Information Maximizing Persistent Homology

Accurate Fisher forecasts for persistent summaries.

Alex Cole
DE Shaw

Matteo Biagetti
AREA Trieste

Karthik Viswanathan
U. Amsterdam

Jacky Yip
U. Wisconsin-Madison

J.P. van der Schaar
U. Amsterdam

Overview

Information Maximizing Persistent Homology (IMPH)

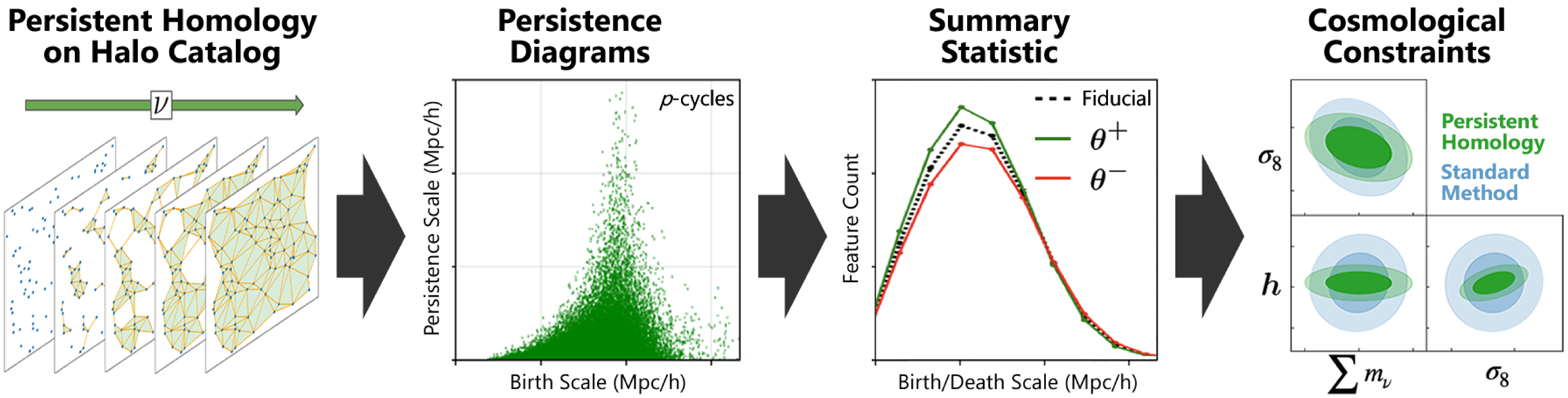
- Basics of Persistent homology

Overview

Information Maximizing Persistent Homology (IMPH)

- Basics of Persistent homology
- Persistent features of Large Scale Structure

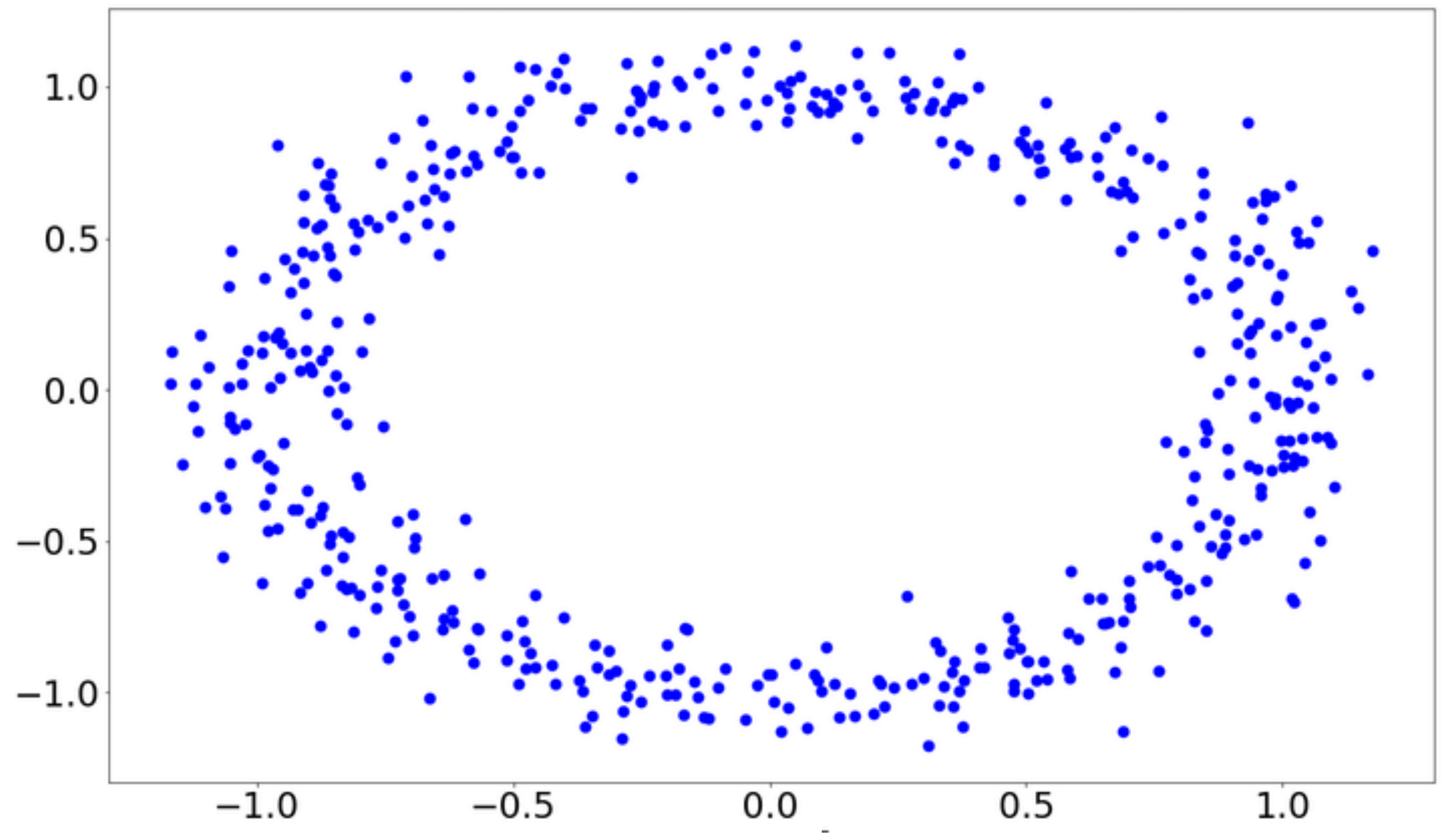
https://arxiv.org/abs/2403.13985 - with Yip, Biagetti et. al.

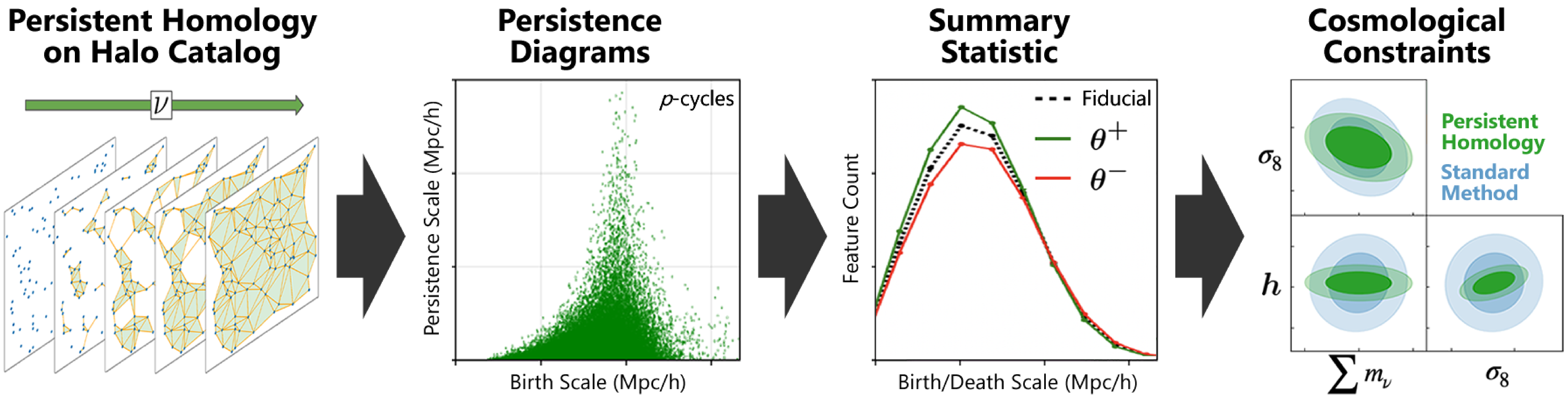


Overview

Information Maximizing Persistent Homology (IMPH)

- Basics of Persistent homology
- Persistent features of Large Scale Structure
- Lessons form the Noisy Circle





Motivation

Information maximizing neural network + differentiability of persistent homology = IMPH

- Persistent homology captures the LSS morphology as a distribution of clusters, loops and voids across scales.
    - It is a geometric way to organise information from higher order correlations.

Motivation

Information maximizing neural network + differentiability of persistent homology = IMPH

- Persistent homology captures the LSS morphology as a distribution of clusters, loops and voids across scales.
    - It is a geometric way to organise information from higher order correlations.
    - Does it give better constraints for the cosmological and PNG parameters?

Motivation

Information maximizing neural network + differentiability of persistent homology = IMPH

- Persistent homology captures the LSS morphology as a distribution of clusters, loops and voids across scales.
- Different hyperparameter choices in TDA probe different physics. We would like a systematic way of making optimal choices to suit the problem at hand.

Motivation

Information maximizing neural network + differentiability of persistent homology = IMPH

- Persistent homology captures the LSS morphology as a distribution of clusters, loops and voids across scales.

- Different hyperparameter choices in TDA probe different physics. We would like a systematic way of making optimal choices to suit the problem at hand.

- Differentiability of persistence homology allows us to employ gradient descent based methods. Can we find the optimal choices by maximizing the Fisher information of the resultant persistent summaries?

Persistent Homology 101

Persistent Homology 101

Multicale decomposition of clusters, loops and voids.

Introduction to Persistent Homology

Topology and Homology

Introduction to Persistent Homology

Topology and Homology

topologically equivalent

=

not topologically equivalent

≠

1 hole

1 hole

3 holes

$b_1(\text{mug}) = b_1(\text{donut}) = 1$

$b_1(\text{pretzel}) = 3$

$b_0 = 1$ # connected components

$b_2 = 0$ # voids

$H_i(X)$

i^{th} homology group of X

$= \{$ i-dim holes $\}$

Betti number

$b_i = \text{rank}(H_i)$

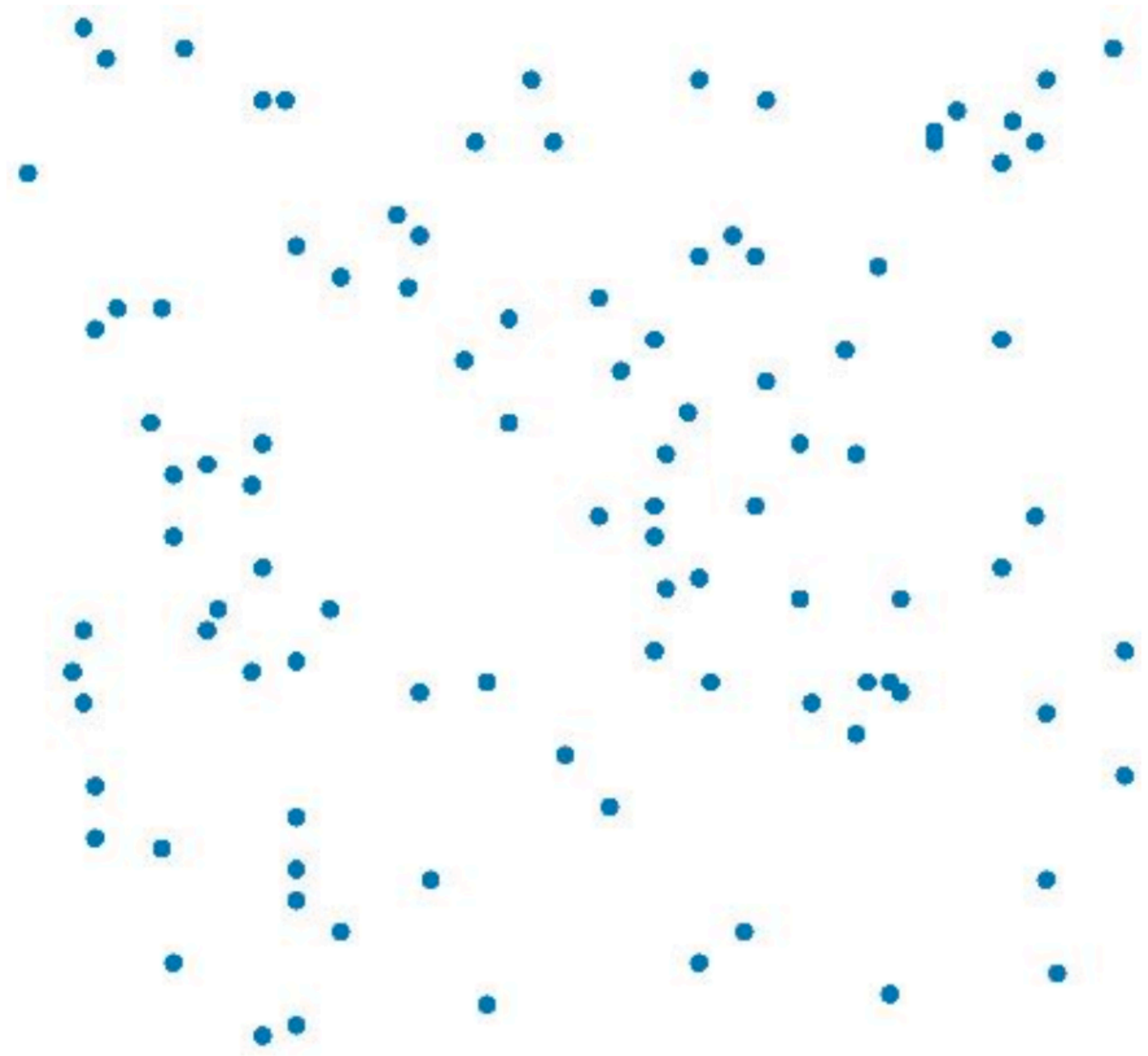
Introduction to Persistent Homology

Topological Data Analysis

- Compute the shape of discrete data via its multiscale topology - clusters, loops and voids.
- Offers a flexible that can also encode local density and knn statistics.
- Applications
    - Sensor networks, image processing, genomics, protein structure, neuroscience, physics and now to study large language models.

Introduction to Persistent Homology

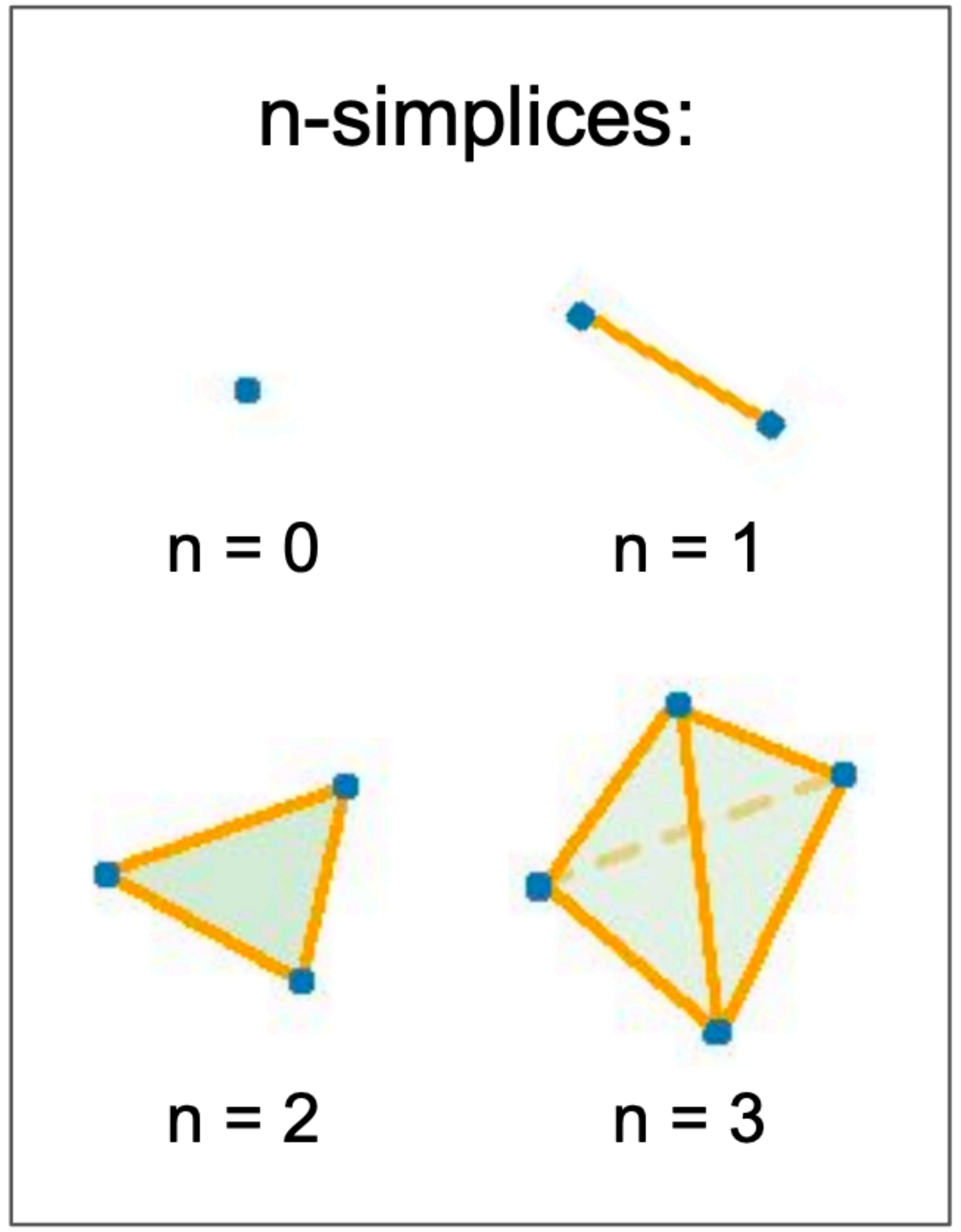
Homology of a point cloud

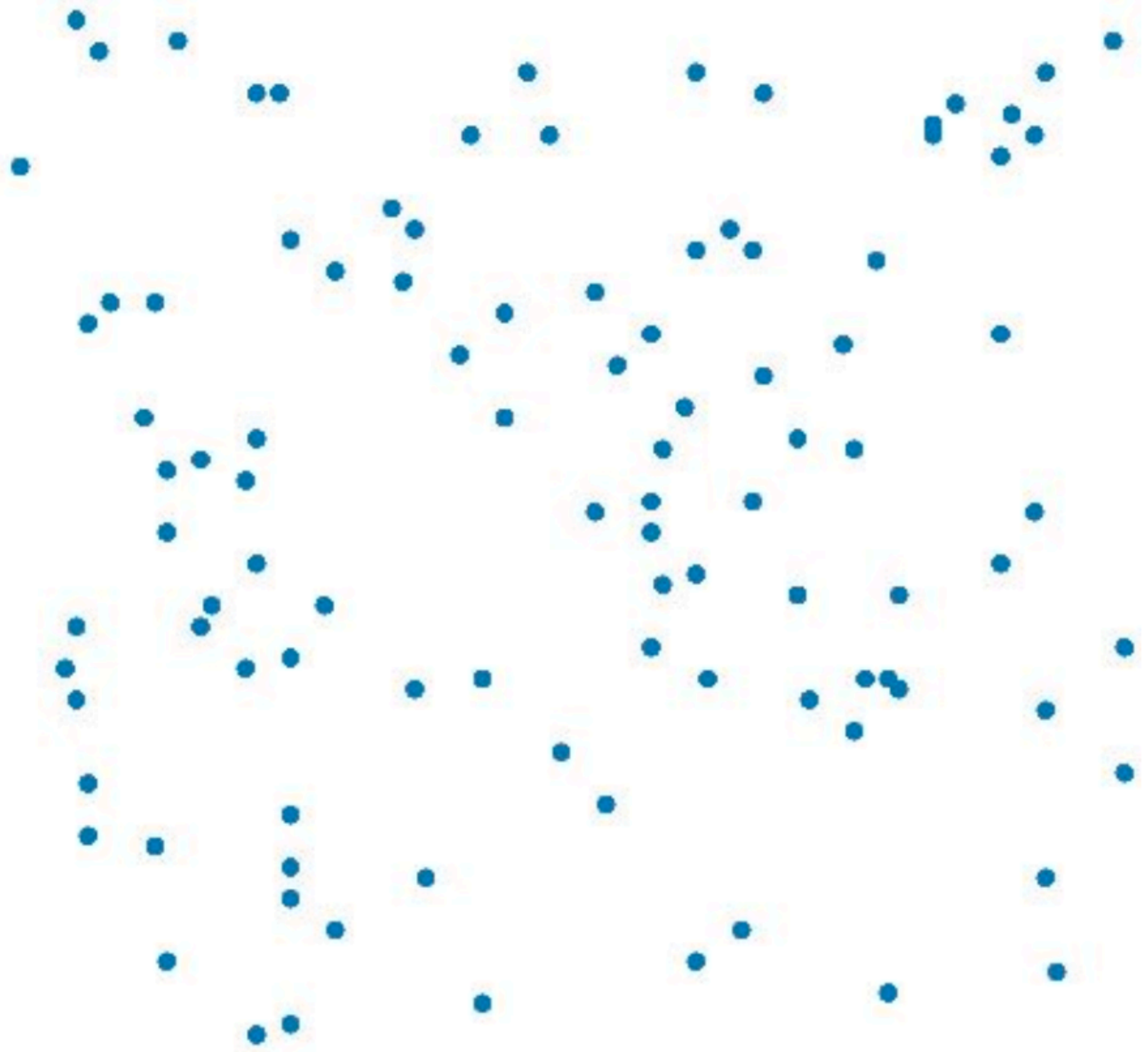


$b_0 = 100$
$b_1 = 0$

Introduction to Persistent Homology

Adding simplices

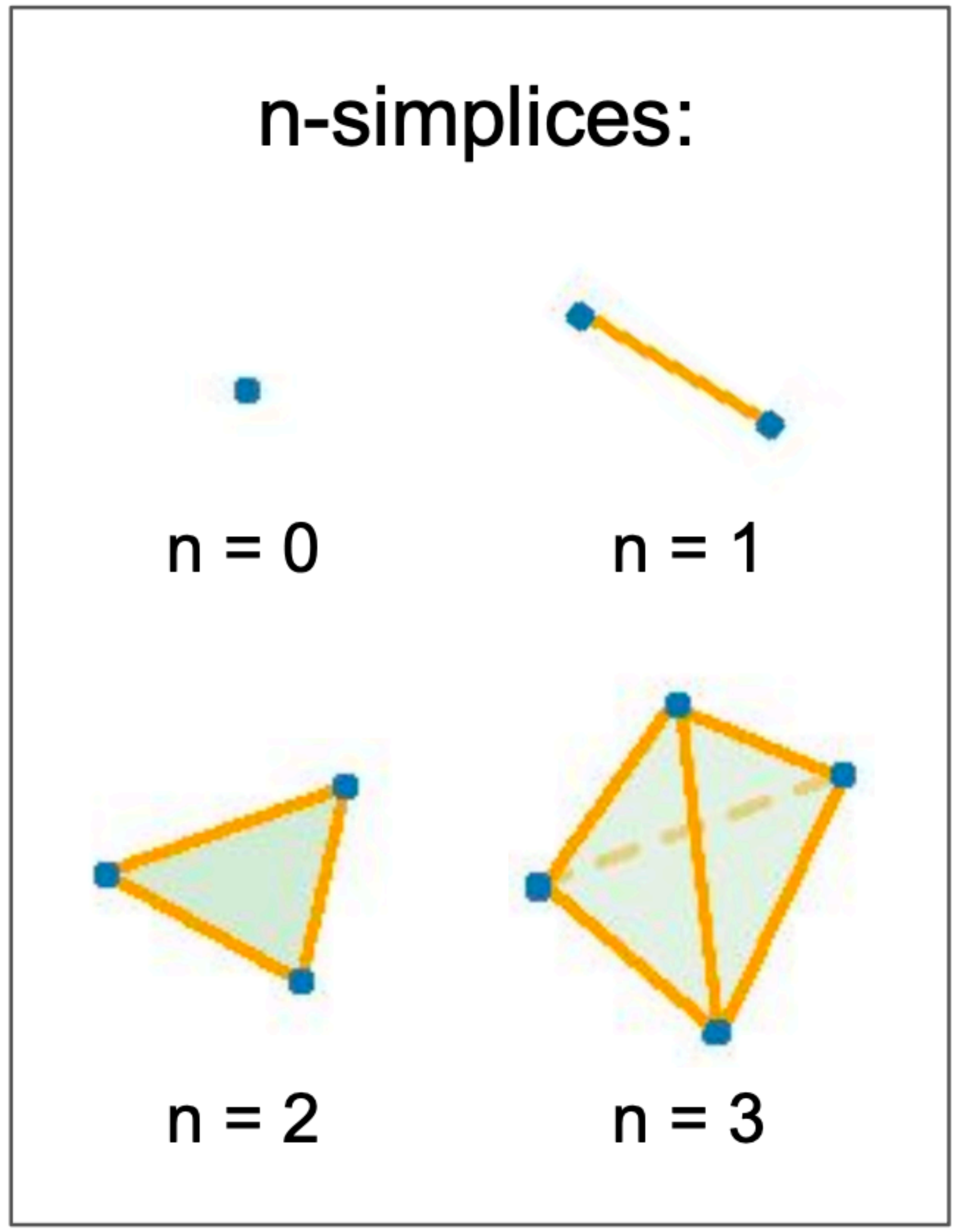


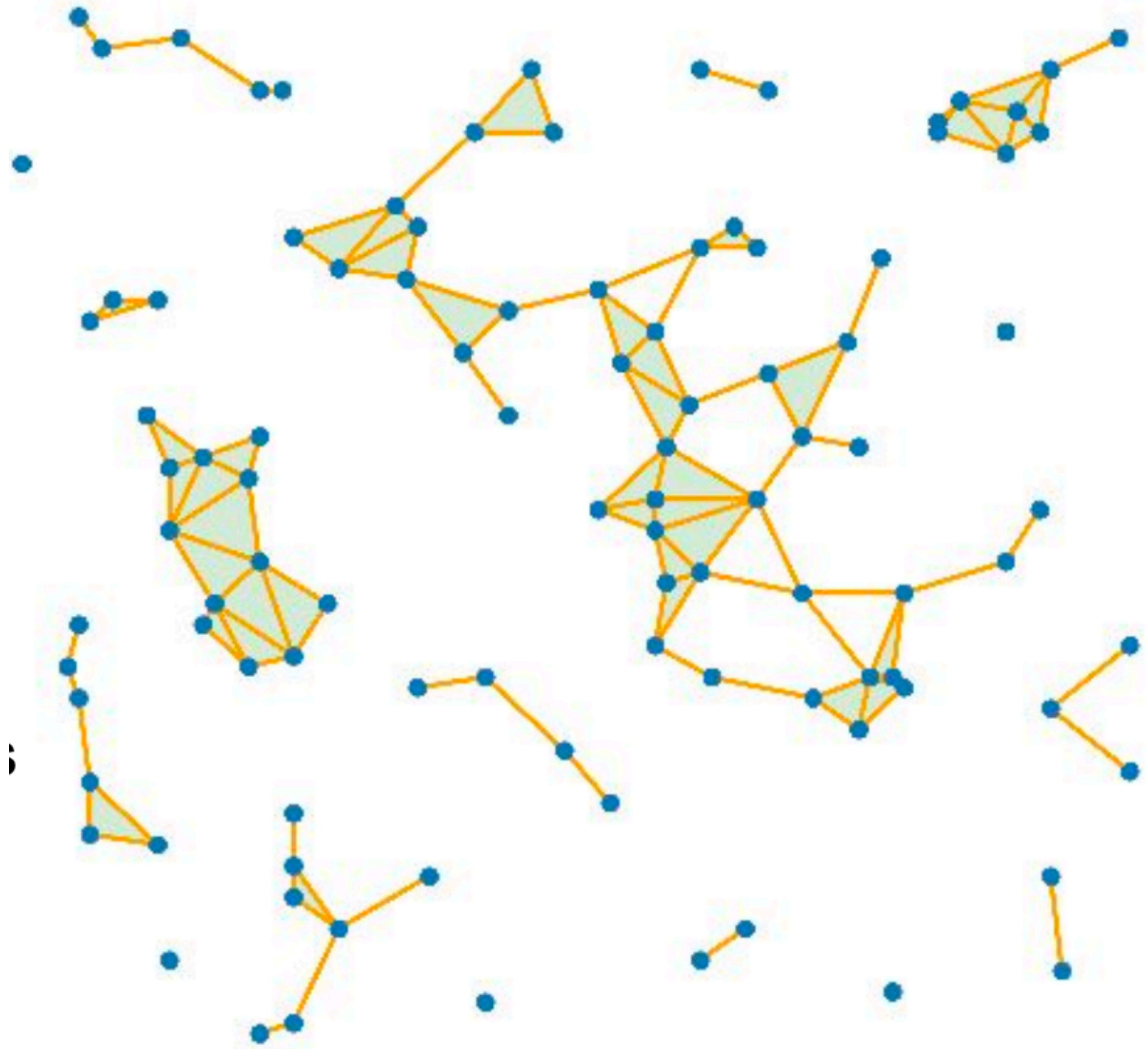


$b_0 = 100$
$b_1 = 0$

Introduction to Persistent Homology

Adding simplices





Simplicial Complex

$b_0 = 17$

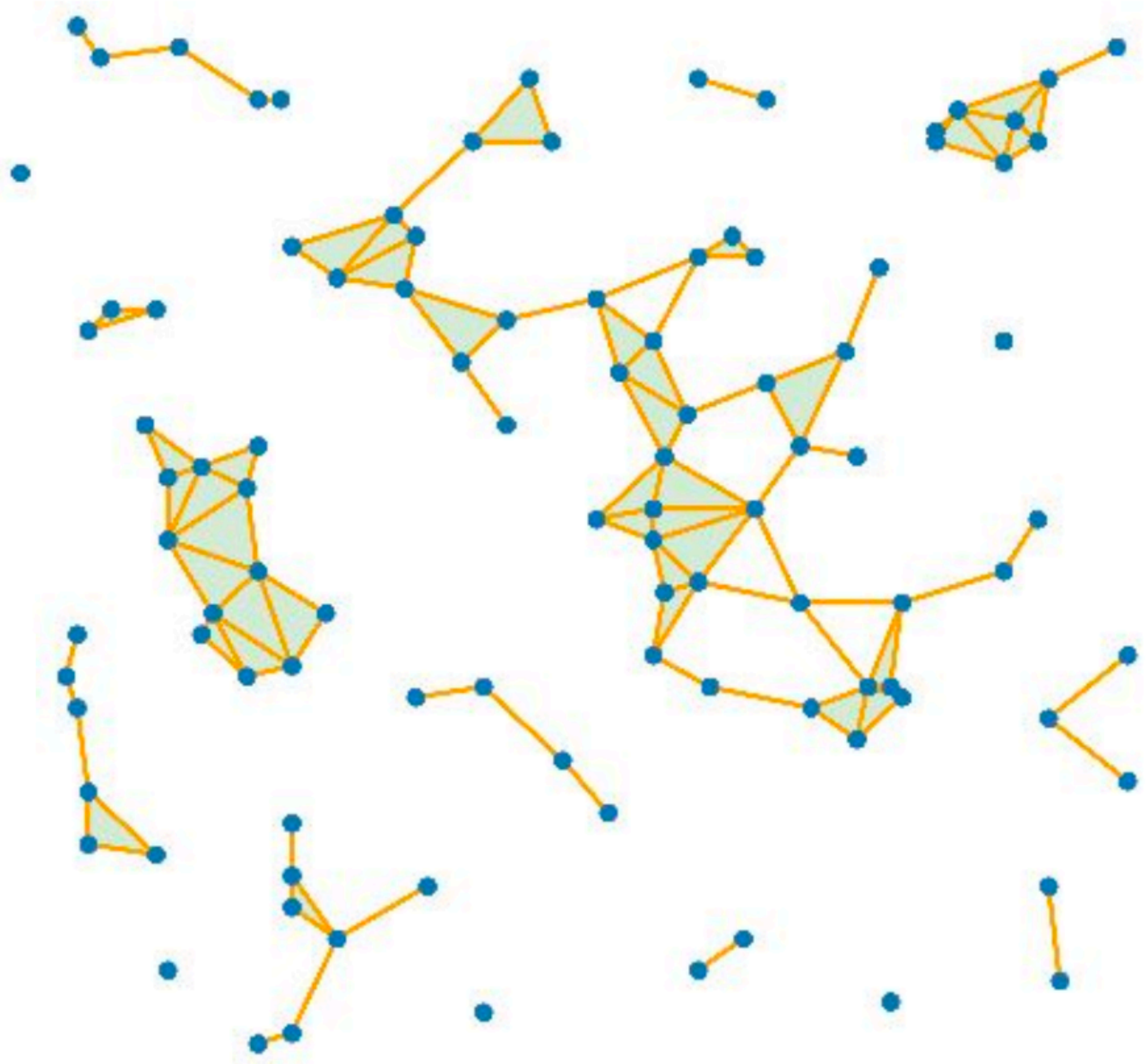
$b_1 = 5$

Introduction to Persistent Homology

Changing homology across scales

$\nu = 5$
Length scale parameter

imagine a ball of radius ν, when balls touch simplices are added to the complex



Simplicial Complex

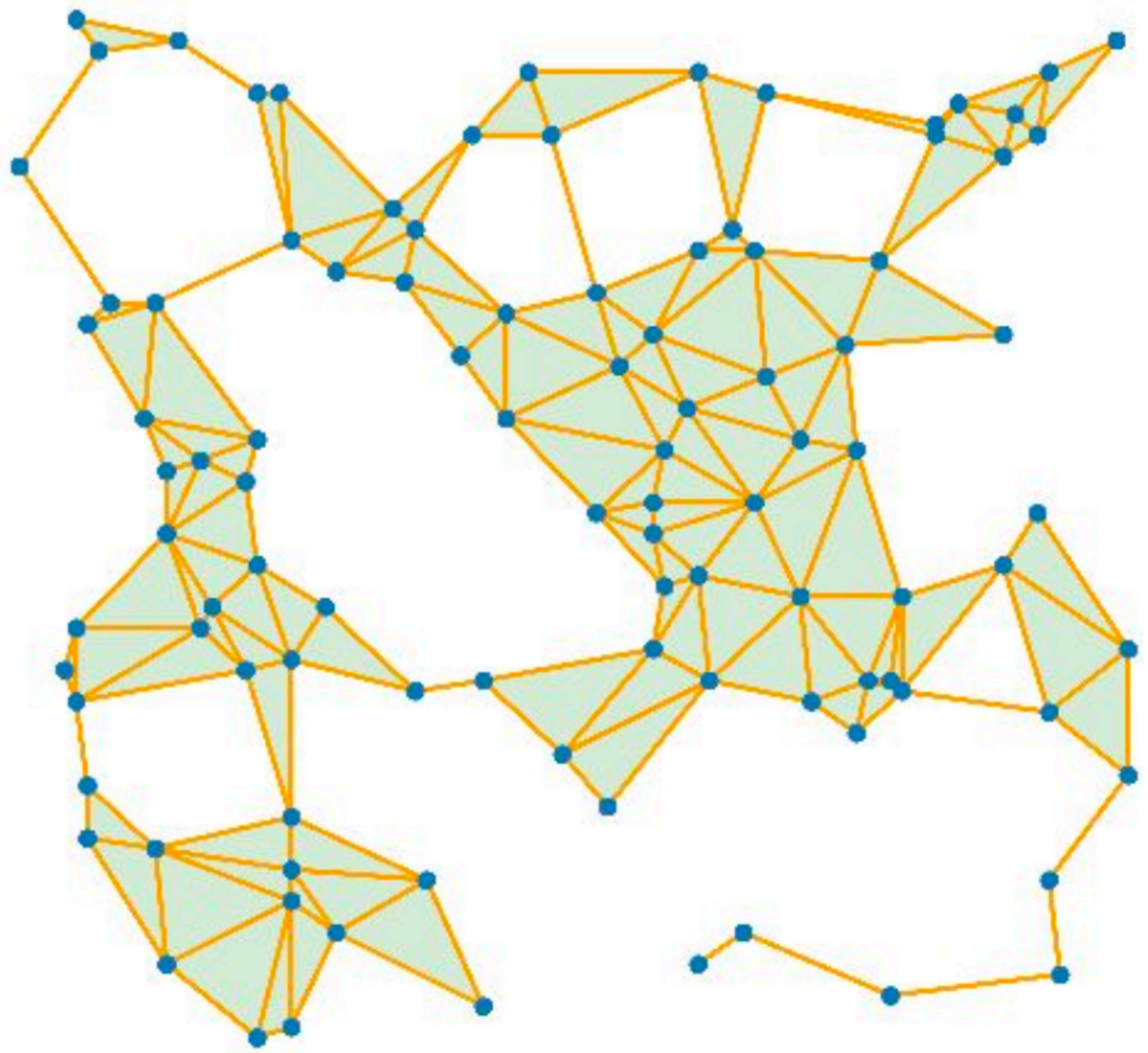
$b_0 = 17$
$b_1 = 5$

Introduction to Persistent Homology

Changing homology across scales

$\nu = 8$
Length scale parameter

imagine a ball of radius ν, when balls touch simplices are added to the complex



Simplicial Complex

$b_0 = 1$
$b_1 = 7$

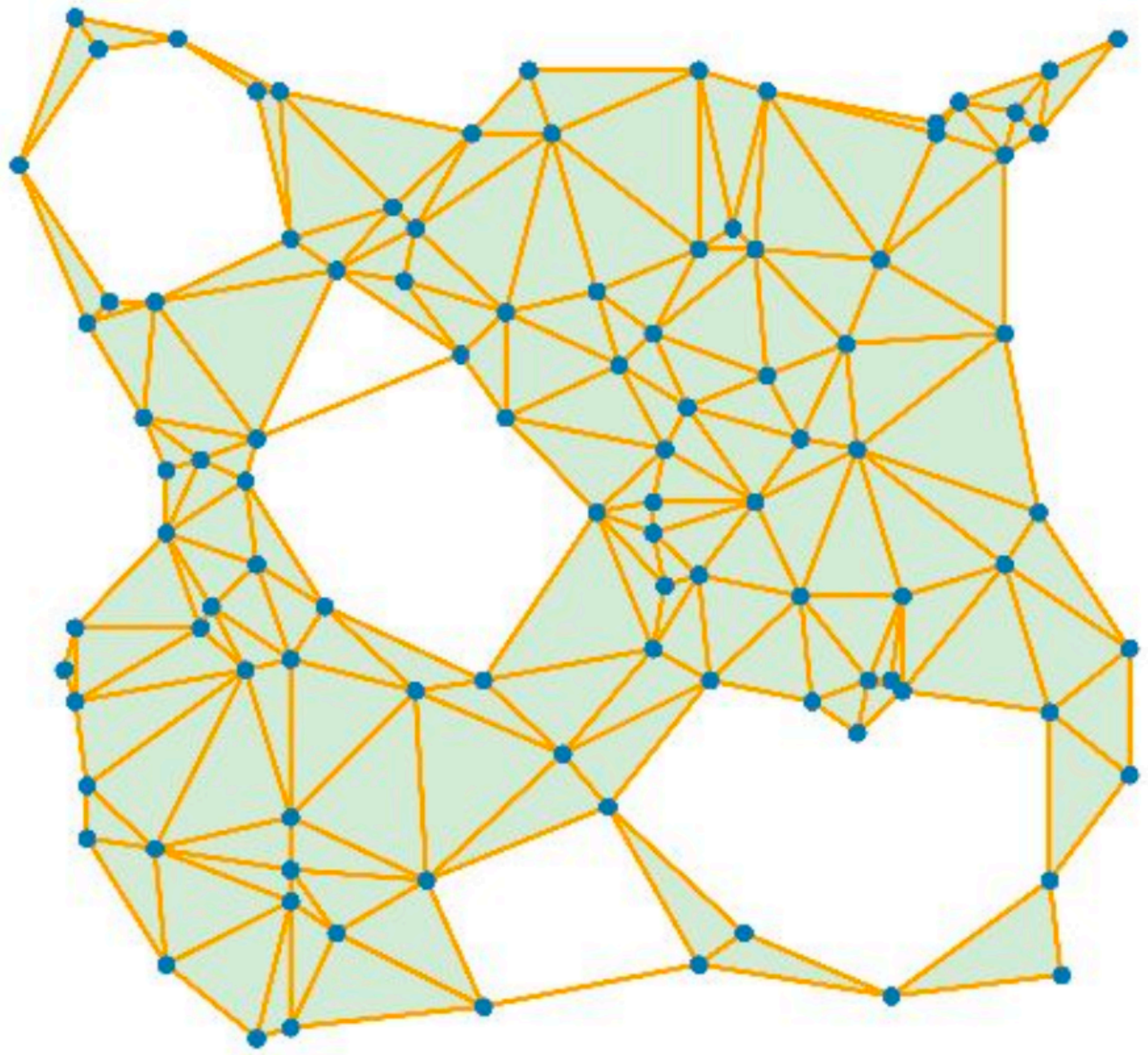
Introduction to Persistent Homology

Changing homology across scales

$\nu = 10$

Length scale parameter

imagine a ball of radius ν, when balls touch simplices are added to the complex



Simplicial Complex

$b_0 = 1$
$b_1 = 5$

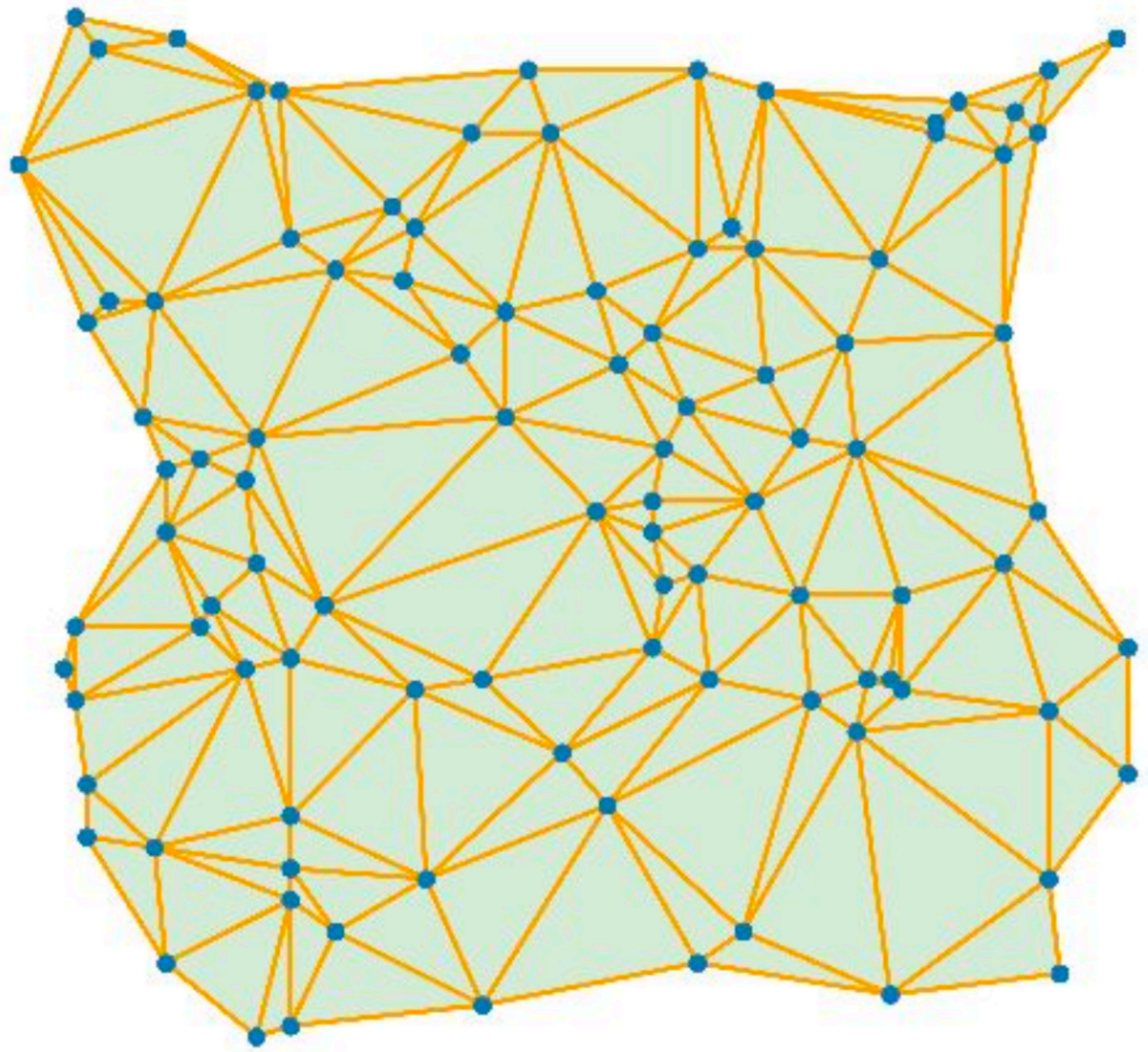
Introduction to Persistent Homology

Changing homology across scales

$\nu = 13$

Length scale parameter

imagine a ball of radius ν, when balls touch simplices are added to the complex



Simplicial Complex

$b_0 = 1$

$b_1 = 0$

The DTM function

A filtration robust to outliers

We can *delay* addition of outliers by penalising them if they are far apart from everything else

Introduce a Distance-To-Measure function

$$DTM(x) = \left(\frac{1}{k} \sum_{x_i \in N_k(x)} \quad x - x_i \; {}^{p} \right)^{1/p}$$

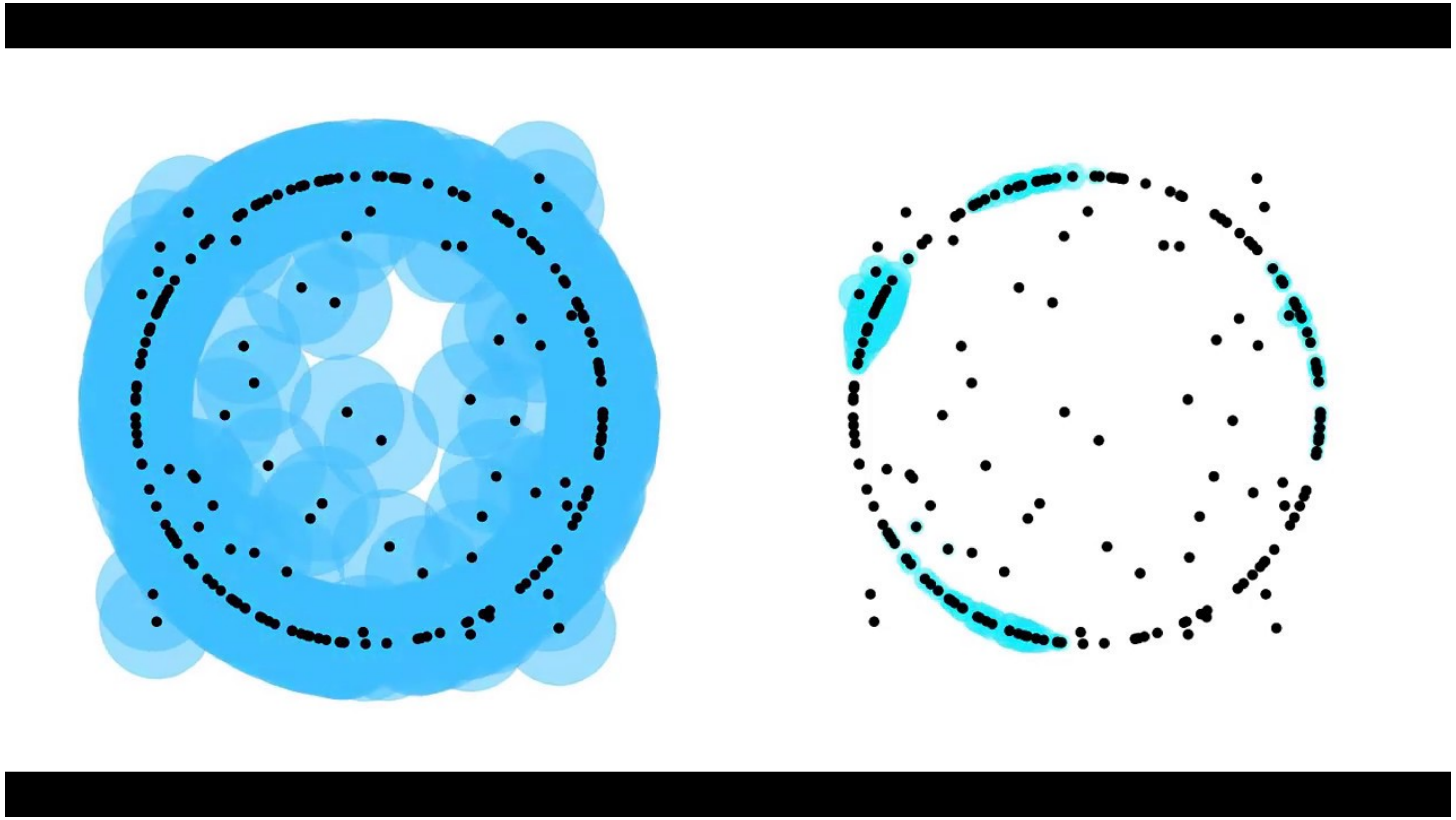
A filtration function takes assigns a real number to each point in the point cloud.

k : # of nearest neighbours
$N_k(x)$: the set of k-nearest neighbours of x
p : a mixing parameter (e.g. $p = 2$)

The DTM function

A filtration robust to outliers

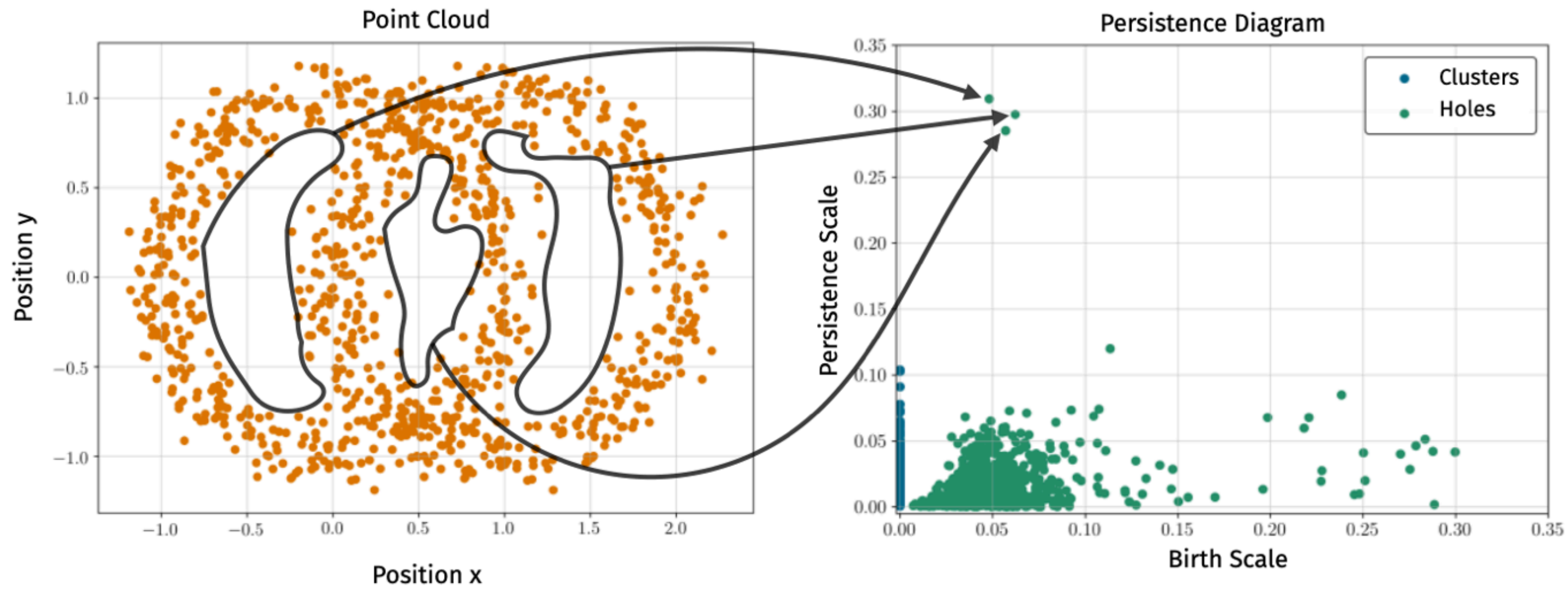
We can *delay* addition of outliers by penalising them if they are far apart from everything else



Anai, Chazal, Glisse, Ike, Inakoshi, Tinarrage, Umeda (2018)

Introduction to Persistent Homology

Tracking persistent features



We can then draw a persistence diagram, $\nu_{\text{persist}} = \nu_{\text{death}} - \nu_{\text{birth}}$, as a function of ν_{birth}

Summarising persistence diagrams

Histograms

ID	Dimension		
1	0	0	0.5
2	0	0	0.707
3	0	0	1
...	...	...	...
100	0	0	inf
101	1	8.732	8.733
...	...	...	...
178	1	7.632	12.029
179	1	8.485	12.889
180	1	6.718	12.905

ν_{birth}

ν_{death}

Persistence diagram as a list of birth and death times.

Summarising persistence diagrams

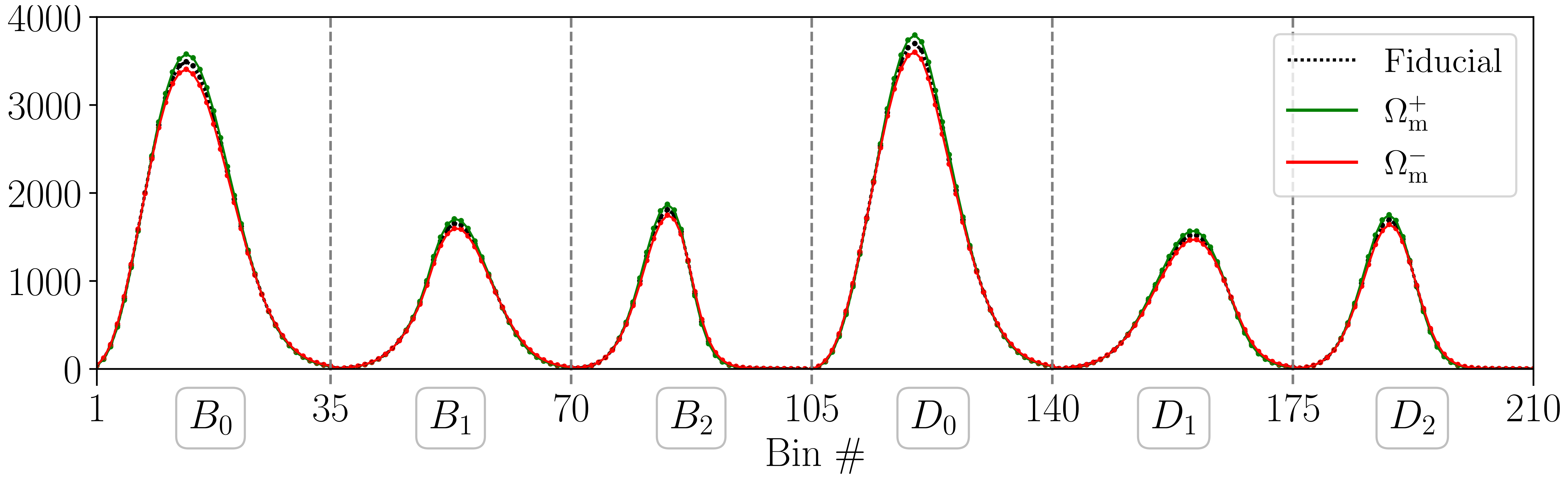
Histograms

ID	Dimension		
1	0	0	0.5
2	0	0	0.707
3	0	0	1
...	...	...	...
100	0	0	inf
101	1	8.732	8.733
...	...	...	...
178	1	7.632	12.029
179	1	8.485	12.889
180	1	6.718	12.905

Persistence diagram as a list of birth and death times.

ν_{birth}

ν_{death}



Different homology dimensions

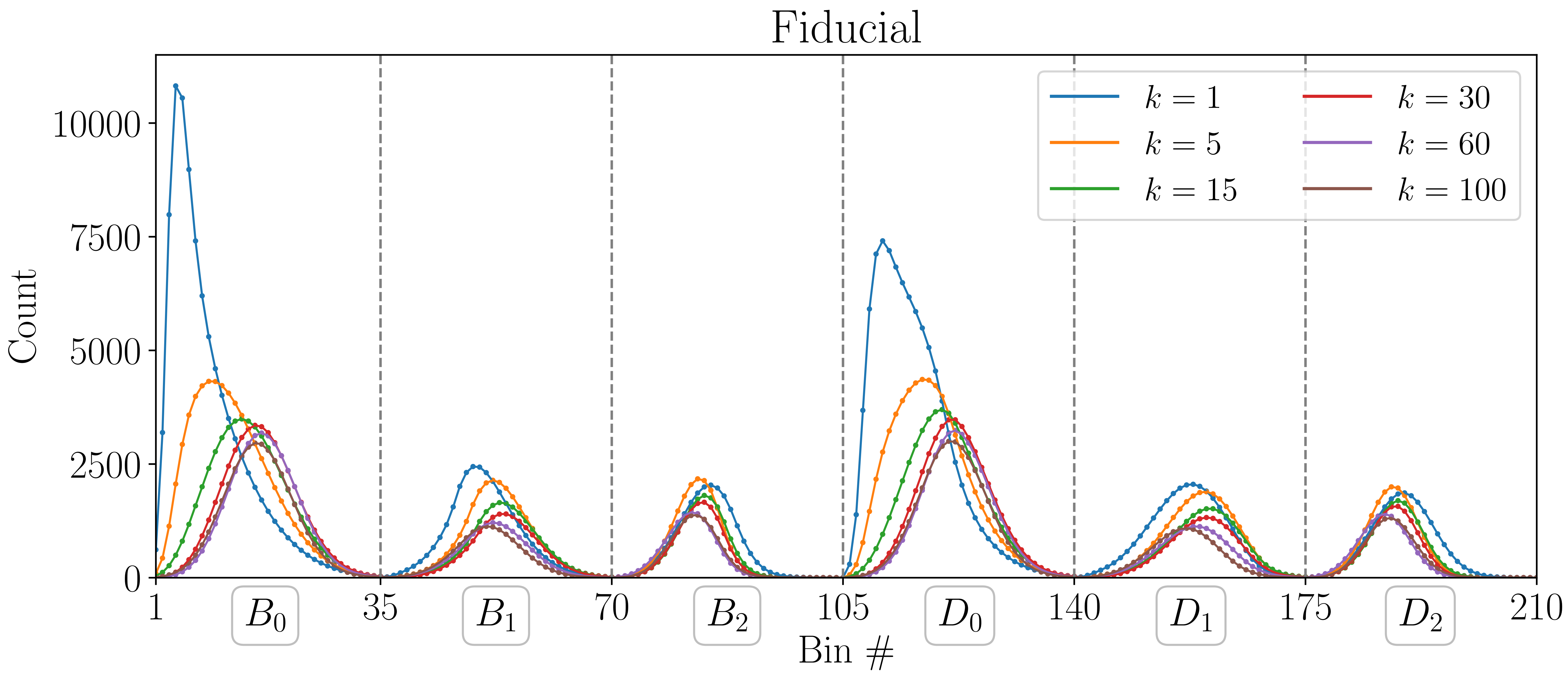
Different homology dimensions

Persistent Features of Large Scale Structure

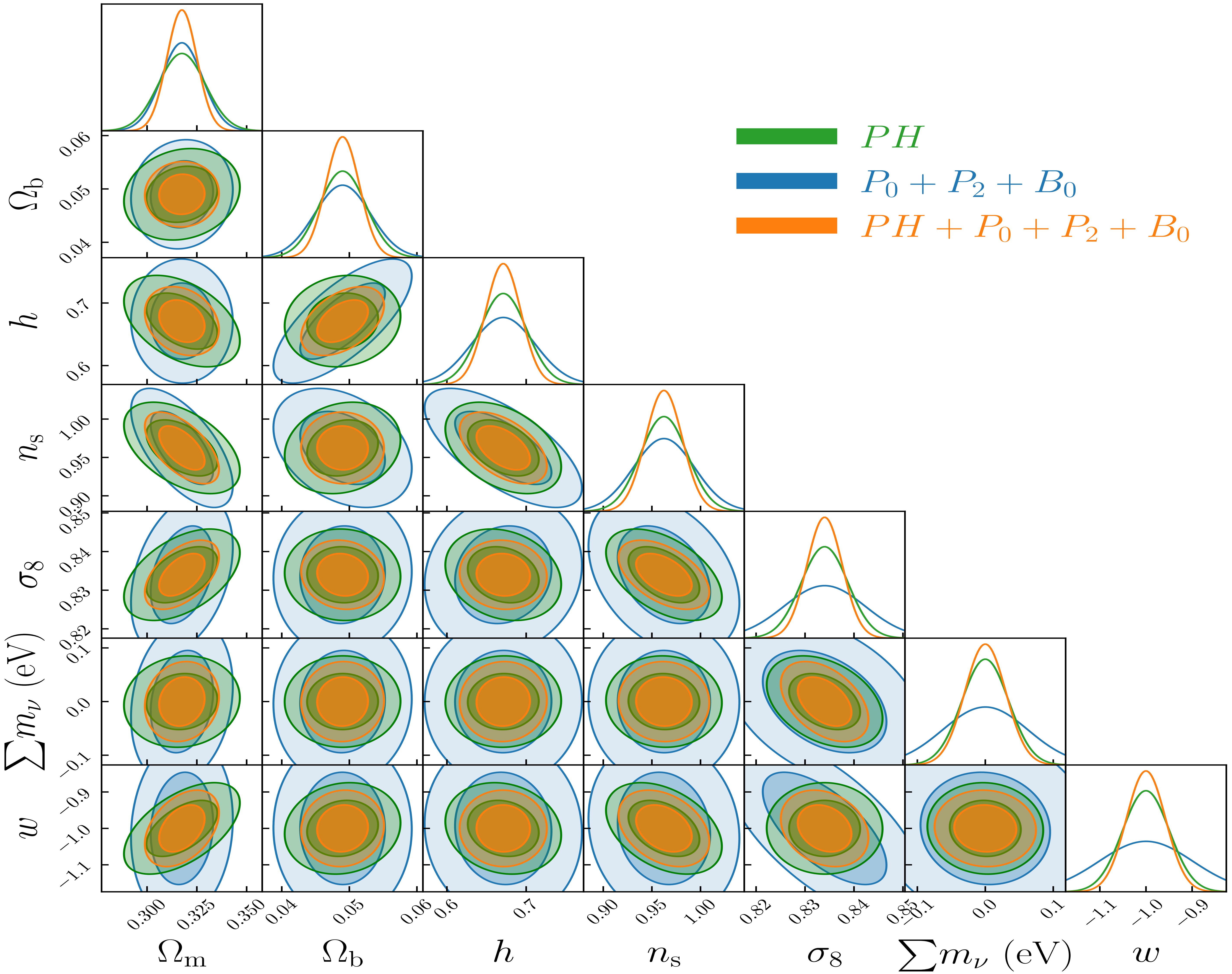
https://arxiv.org/abs/2403.13985 - with Yip, Biagetti et. al.

Implementation Details

- Dataset - Quijote simulations
- Filtration - AlphaDTM filtration for k = (1, 5, 15, 30, 60, 100).
- Vectorization - Histogram of counts.

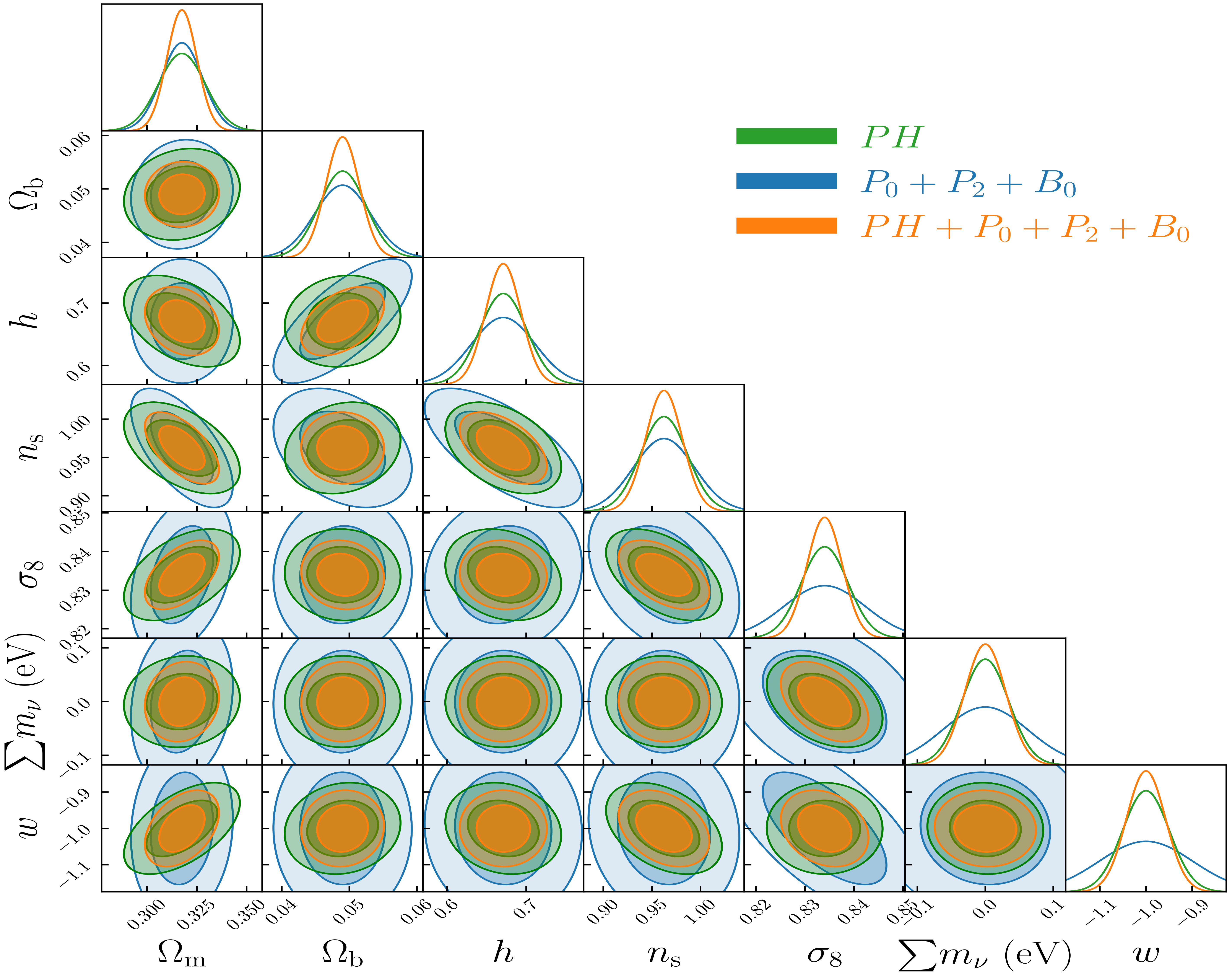


Fisher contours for cosmological parameters



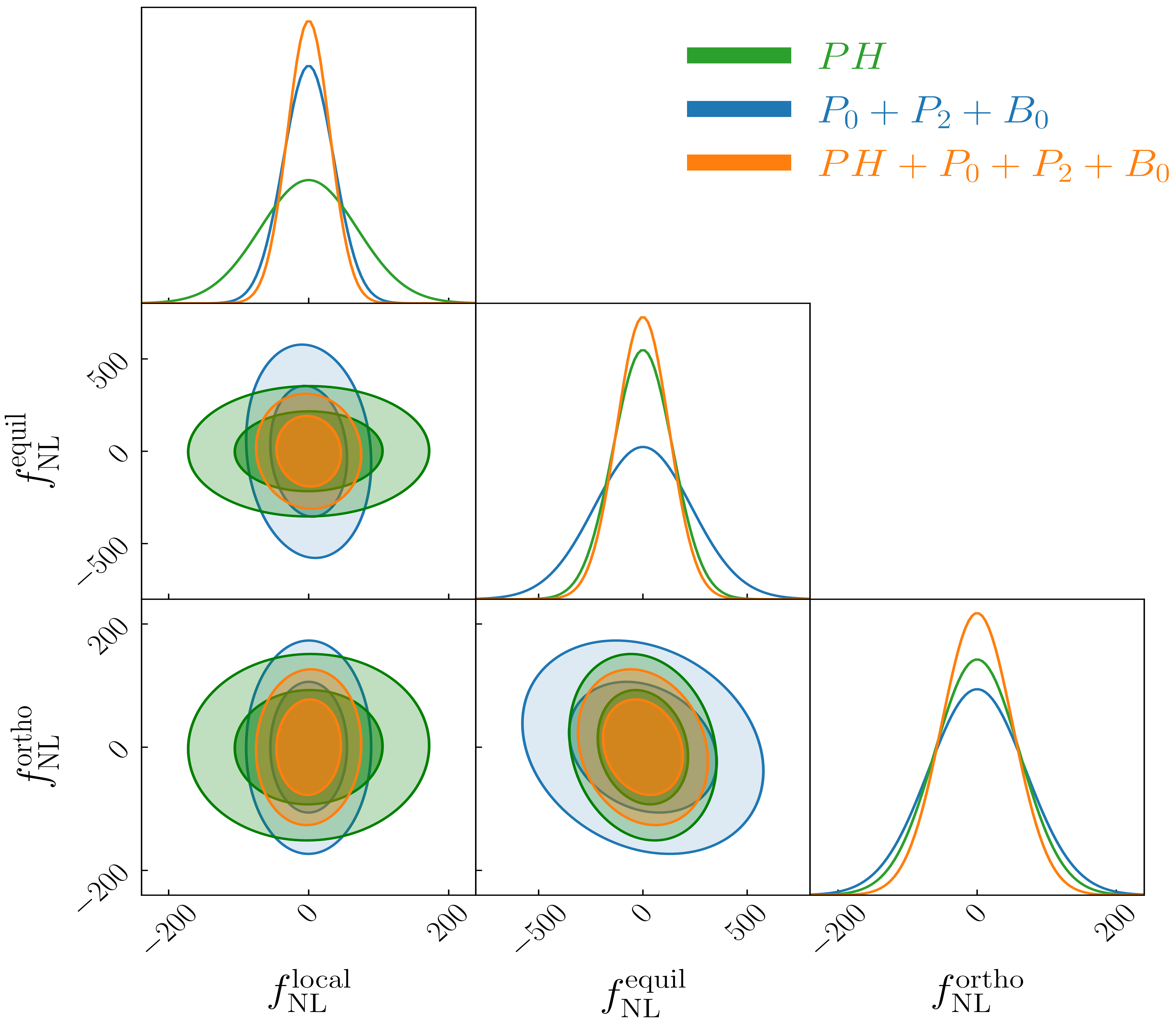
- The contours from combining PH and PS + BS are more constraining in most of the cases.
- This could be because PH is assessing information of higher order correlators.

Fisher contours for cosmological parameters



- The contours from combining PH and PS + BS are more constraining in most of the cases.
- This could be because PH is assessing information of higher order correlators.
- The parameter degeneracies for our statistic are in directions fairly different from those for the joint power spectrum and bispectrum statistic

Fisher contours for PNG amplitudes



- Tighter constraints for equilateral and orthogonal PNG.
- Local PNG better constrained by joint power spectrum and bispectrum since most of the information is in the larger scale and not many cycles persist in the large scales.

Summary of choices

A set of discrete choices that work well together

- AlphaDTM filtration with k = (1, 5, 15, 30, 60, 100)
- Histogram of counts to summarise the resulting persistence diagrams.

Summary of choices

A set of discrete choices that work well together

- AlphaDTM filtration with k = (1, 5, 15, 30, 60, 100)
- Histogram of counts to summarise the resulting persistence diagrams.
- These were empirical choice. Can we come up with a more versatile way of deciding on the filtration and vectorization?
- The resulting summaries can then be used as a part of an inference pipeline.

Before checking this on DM halo simulations, we try our method on a simpler example - the noisy circle.

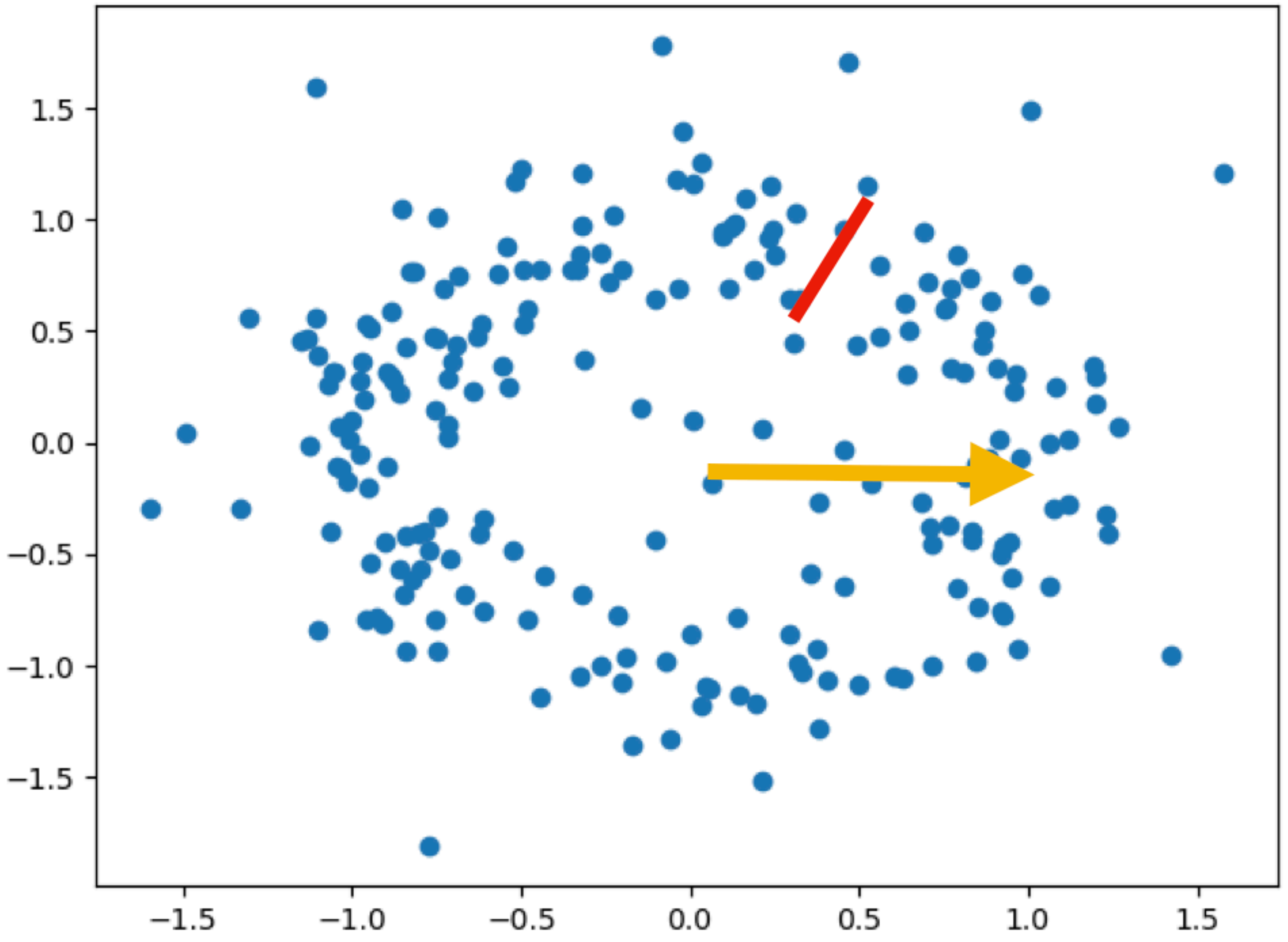
Information Maximizing Persistent Homology

The noisy circle

To appear- with Biagetti, Yip, van der Schaar, et. al.

Extract information about radius (and variance) from a noisy ring

Tractable Fisher information!



Noisy ring: mixture of uniform distribution and gaussian distribution around ring of unit radius (200 points) + 20 background points as noise (uniformly distributed)

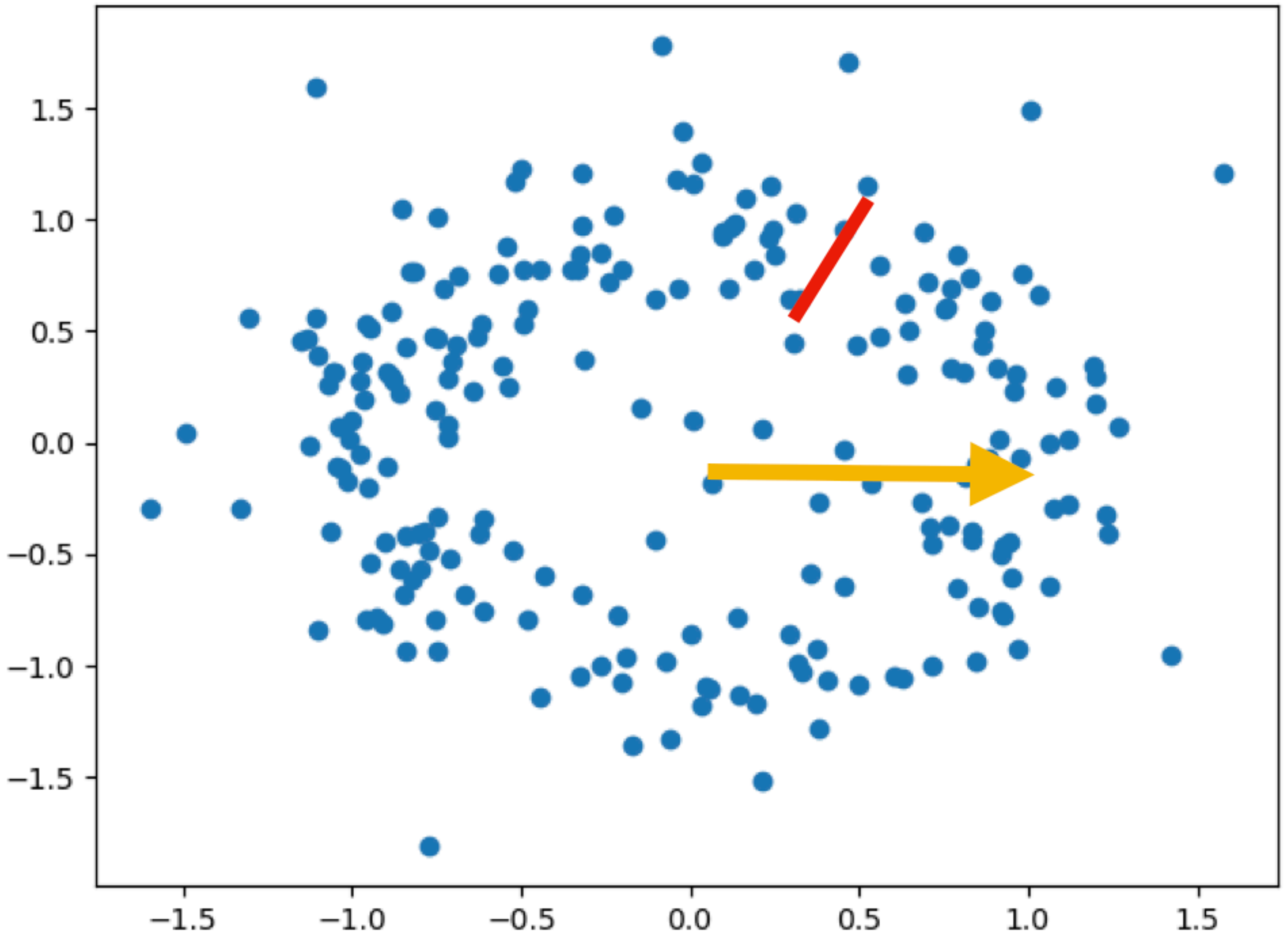
Information Maximizing Persistent Homology

The noisy circle

To appear- with Biagetti, Yip, van der Schaar, et. al.

Extract information about radius (and variance) from a noisy ring

Tractable Fisher information!



Can we learn an optimal filtration function that optimises the Fisher Information?

Noisy ring: mixture of uniform distribution and gaussian distribution around ring of unit radius (200 points) + 20 background points as noise (uniformly distributed)

Information Maximizing Persistent Homology

The noisy circle

To appear- with Biagetti, Yip, van der Schaar, et. al.

- We consider a neural network that takes the *k* nearest neighbour distances as input and outputs the filtration value for each vertex.

- The Fisher Information is calculated on the resulting persistence summaries.

- The filtration function is learnt to maximise the Fisher Information.

Information Maximizing Persistent Homology

The noisy circle

To appear- with Biagetti, Yip, van der Schaar, et. al.

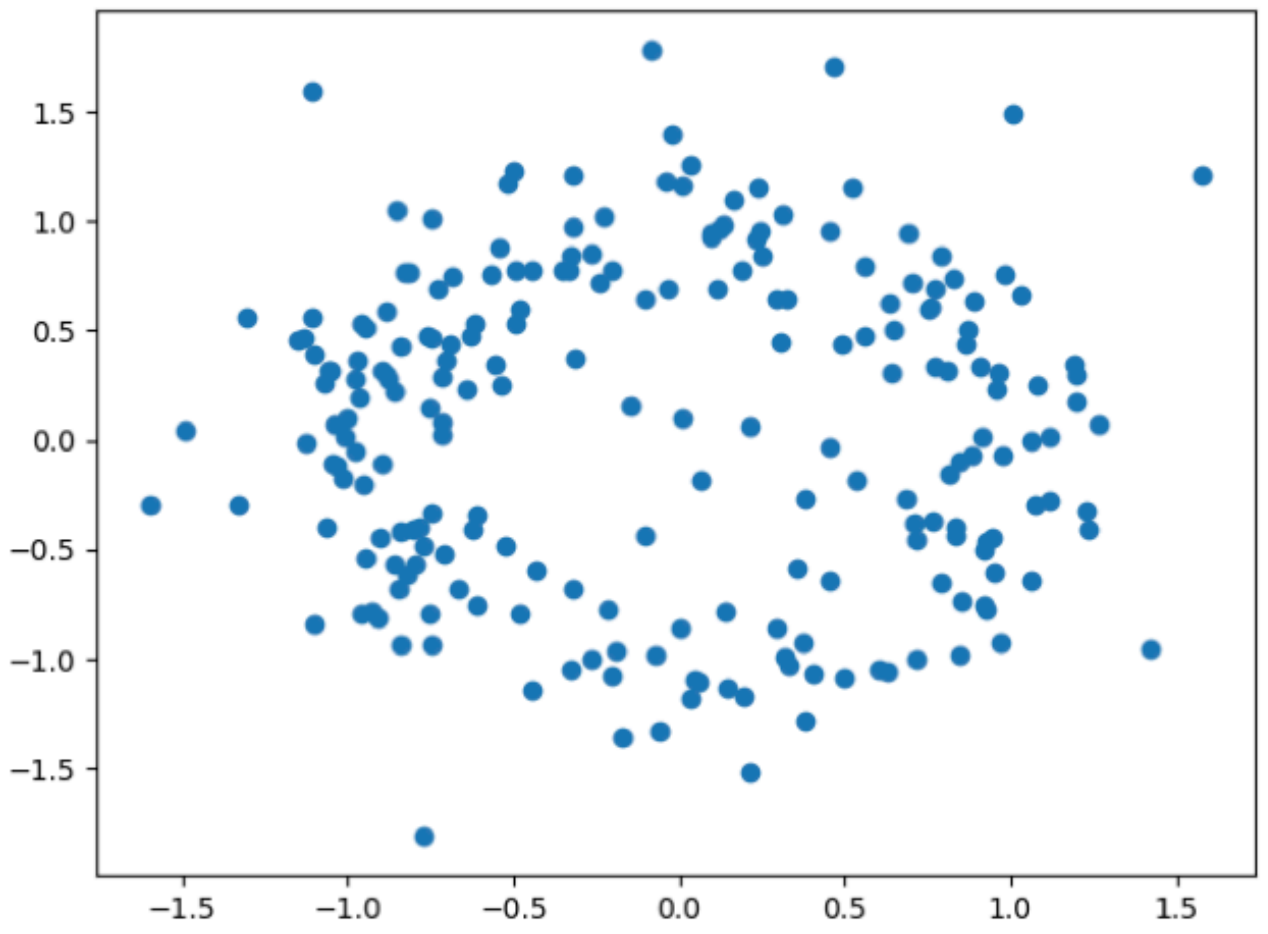
- We consider a neural network that takes the *k* nearest neighbour distances as input and outputs the filtration value for each vertex.
- The Fisher Information is calculated on the resulting persistence summaries.
- The filtration function is learnt to maximise the Fisher Information.
- To summarise the persistence diagram
  - We compress summaries using MOPED and IMNN that give more accurate and trustable estimates.
  - Uncompressed summaries overestimate the Fisher information due to imprecise derivatives and presence of high dimensional non-Gaussianity.

Information Maximizing Filtrations

Can we learn the optimal filtration?

To appear- with Biagetti, Yip, van der Schaar, et. al.

Point cloud

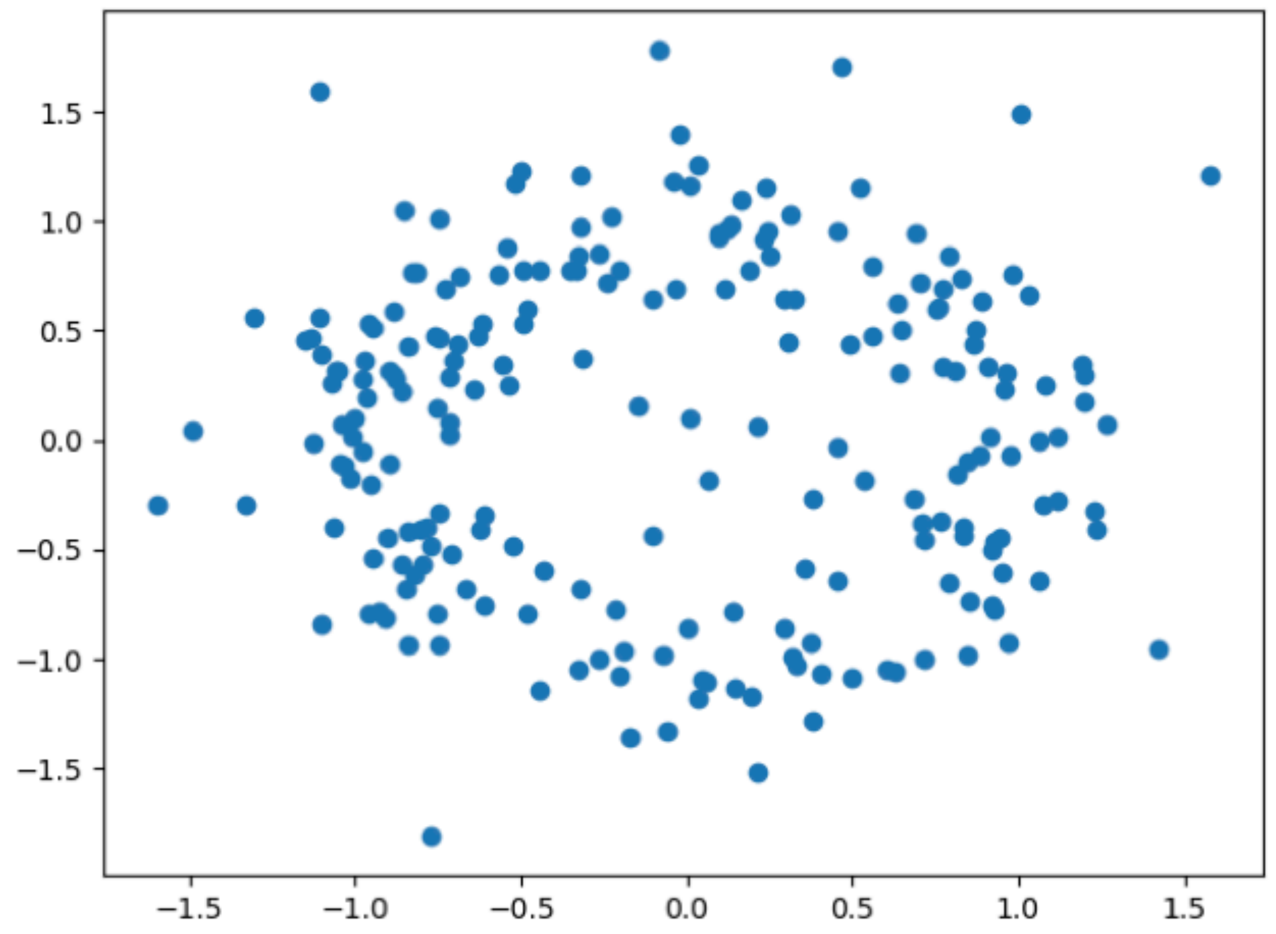


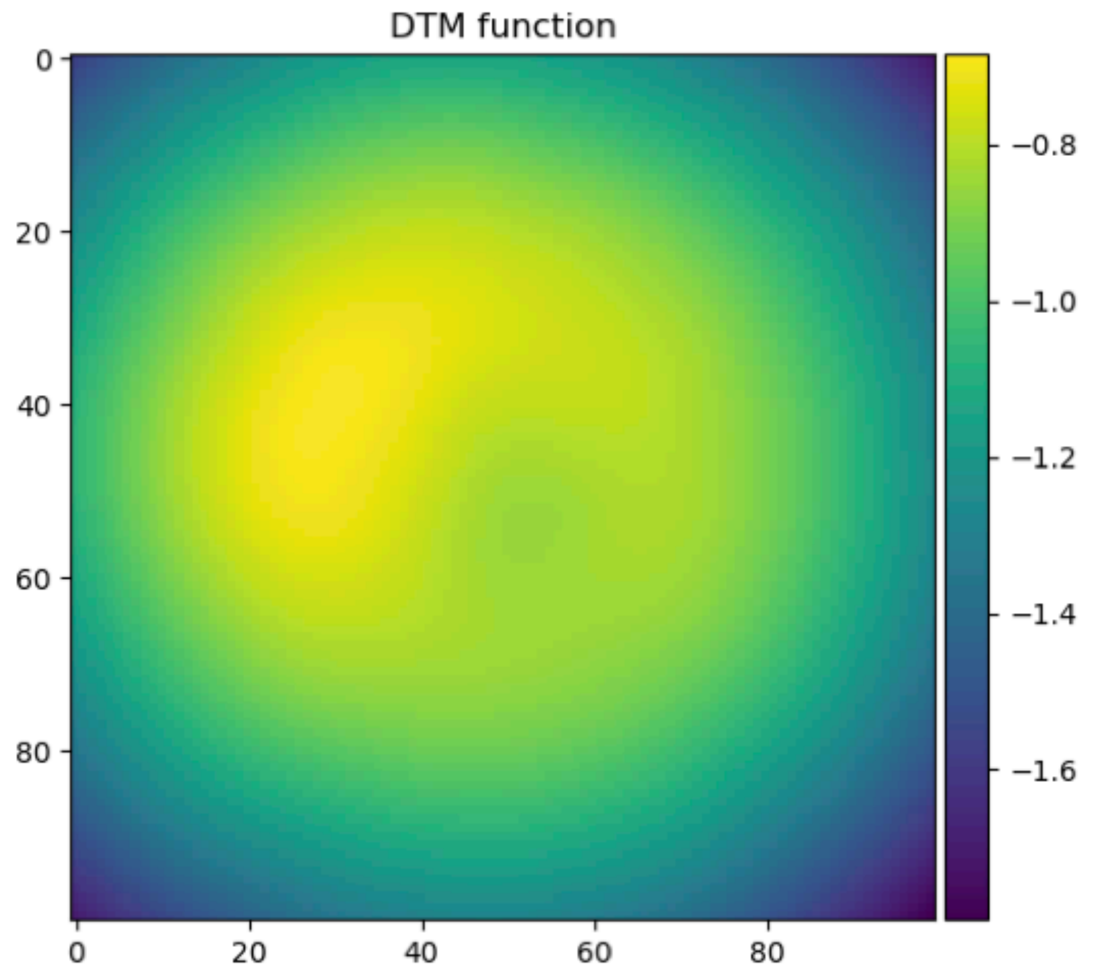
Information Maximizing Filtrations

Can we learn the optimal filtration?

To appear- with Biagetti, Yip, van der Schaar, et. al.

Point cloud





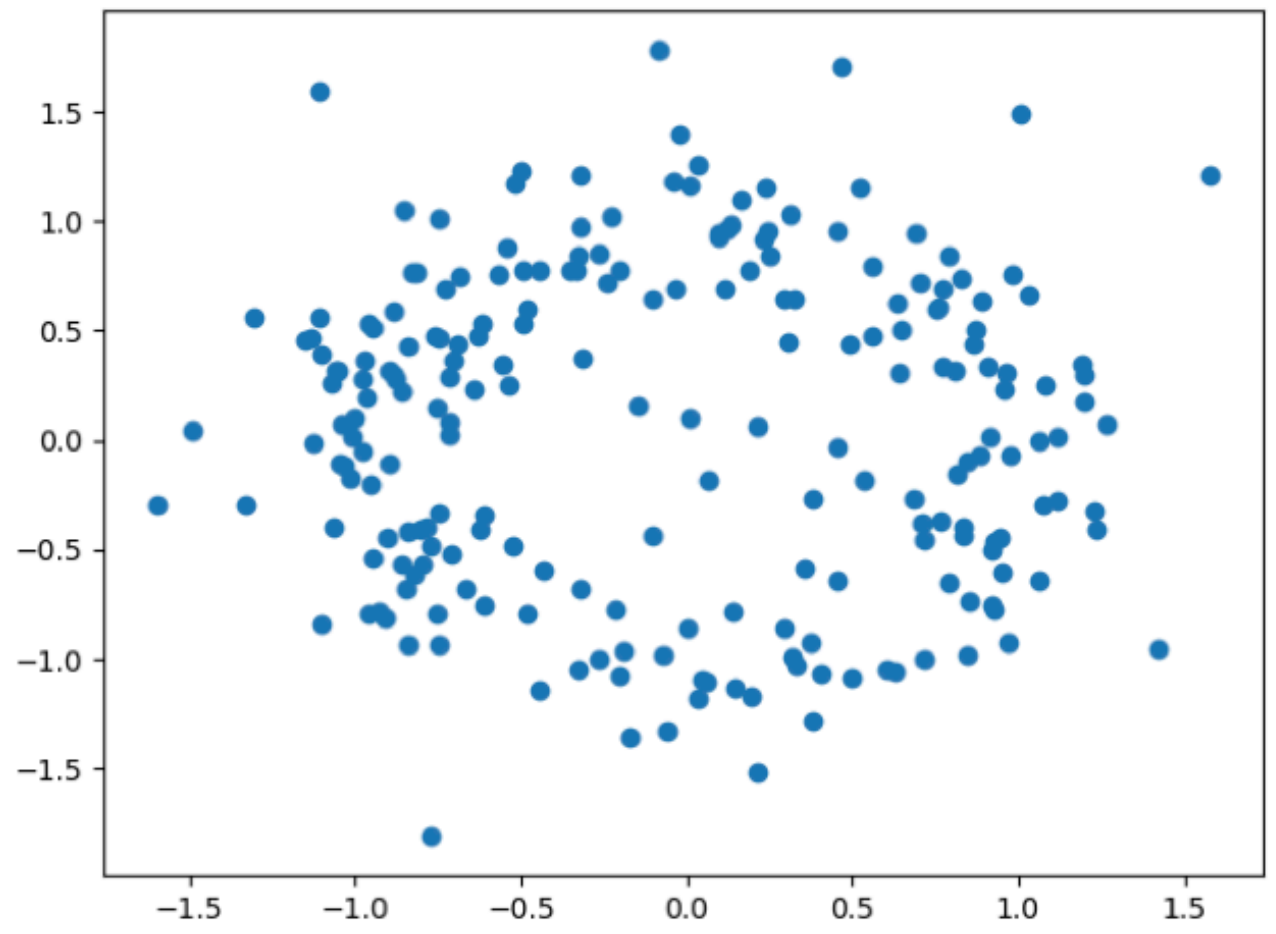
The distance to measure function

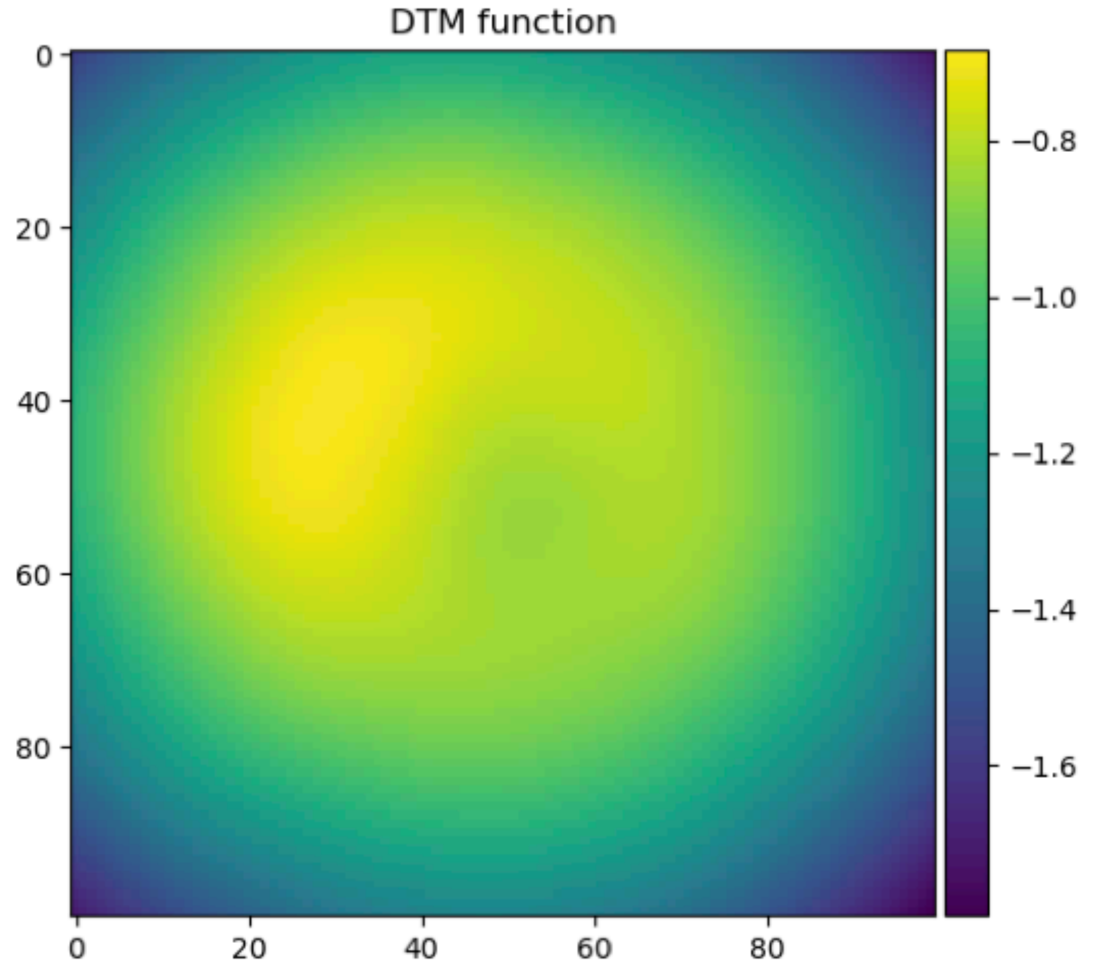
Information Maximizing Filtrations

Can we learn the optimal filtration?

To appear- with Biagetti, Yip, van der Schaar, et. al.

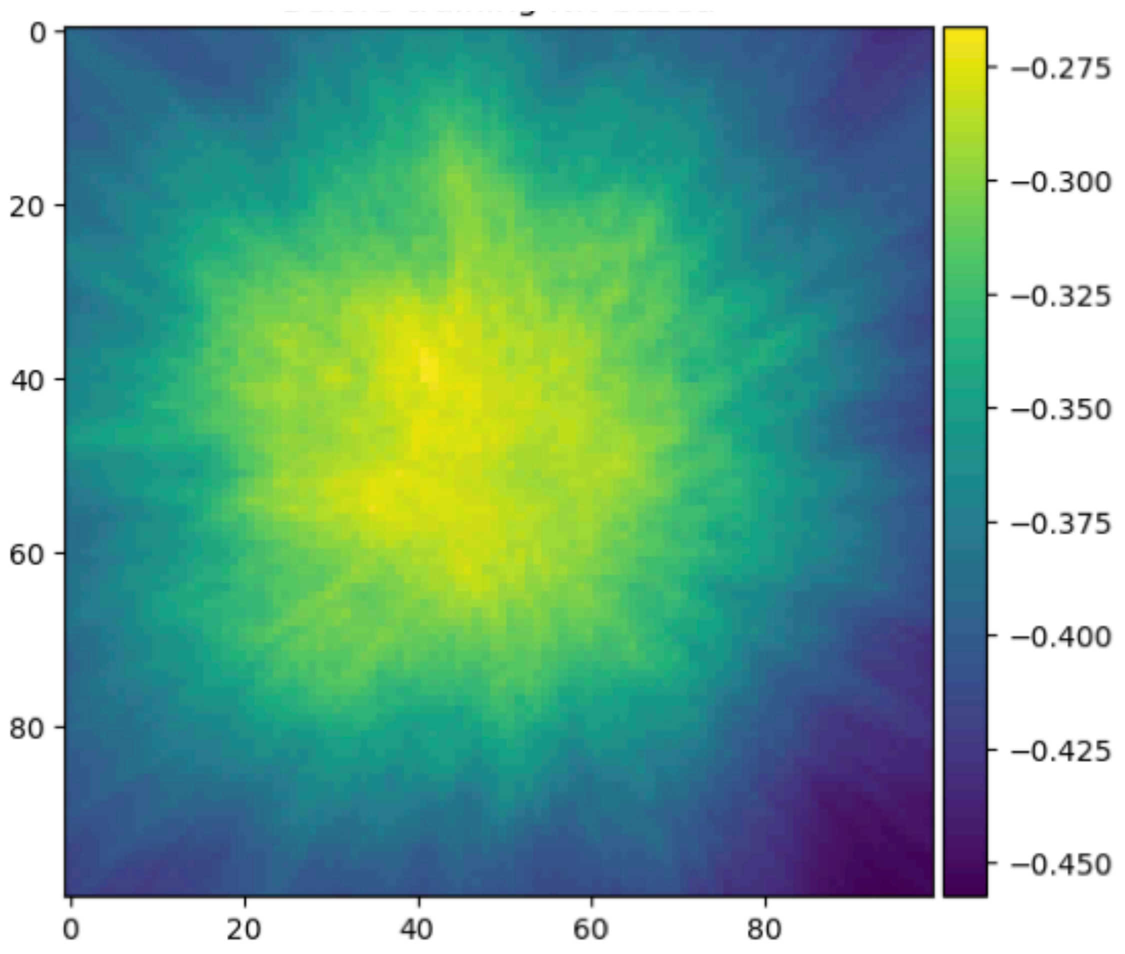
Point cloud





The distance to measure function

Untrained filtration function

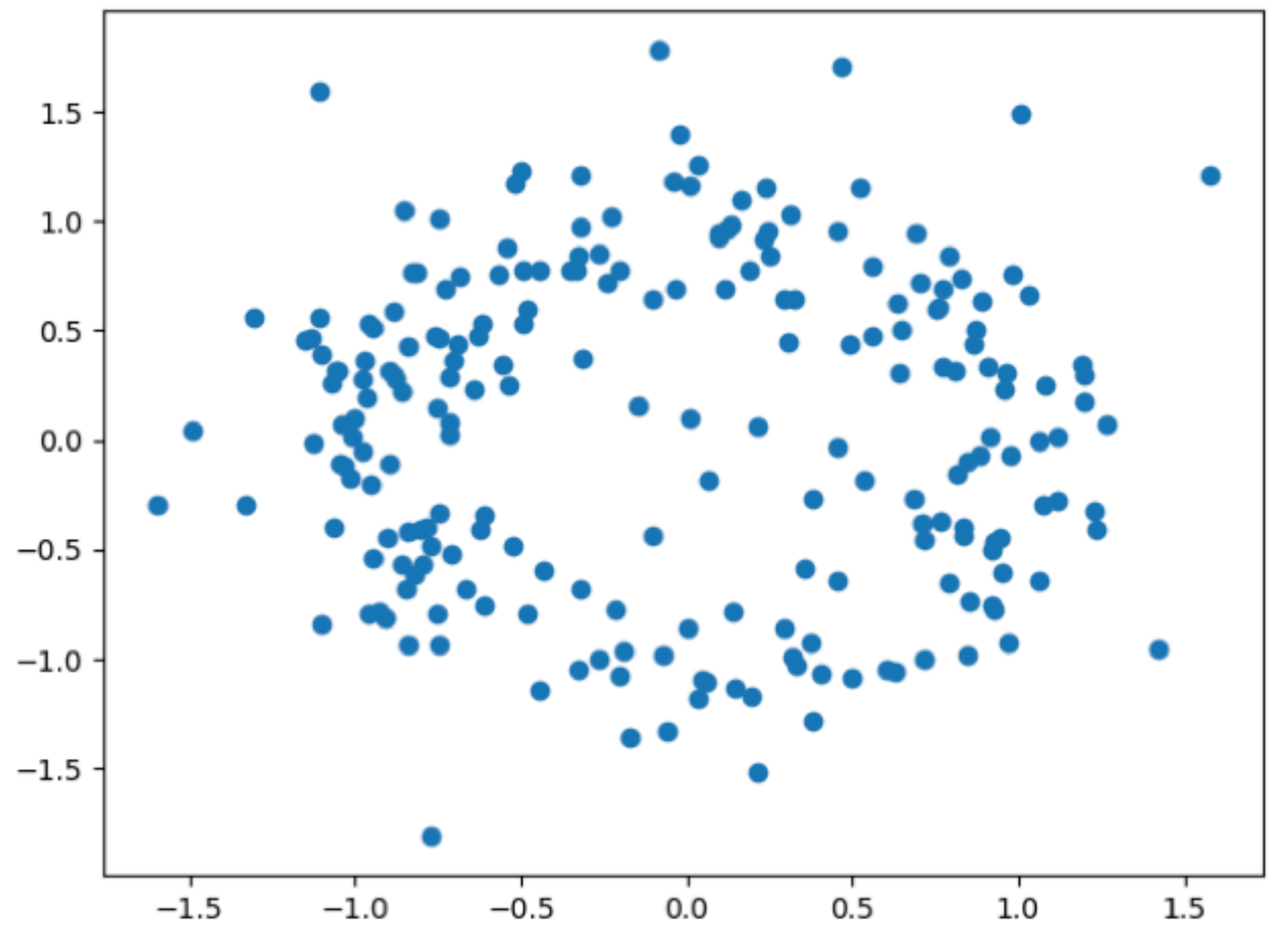


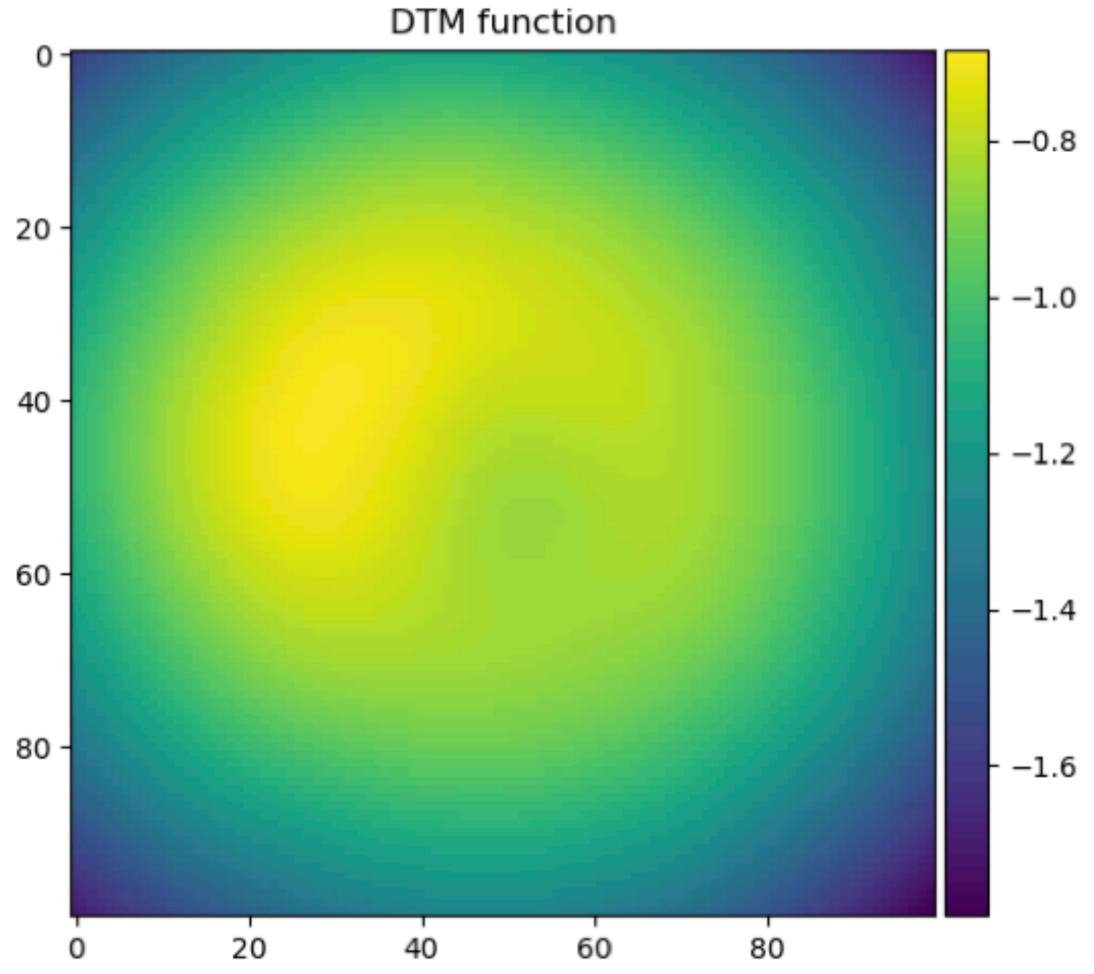
Information Maximizing Filtrations

Can we learn the optimal filtration?

To appear- with Biagetti, Yip, van der Schaar, et. al.

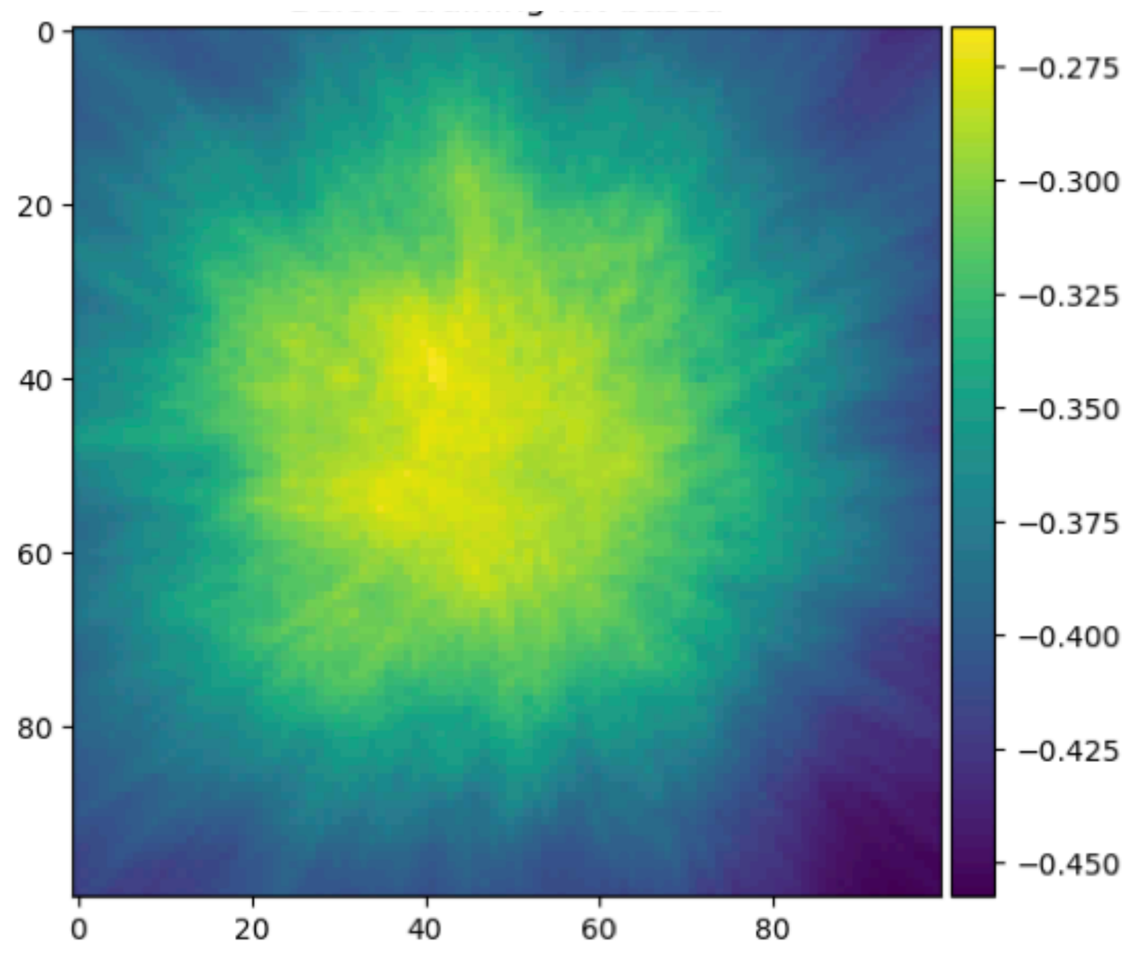
Point cloud

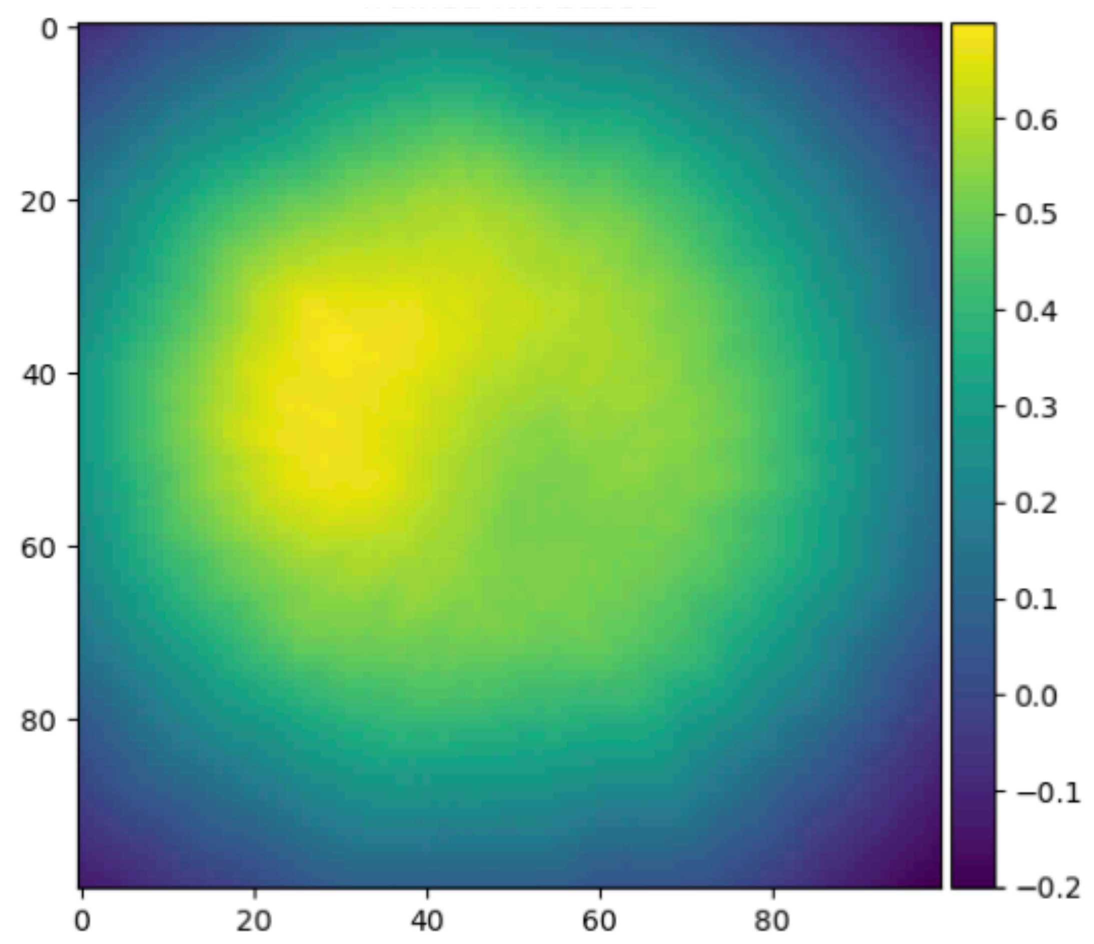




The distance to measure function

Untrained filtration function





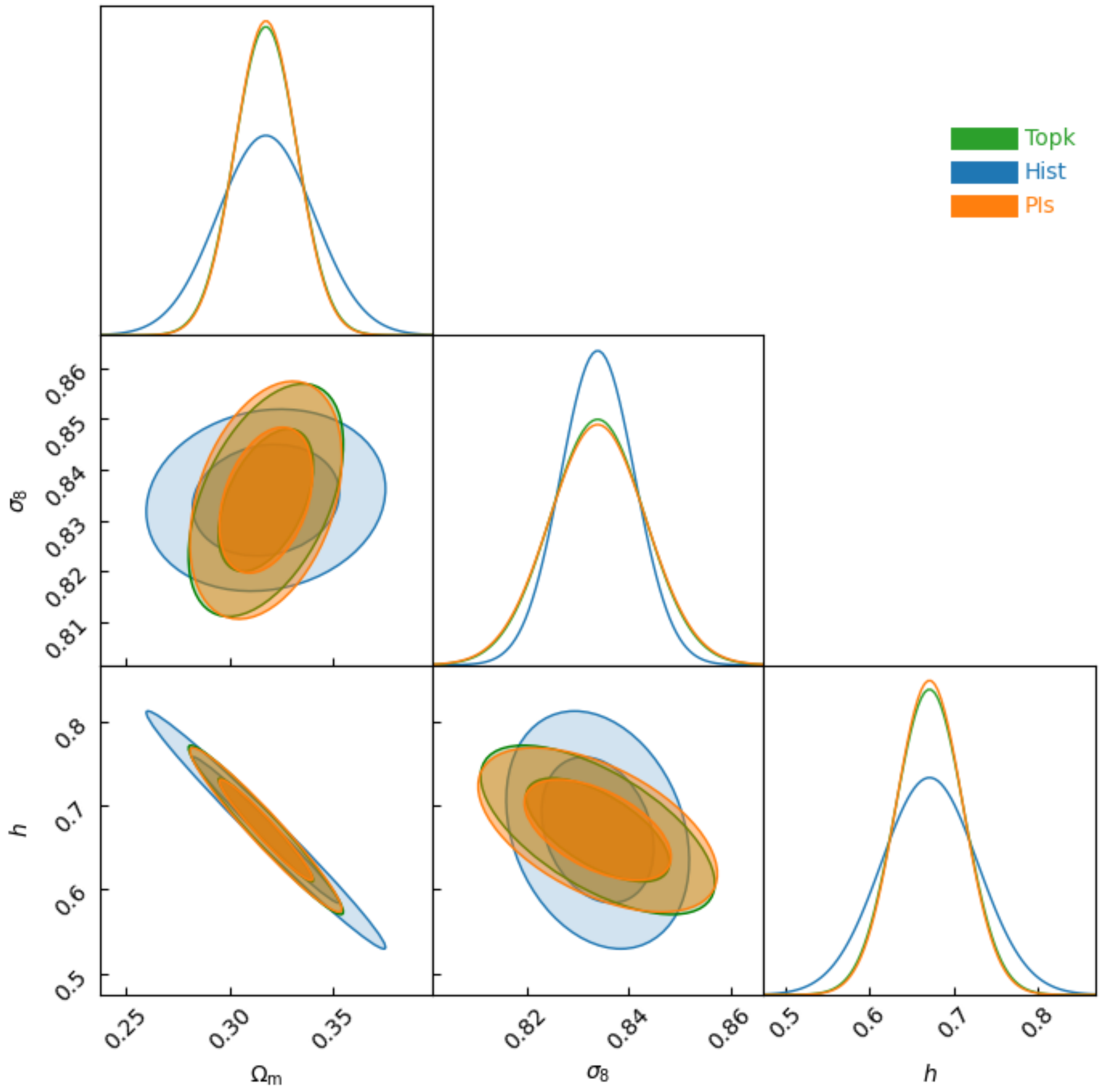
Trained filtration function

Information Maximizing Persistent Homology

Proof of concept using galaxy catalogs

To appear- with Biagetti, Yip, van der Schaar, et. al.

Sancho galaxy catalogs developed by Biagetti et. al.



Conclusion and Outlook

- Persistent Homology combined with Power spectrum and Bispectrum gives more constrained contours.

Conclusion and Outlook

- Persistent Homology combined with Power spectrum and Bispectrum gives more constrained contours.
- Can we further improve these contours by using the IMPH -
  - Is there an information maximizing filtration?
  - Can other persistence summaries give more information?
  - Why are parameter degeneracies for PH statistics different from combined power spectrum and bispectrum statistics?
  - Coming soon in our next paper.

Thank you!