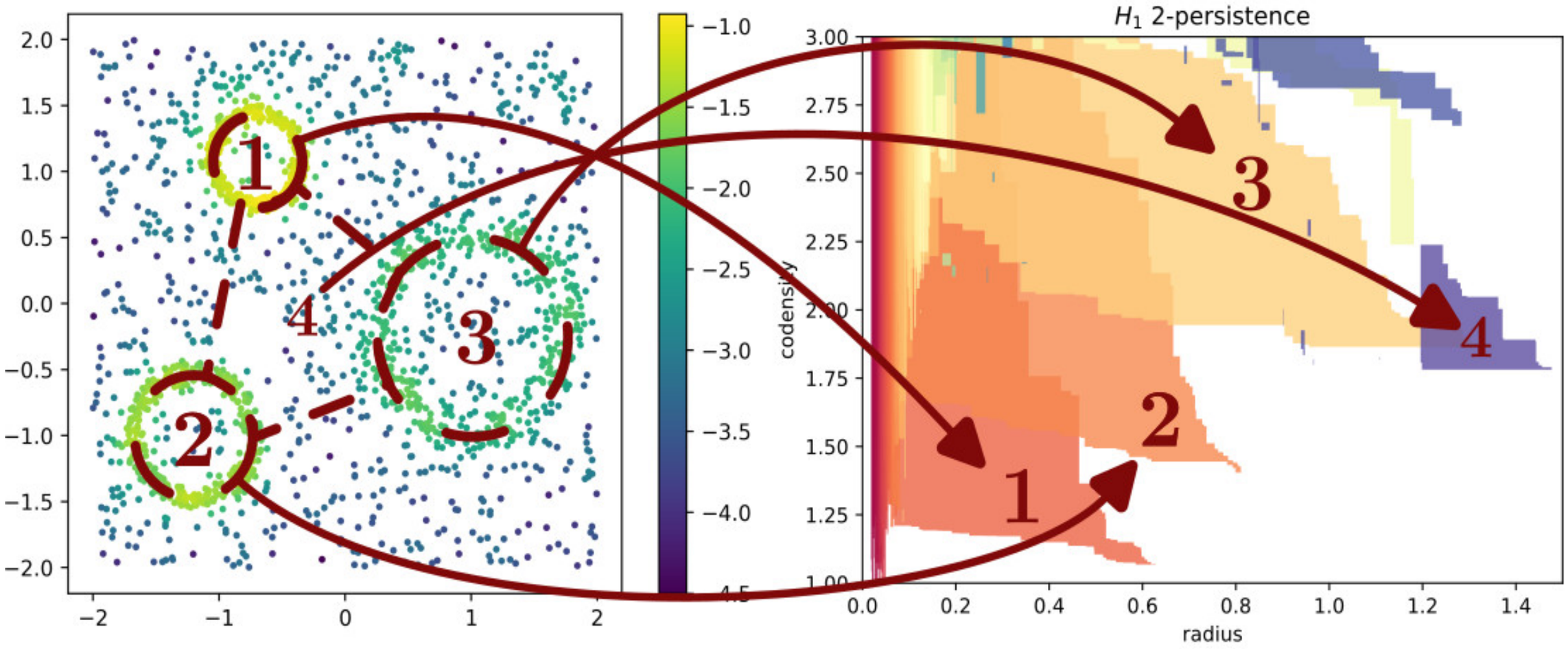
$$d_I(M, \tilde{M}) \leq \delta/2$$

n-parameter Persistent Homology

[Loiseaux et al. - 2022 - *Fast, stable and efficient approximation of multi-parameter persistence modules with MMA*]

A: Get back to the 1-parameter case by **slicing + matching**.

Approximate persistence module



Thm: Let M be an *interval decomposable* multiparameter persistence module, and $\tilde{M}$ be its approximation with a family of lines at distance δ. Then:

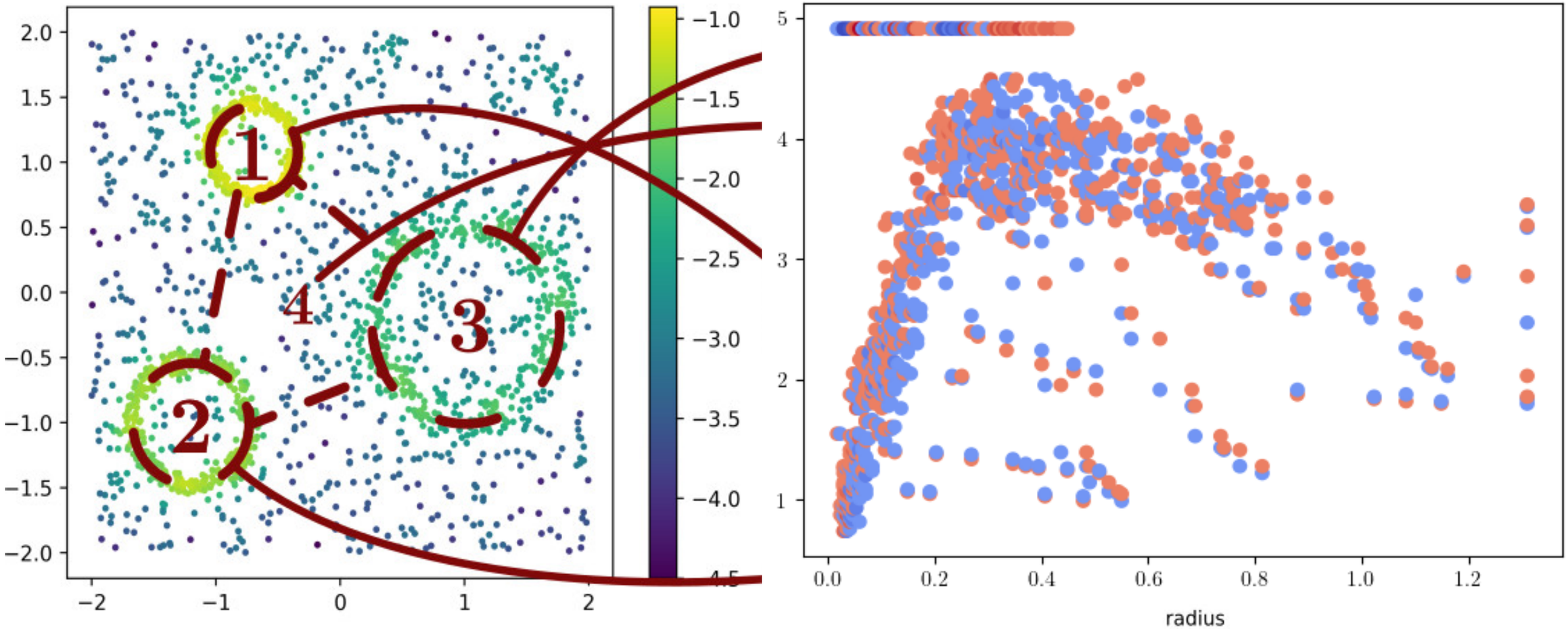
$$d_I(M, \tilde{M}) \leq \delta/2$$

n-parameter Persistent Homology

[Loiseaux et al. - 2022 - *Fast, stable and efficient approximation of multi-parameter persistence modules with MMA*]

A: Get back to the 1-parameter case by **slicing + matching**.

Representation with measure (of mass 0)



[Loiseaux et al. - 2023 - *Stable vectorization of multiparameter persistent homology using signed barcodes as measures*]

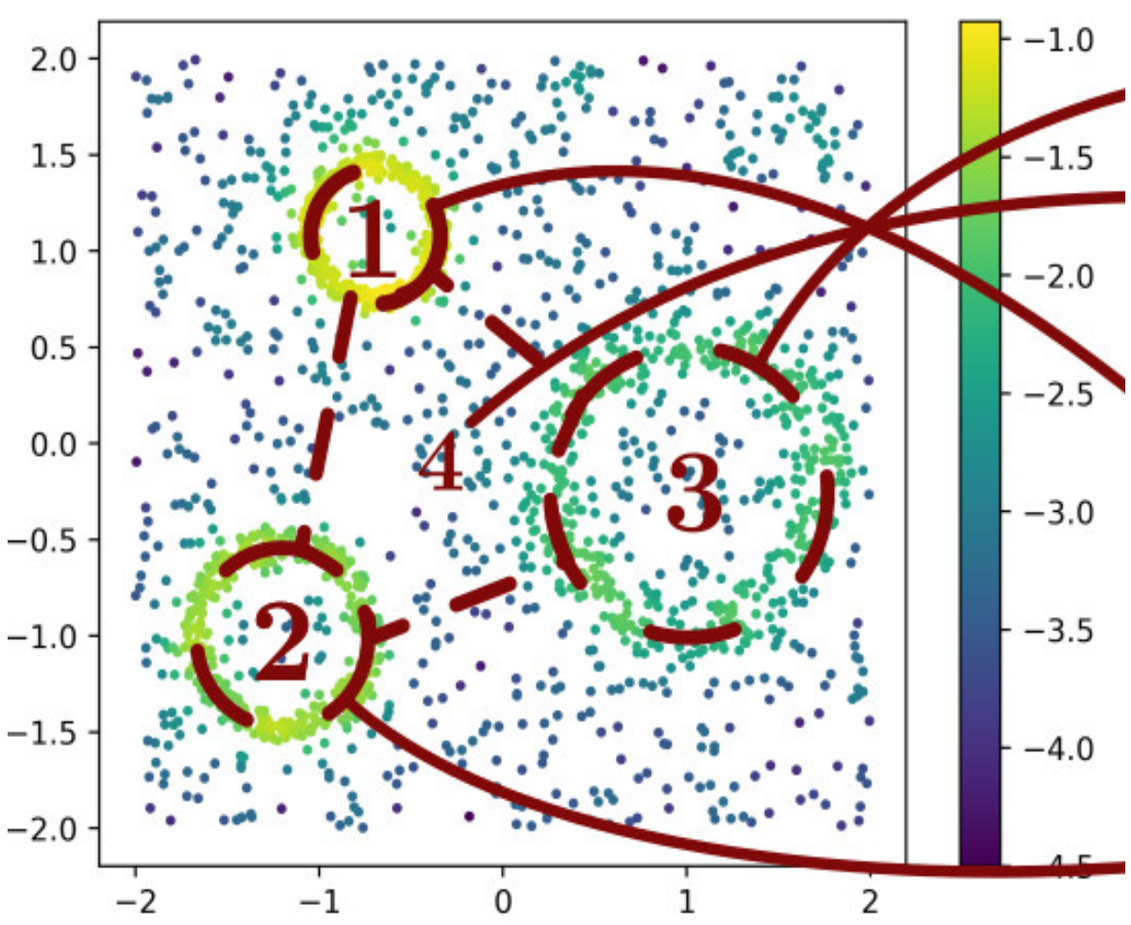
Thm: $\|\mu_f - \mu_g\|_1^K \leq \|f - g\|_1$

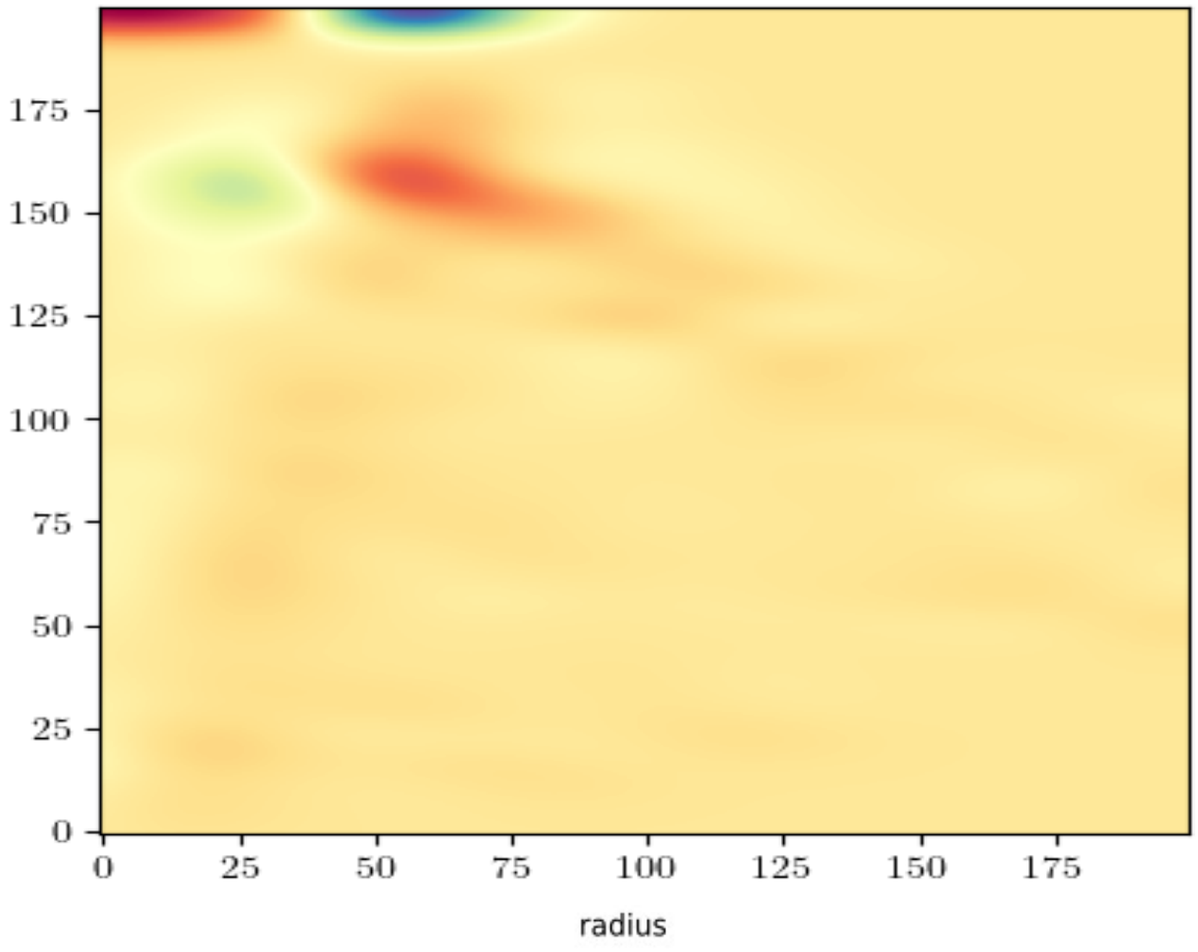
n-parameter Persistent Homology

[Loiseaux et al. - 2022 - *Fast, stable and efficient approximation of multi-parameter persistence modules with MMA*]

A: Get back to the 1-parameter case by **slicing + matching**.

Representation with measure (of mass 0)





[Loiseaux et al. - 2023 - *Stable vectorization of multiparameter persistent homology using signed barcodes as measures*]

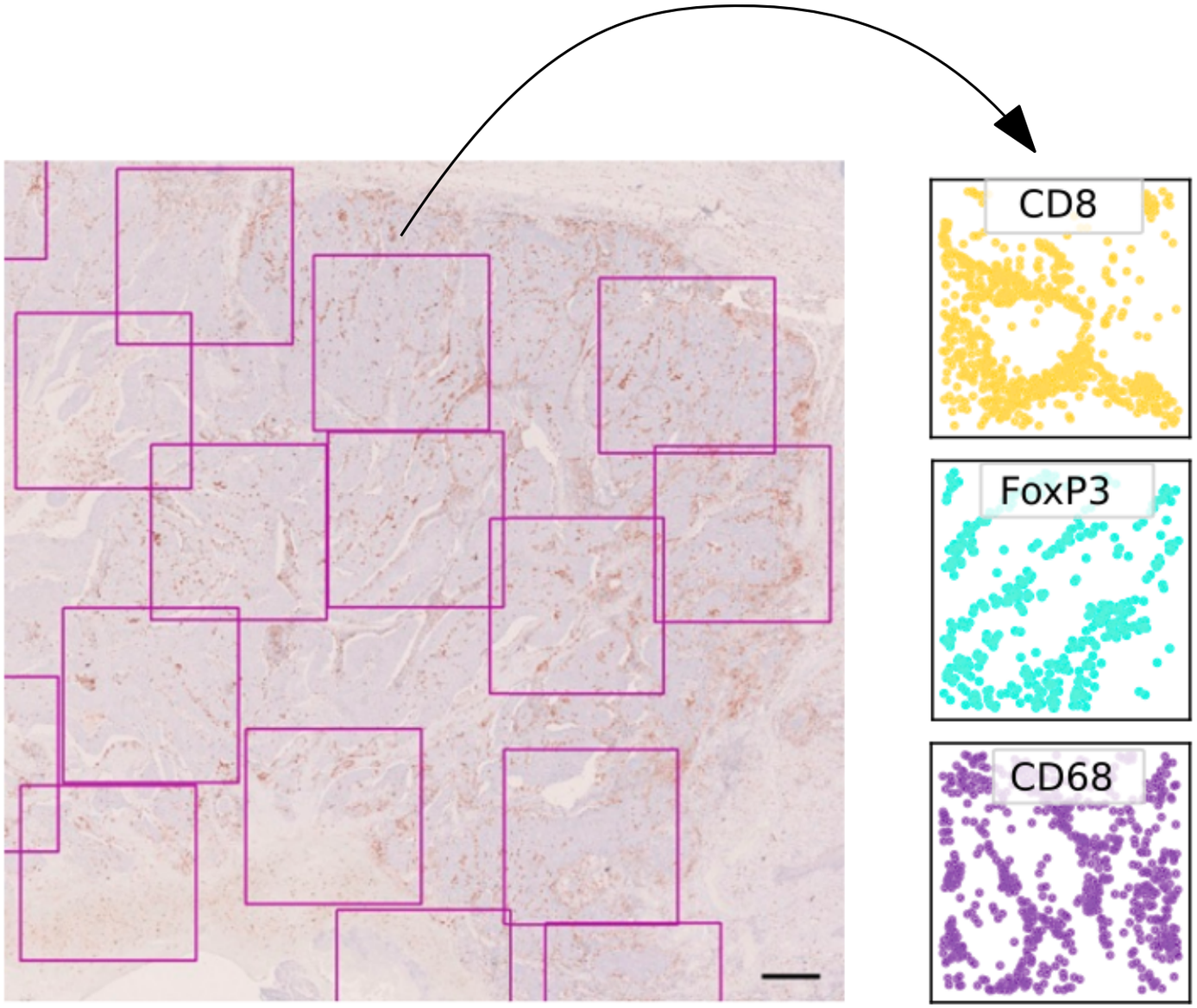
Thm: $\|\mu_f - \mu_g\|_1^K \leq \|f - g\|_1$

Thm: $\|K * \mu_f - K * \mu_g\|_2 \leq \|\mu_f - \mu_g\|_2^K$

Application to cell arrangements

Method:

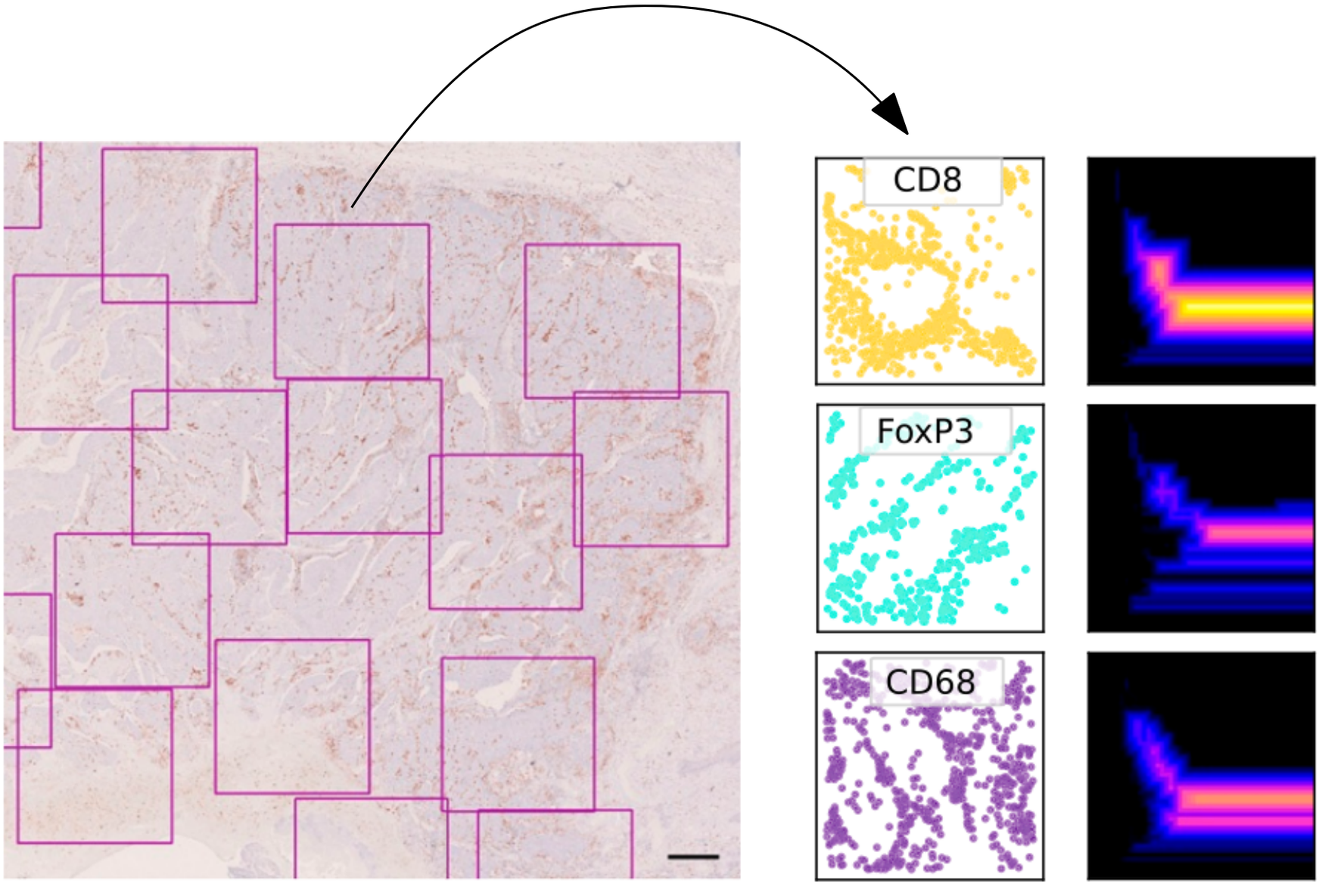
1. Extract local point clouds and cell types corresponding to several measurement spots.



Application to cell arrangements

Method:

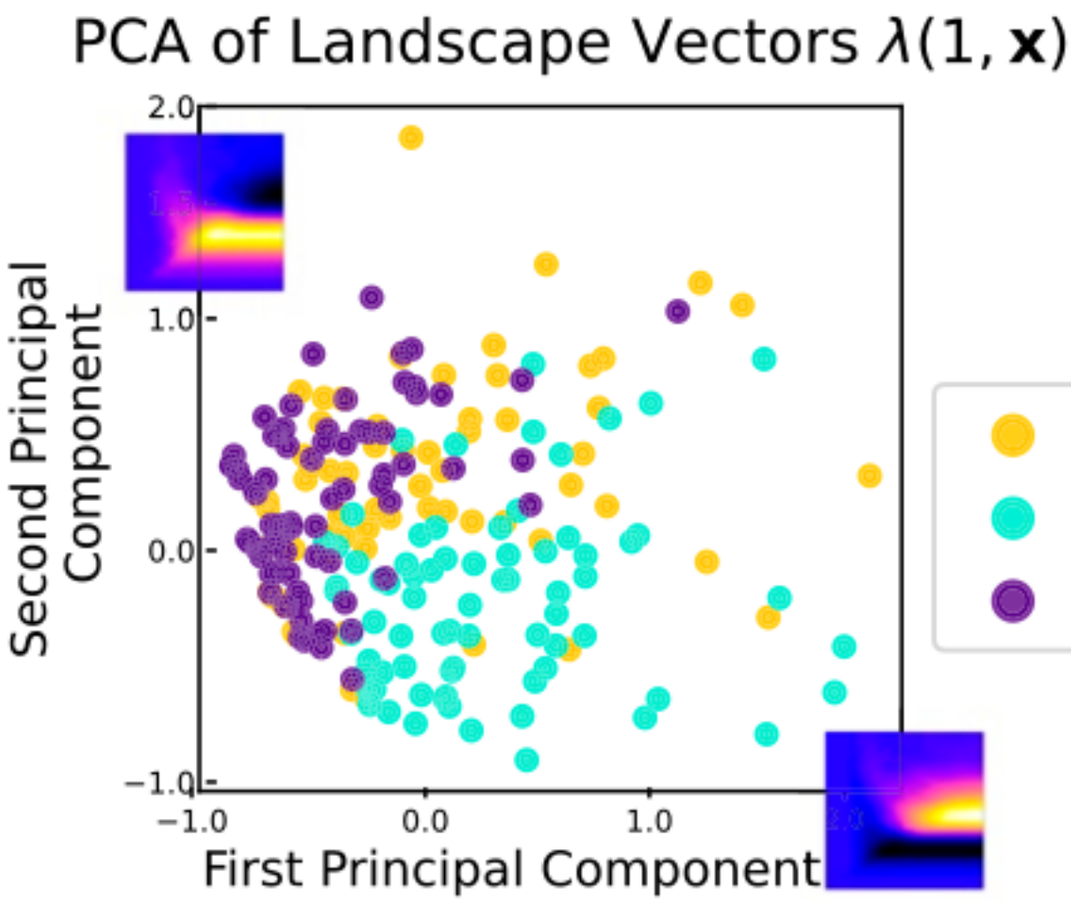
1. Extract local point clouds and cell types corresponding to several measurement spots.
2. Compute their multi-parameter persistence landscapes (MPLs) in dimension 1.

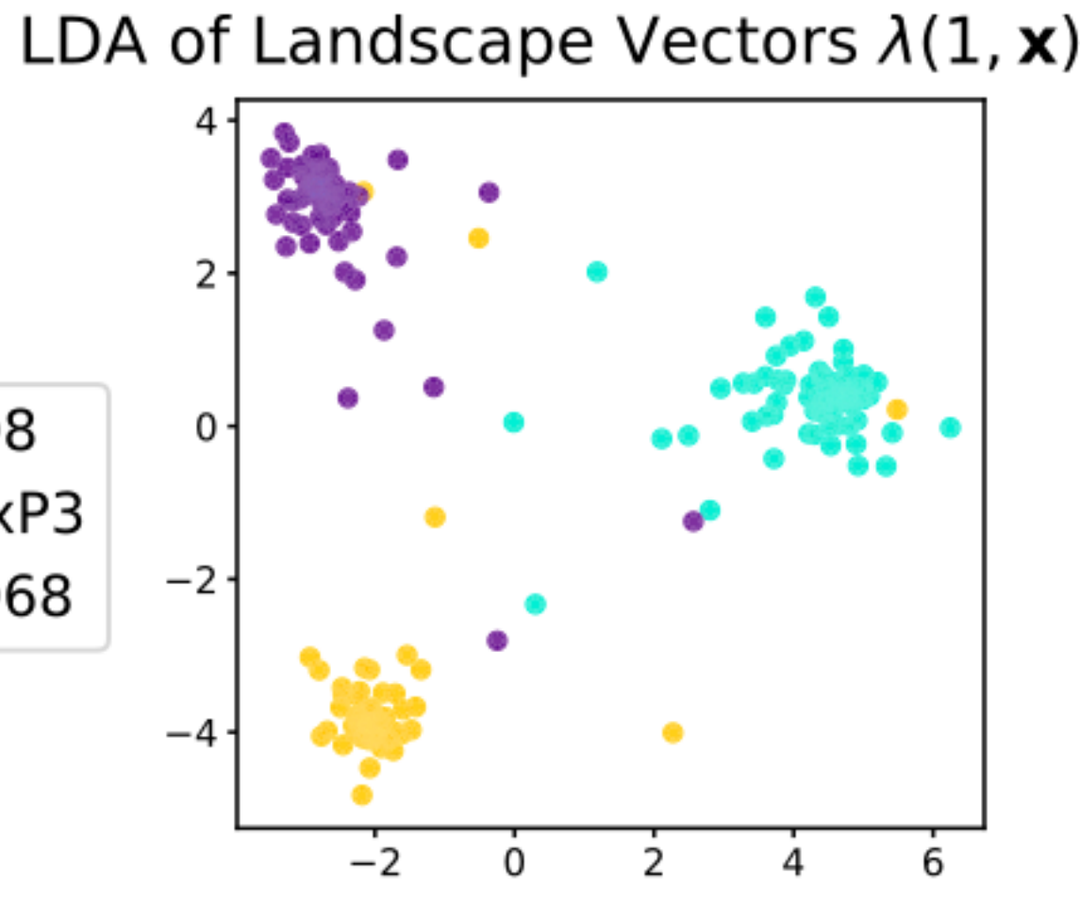


Application to cell arrangements

Method:

1. Extract local point clouds and cell types corresponding to several measurement spots.
2. Compute their multi-parameter persistence landscapes (MPLs) in dimension 1.
3. Visualize and predict cell types with MPLs.

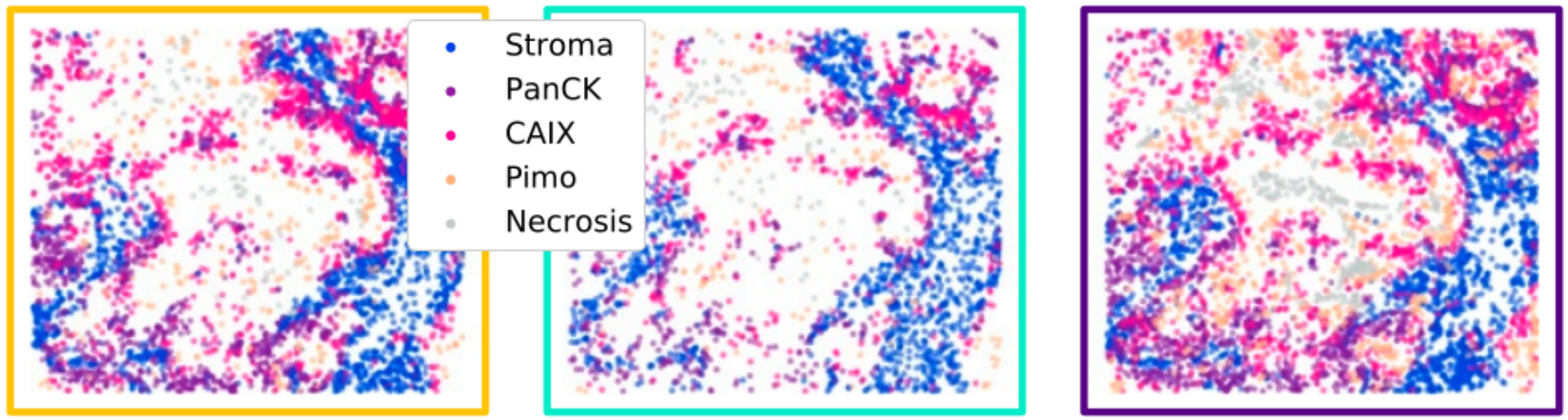




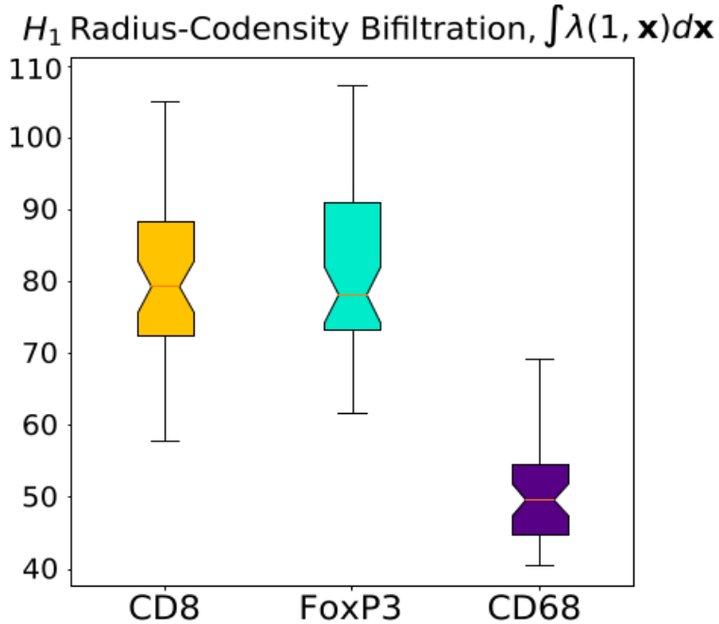
Application to cell arrangements

Method:

1. Extract local point clouds and cell types corresponding to several measurement spots.
2. Compute their multi-parameter persistence landscapes (MPLs) in dimension 1.
3. Visualize and predict cell types with MPLs.
4. Perform biological inference from additional cell annotations.



The fact that the norm of CD68 MPL is smaller than the others means that CD68 cells have a bigger tendancy to infiltrate necrosis cells.



Application to cell arrangements

Method:

1. Extract local point clouds and cell types corresponding to several measurement spots.
2. Compute their multi-parameter persistence landscapes (MPLs) in dimension 1.
3. Visualize and predict cell types with MPLs.
4. Perform biological inference from additional cell annotations.

MPLs were also used recently for studying immune responses in kidney glomeruli from mice with renal disease: a high MPL norm (in dimension 1) was shown to indicate high immune activity, exhibiting ring structures as spatial patterns.

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Plan of the course

1. ToMATo for colocalizing cell types

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

2. Rips persistence for marker gene correlations

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

3. Multi-persistence for immune cell arrangements

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

4. Future research directions

4. Future research directions

Summary of what we have seen:

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

4. Future research directions

Summary of what we have seen:

Limitation: too many marker genes for running multi-persistence, which forces to pick (combinations of) maker genes *a priori*

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

Limitation: 2-parameter PH is robust to outliers, but
(a) Outliers are less and less likely to happen due to progress in integration techniques, thus density could end up *removing accurate cells*.
(b) What about cells for which integration method is unsure (low confidence in the predicted cell types)? This could even happen systematically in some regions with high density.

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

To handle these, one could use the integration method confidence to process (Spot in $\mathbb{R}^2$, Cell Type **Probabilities**) pairs with TDA!

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

To handle these, one could use the integration method confidence to process (Spot in $\mathbb{R}^2$, Cell Type **Probabilities**) pairs with TDA!

$\longrightarrow$ no need to cherry pick gene combinations

$\longrightarrow$ robust to outliers *and* integration methods

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

To handle these, one could use the integration method confidence to process (Spot in $\mathbb{R}^2$, Cell Type **Probabilities**) pairs with TDA!

Moreover, descriptors from n-parameter PH (approximate decompositions of persistence modules, signed barcodes, multivariate Mappers) could be used.

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

Other recent (August 2023) directions on synthetic data:

[Stolz et al. - 2023 - *Relational persistent homology for multi-species data with application to the tumor microenvironment*]

[Yang et al. - 2023 - *Topological classification of tumour-immune interactions and dynamics*]

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

Other recent (August 2023) directions on synthetic data:

[Stolz et al. - 2023 - *Relational persistent homology for multi-species data with application to the tumor microenvironment*]

[Yang et al. - 2023 - *Topological classification of tumour-immune interactions and dynamics*]

→ Characterize interactions between cell types with PH of *Dowker and Witness complexes*

4. Future research directions

Summary of what we have seen:

TDA + Raw Spatial Transcriptomics

[Bae et al. - 2022 - *STopover captures spatial colocalization and interaction in the tumor microenvironment using topological analysis in spatial transcriptomics data*]

[Aukerman et al. - 2022 - *Persistent homology based characterization of the breast cancer immune microenvironment: a feasibility study*]

Process (Spot in $\mathbb{R}^2$, Expression vector) pairs

[Alsaleh et al. - 2022 - *Spatial transcriptomic analysis reveals associations between genes and cellular topology in breast and prostate cancers*]

Process (Local point cloud $\subset \mathbb{R}^2$, Expression vector) pairs

TDA + Integrated (w/ scRNA-seq) Spatial Transcriptomics

[Vipond et al. - 2021 - *Multiparameter persistent homology landscapes identify immune cell spatial patterns in tumors*]

[Benjamin et al. - 2022 - *Multiscale topology classifies and quantifies cell types in subcellular spatial transcriptomics*]

Process (Spot in $\mathbb{R}^2$, Cell Type) pairs

Other recent (August 2023) directions on synthetic data:

[Stolz et al. - 2023 - *Relational persistent homology for multi-species data with application to the tumor microenvironment*]

[Yang et al. - 2023 - *Topological classification of tumour-immune interactions and dynamics*]

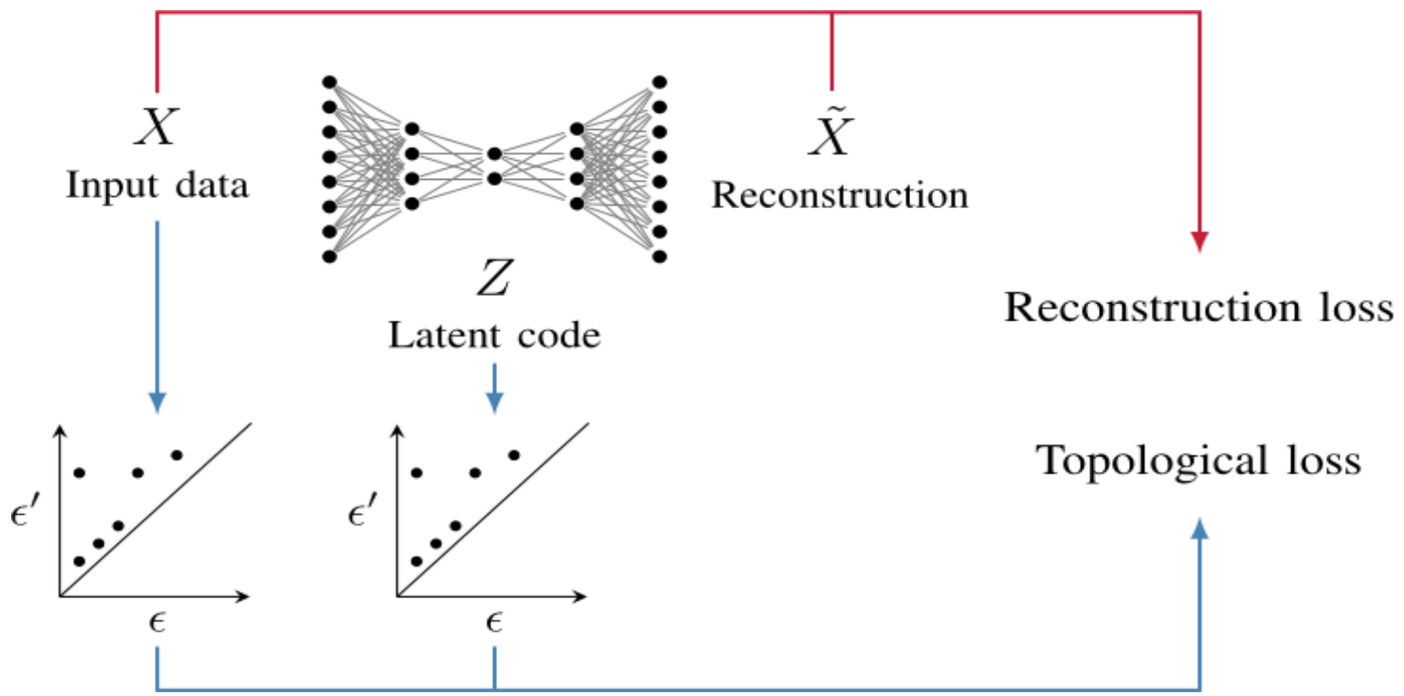
→ Characterize *time-dependent* spatial transcriptomics with *Zigzag and Vineyard* Rips PH

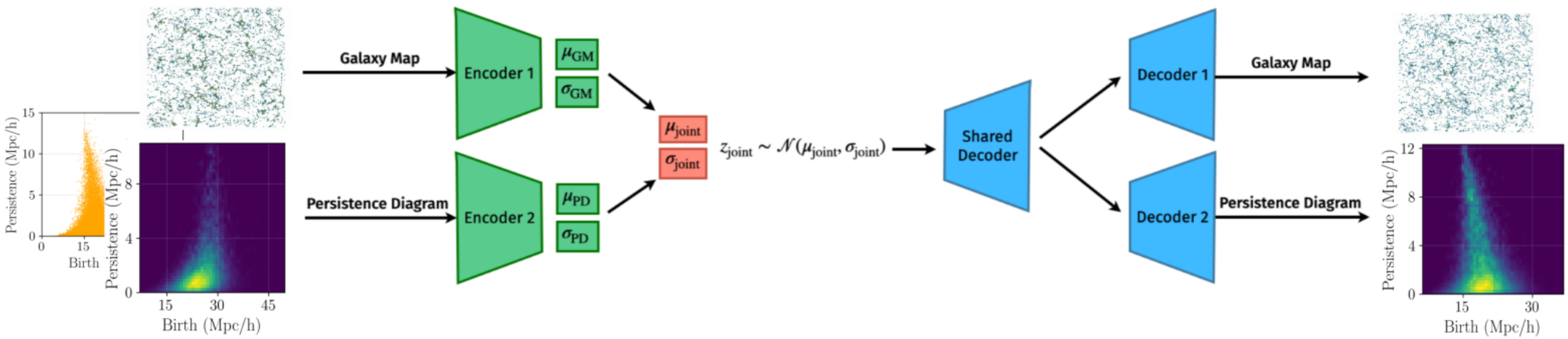
Other Topics: Differentiating PDs

Define gradient of persistence with

$$\forall p = (f(x_+), f(x_-)),\ \nabla p = [\nabla f(x_+), \nabla f(x_-)]$$

`https://github.com/GUDHI/TDA-tutorial/blob/master/Tuto-GUDHI-optimization.ipynb`

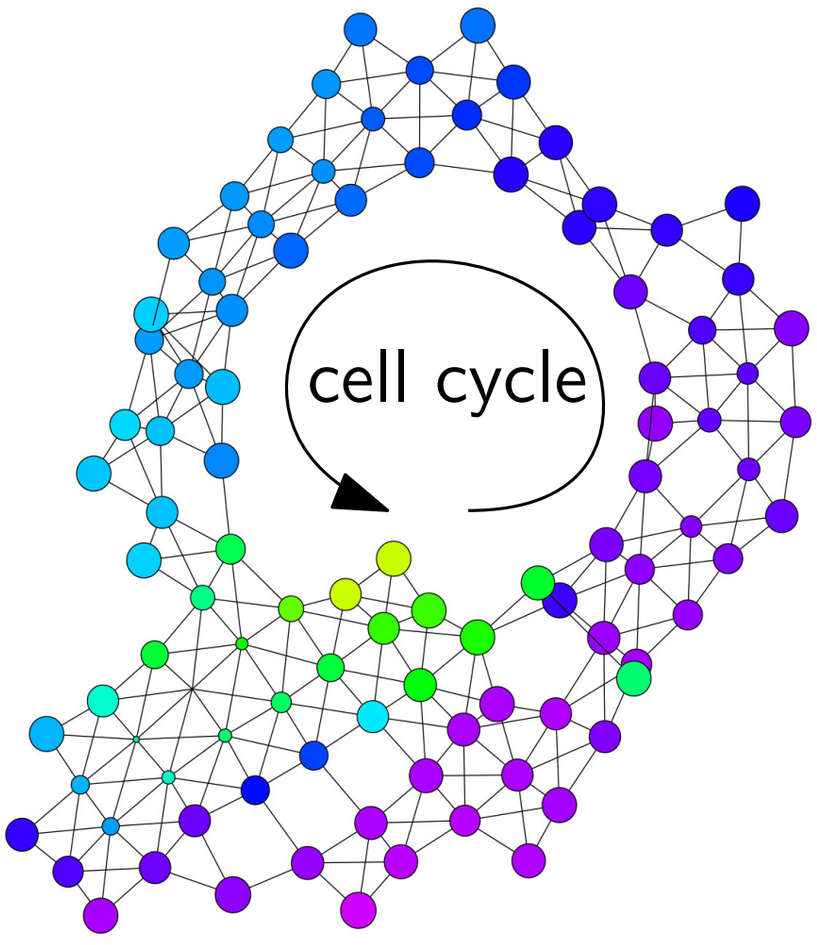




Other Topics: Exploratory Data Analysis with Mapper

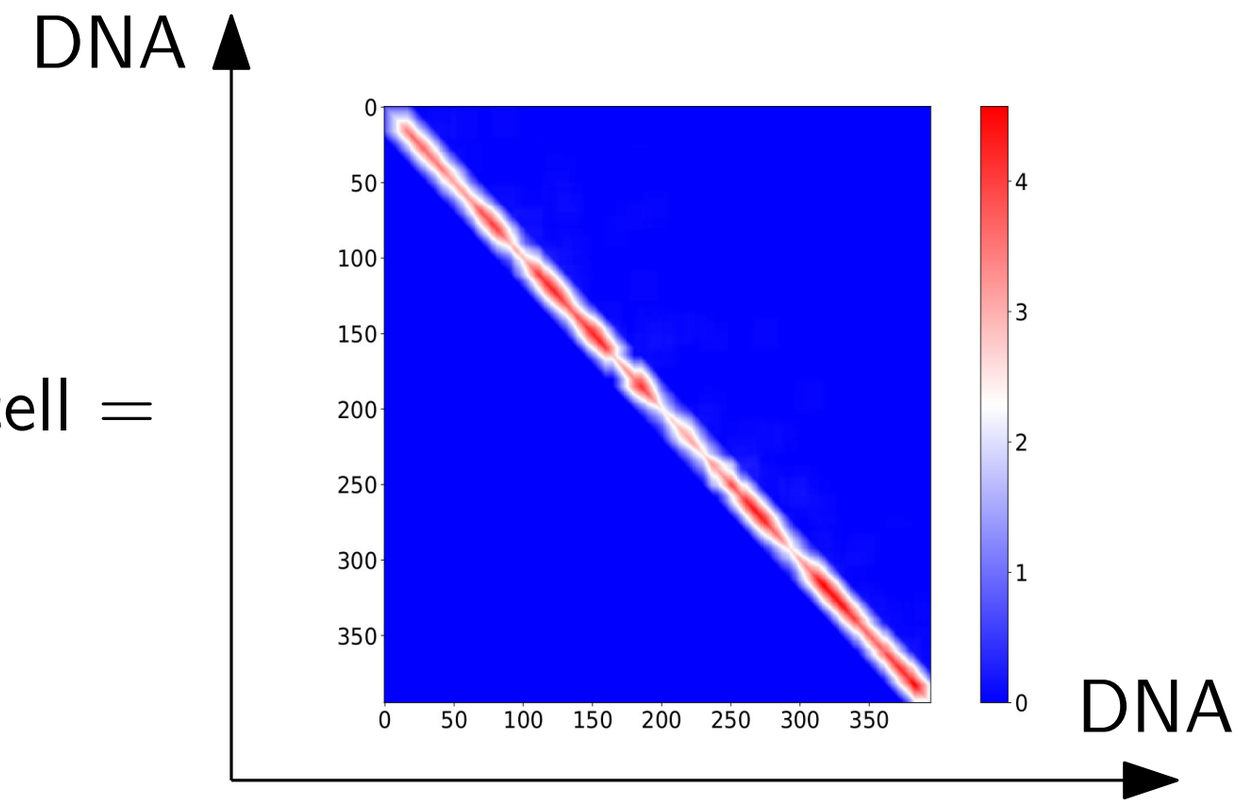
Mapper (and other cover complexes) allow to
represent and visualize datasets

`https://github.com/GUDHI/TDA-tutorial/blob/master/Tuto-GUDHI-cover-complex.ipynb`



Hi-C datasets =
Count matrices of *physically close* DNA fragments

cell =



Thank you!