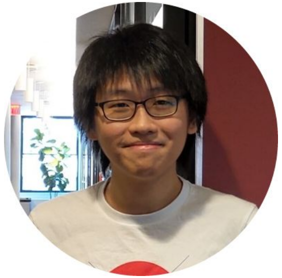
June 2022, Trieste

Constraining Cosmology with Persistent Homology

Matteo Biagetti, Alex Cole, Gary Shiu and Jacky Yip



IFPU/SISSA

University of Amsterdam

University of Wisconsin-Madison

Overview of This Talk

Persistent Homology Basics

↓

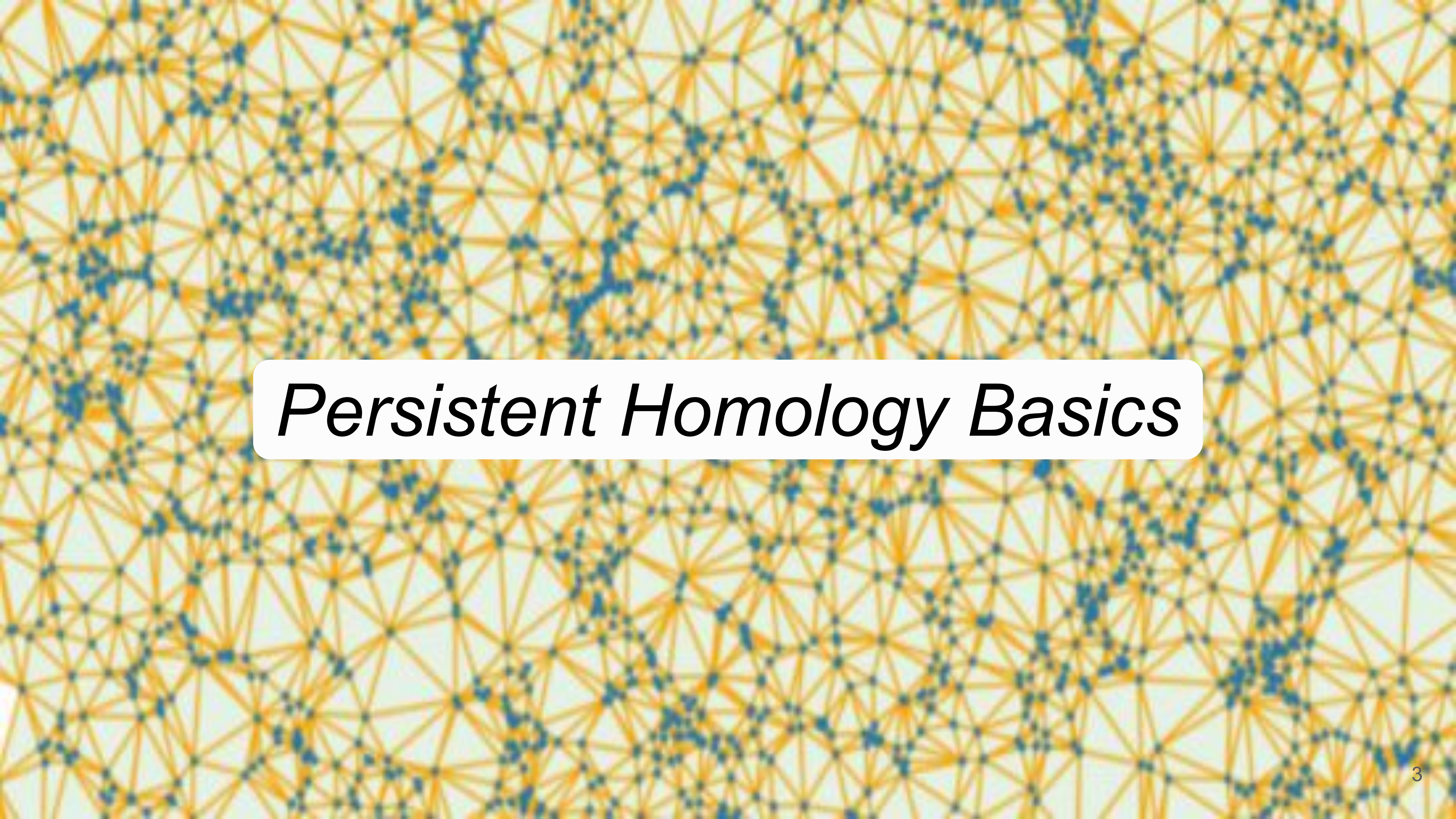
Persistence of the Large Scale Structures

↓

Summary Statistics

↓

Constraining Cosmology



Persistent Homology Basics

Topology and Homology

topologically
equivalent

=

not
topologically
equivalent

≠

Topology and Homology

topologically
equivalent

$=$

not
topologically
equivalent

$\neq$

$$H_i(X)$$

i^th homology group of X

$$= \left\{ \begin{array}{c} \text{i-dim} \\ \text{holes} \end{array} \right\}$$

Betti number

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

Topology and Homology

topologically equivalent

$=$

not topologically equivalent

$\neq$

1 hole

1 hole

3 holes

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

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

$$b_0 = 1 \quad \text{(\# of connected components)}$$

$$b_2 = 0 \quad \text{(\# of voids)}$$

$$H_i(X)$$

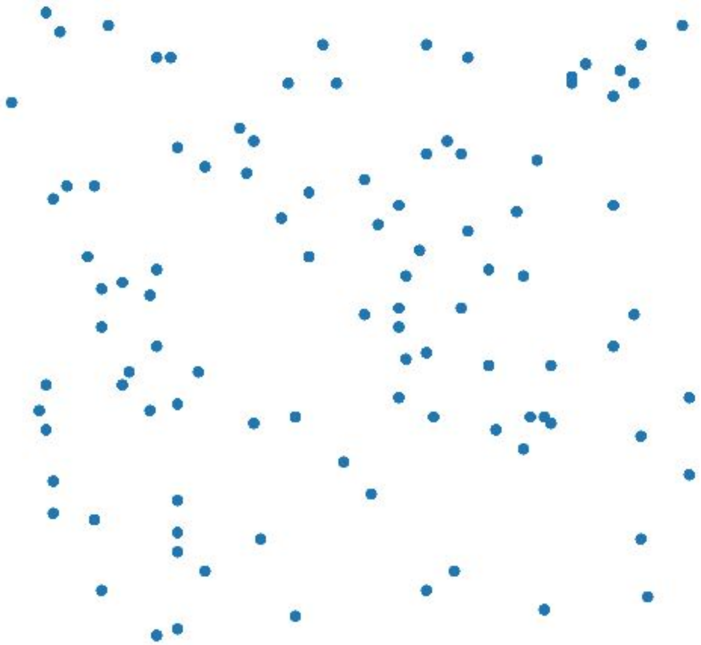
i^{th} homology group of X

$$= \left\{ \begin{array}{c} \text{i-dim} \\ \text{holes} \end{array} \right\}$$

Betti number

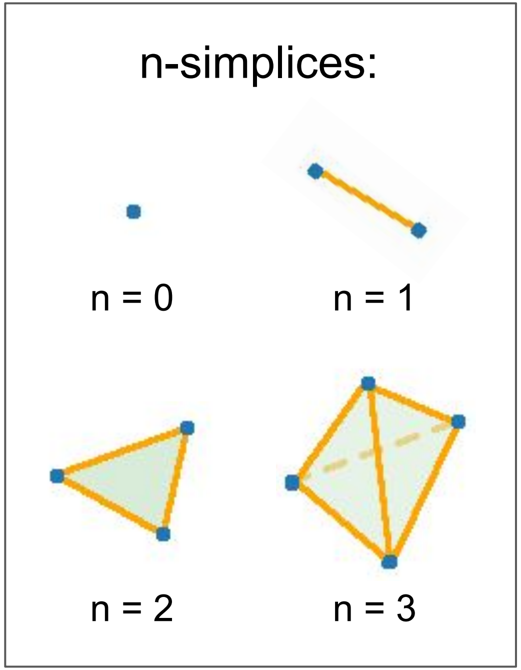
$$b_i = \text{rank}(H_i)$$

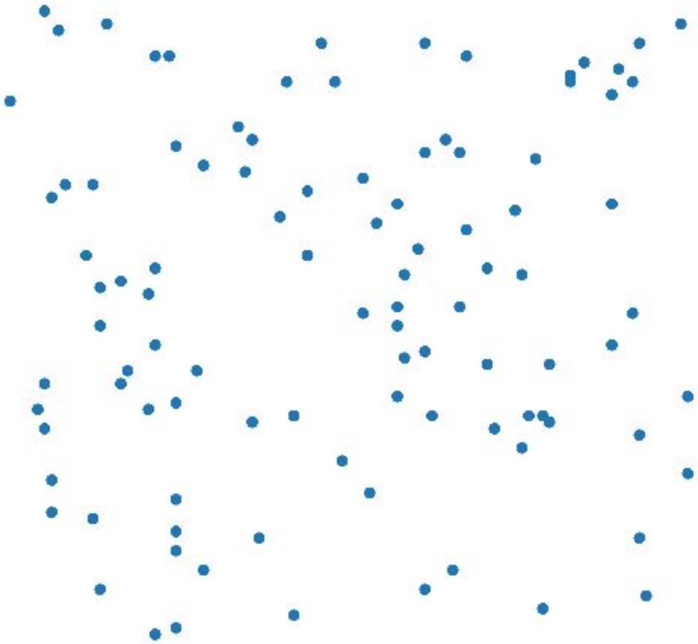
Homology of a Point Cloud



$$b_0 = 100$$
$$b_1 = 0$$

Adding Simplices



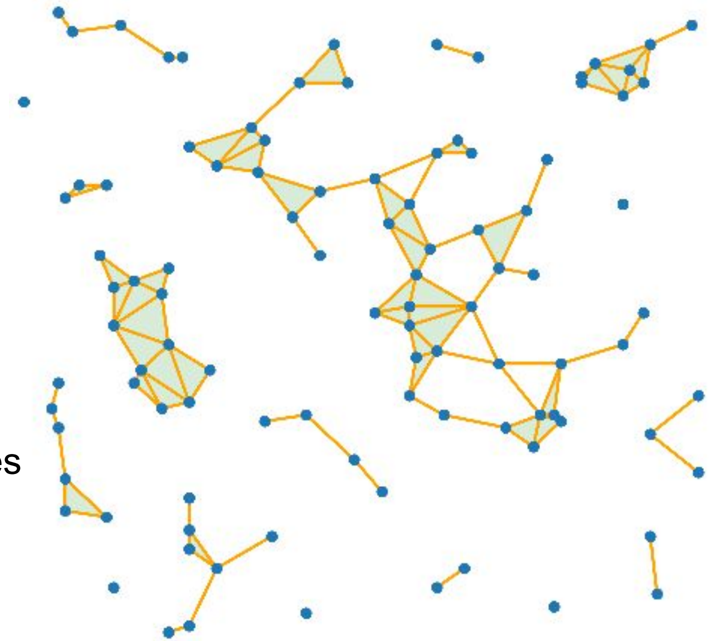


$$b_0 = 100$$
$$b_1 = 0$$

Changing Homology Across Length Scales

$\nu = 5$

- Length scale parameter
- Decides which simplices are added along with a set of rules

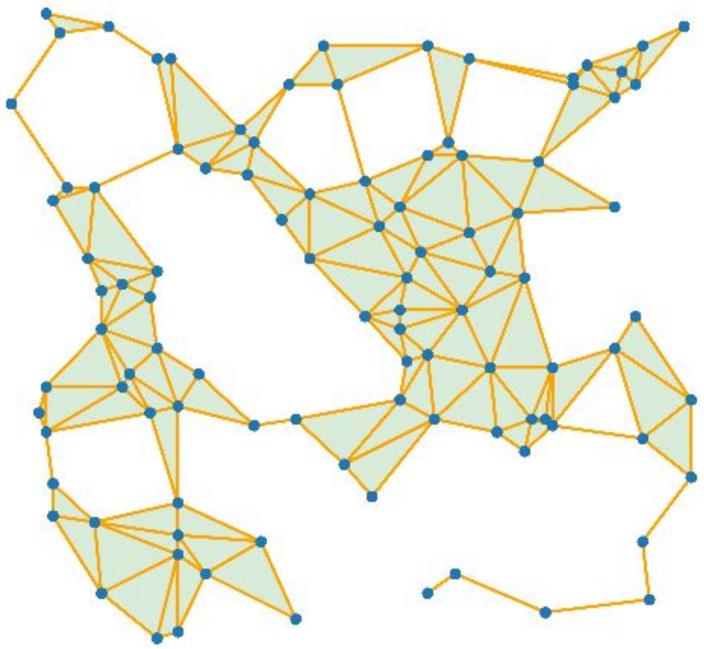


$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{\text{a simplicial complex}}$$

$$b_0 = 17$$
$$b_1 = 5$$

Changing Homology Across Length Scales

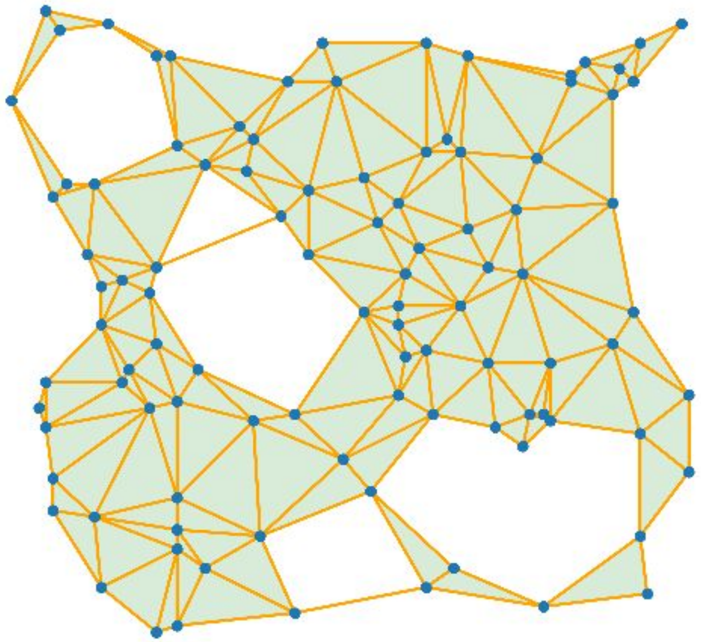
$\nu = 8$



$$b_0 = 1$$
$$b_1 = 7$$

Changing Homology Across Length Scales

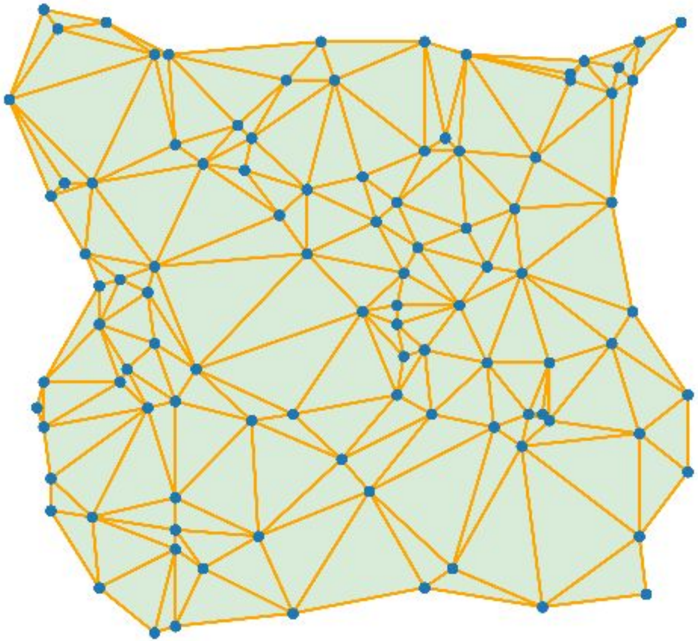
$\nu = 10$



$$b_0 = 1$$
$$b_1 = 5$$

Changing Homology Across Length Scales

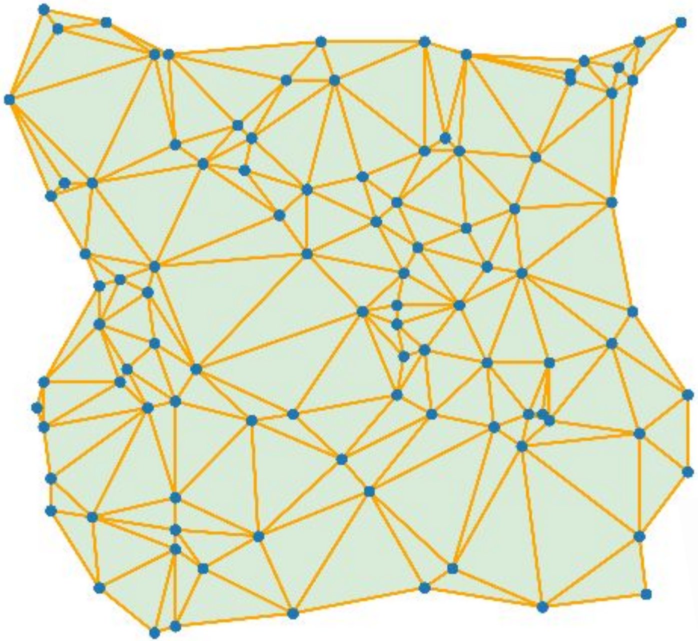
$\nu = 13$



$b_0 = 1$
$b_1 = 0$

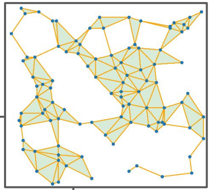
Changing Homology Across Length Scales

$$\nu = 13$$



$$b_0 = 1$$
$$b_1 = 0$$

Tracking Persistent Features

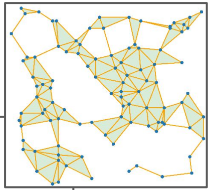


Apart from

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

describe the topological object as a whole

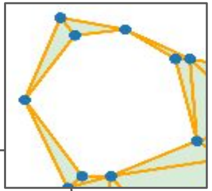
Tracking Persistent Features



Apart from

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

describe the topological object as a whole

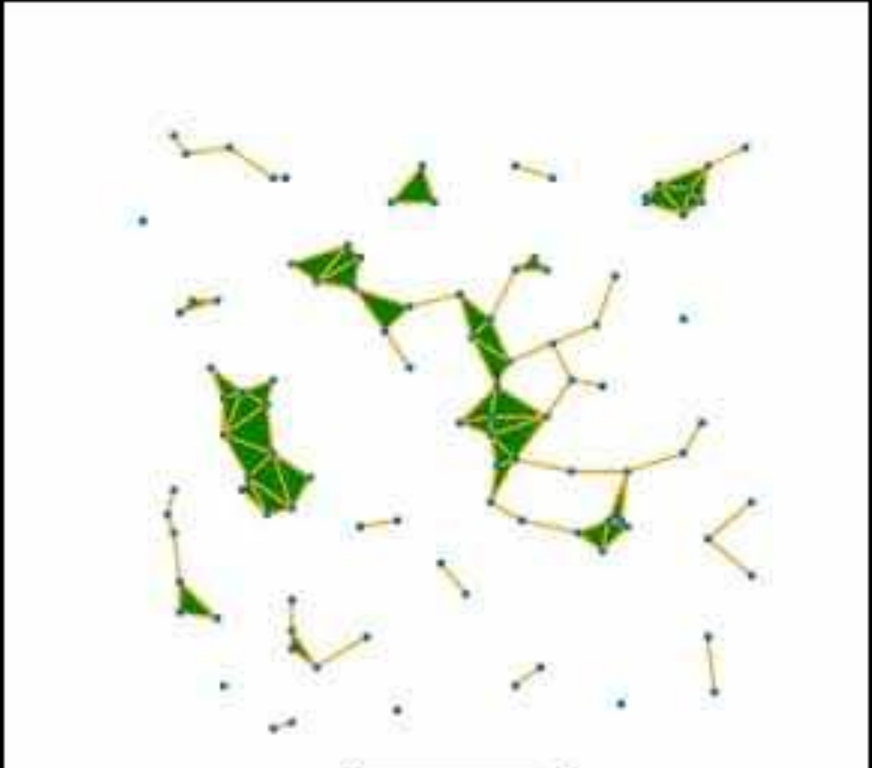


we can

Track individual features

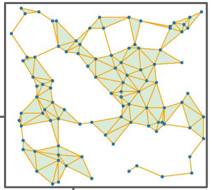
(n-dimensional holes / n-cycles)

that persist across scales



$\nu = 4.87$

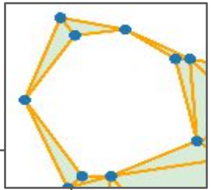
Tracking Persistent Features



Apart from

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

describe the topological object as a whole



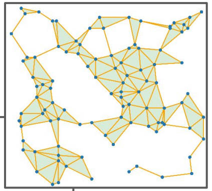
we can

Track individual features

(n-dimensional holes / n-cycles)

that persist across scales

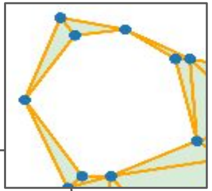
Tracking Persistent Features



Apart from

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

describe the topological object as a whole



we can

Track individual features

(n-dimensional holes / n-cycles)

that persist across scales

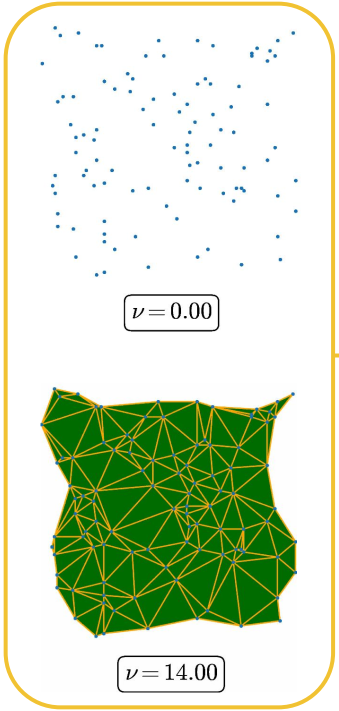
$$(\nu_{\text{birth}}, \nu_{\text{death}})$$

for each feature

Tracking Persistent Features

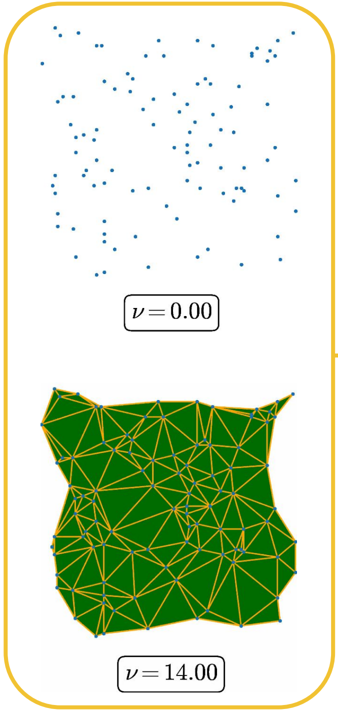
ID	Dimension	ν_{birth}	ν_{death}
1	0	0	0.5
2	0	0	0.707
3	0	0	1
…	…	…	…
100	0	0	inf
101	1	8.732	8.733
…	…	…	…
178	1	7.632	12.029
179	1	8.485	12.889
180	1	6.718	12.905

Tracking Persistent Features

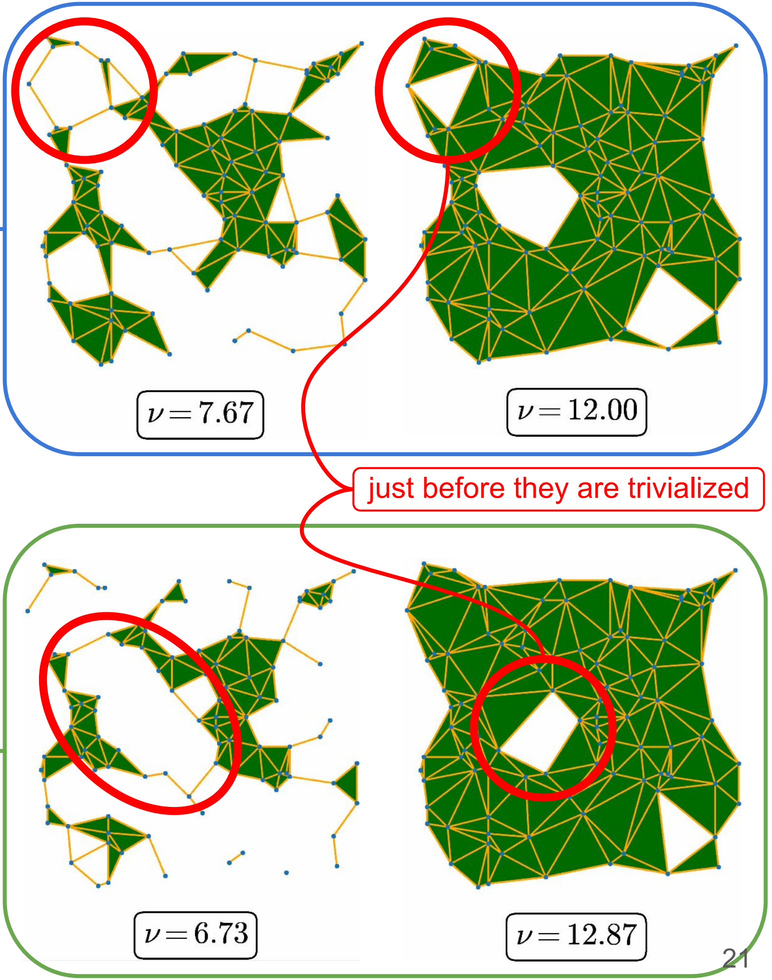


ID	Dimension	ν_{birth}	ν_{death}
1	0	0	0.5
2	0	0	0.707
3	0	0	1
...	...	...	...
100	0	0	inf
101	1	8.732	8.733
...	...	...	...
178	1	7.632	12.029
179	1	8.485	12.889
180	1	6.718	12.905

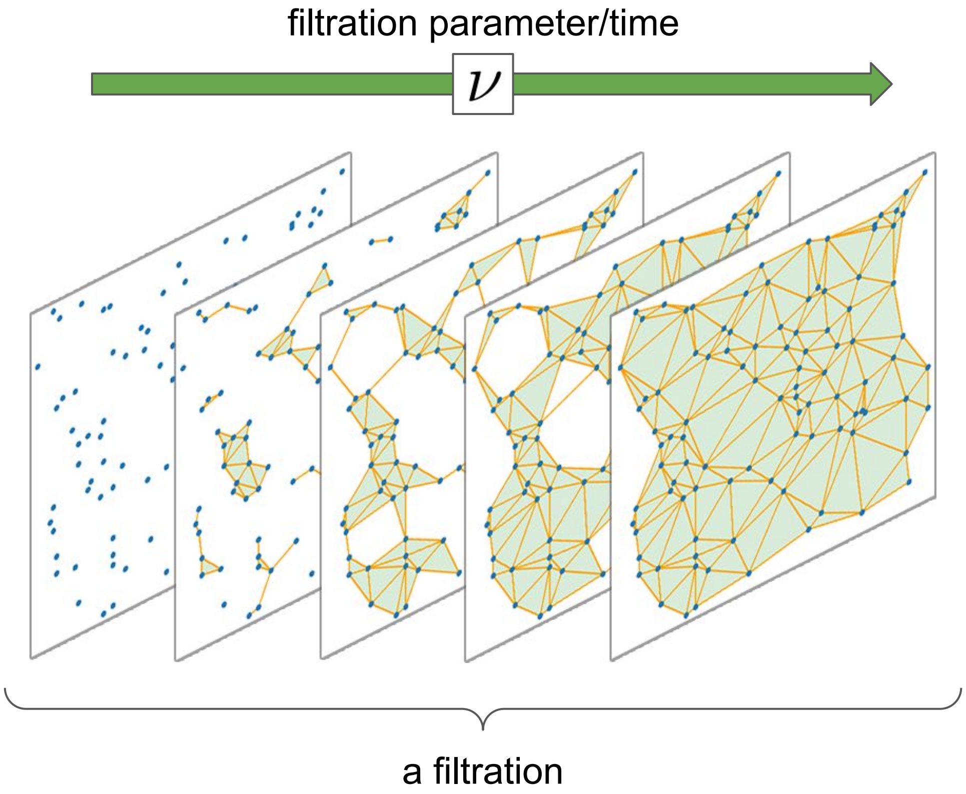
Tracking Persistent Features



ID	Dimension	ν_{birth}	ν_{death}
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



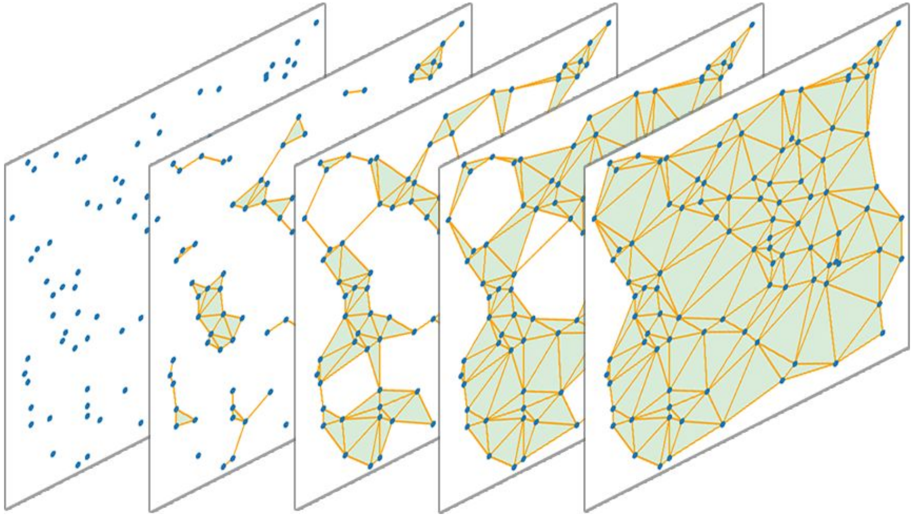
Persistent Homology



Persistent Homology

filtration parameter/time

ν

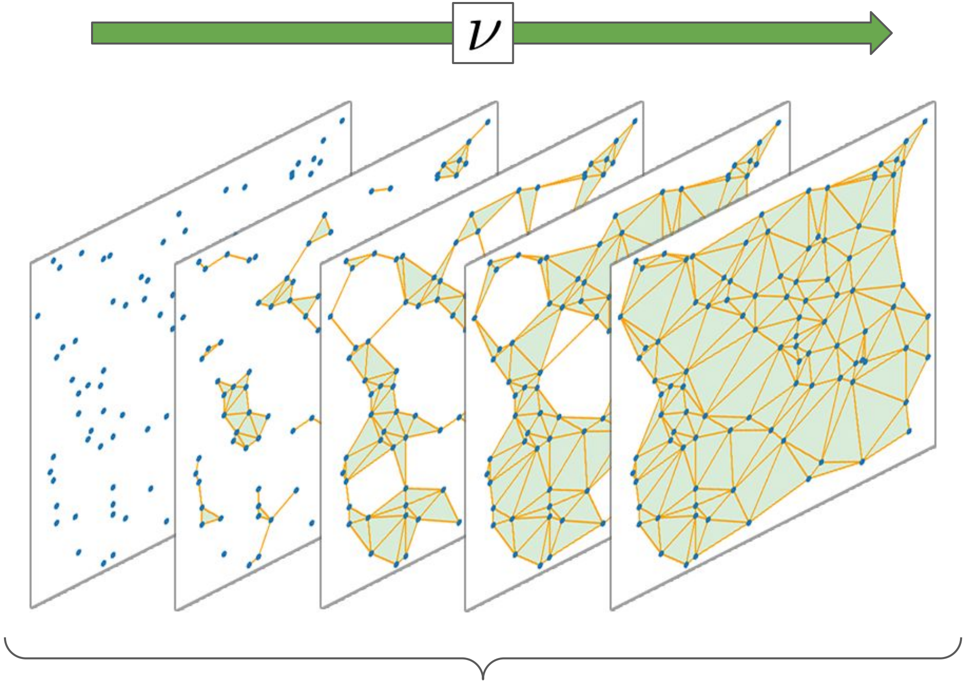


a filtration

ID	Dimension	ν_{birth}	ν_{death}
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

Persistent Homology

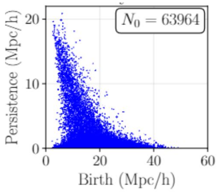
filtration parameter/time

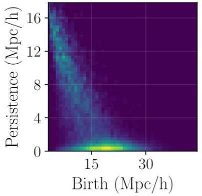


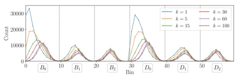
a filtration

ID	Dimension	ν_{birth}	ν_{death}
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

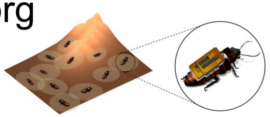
persistence diagrams, persistence images, data vectors…





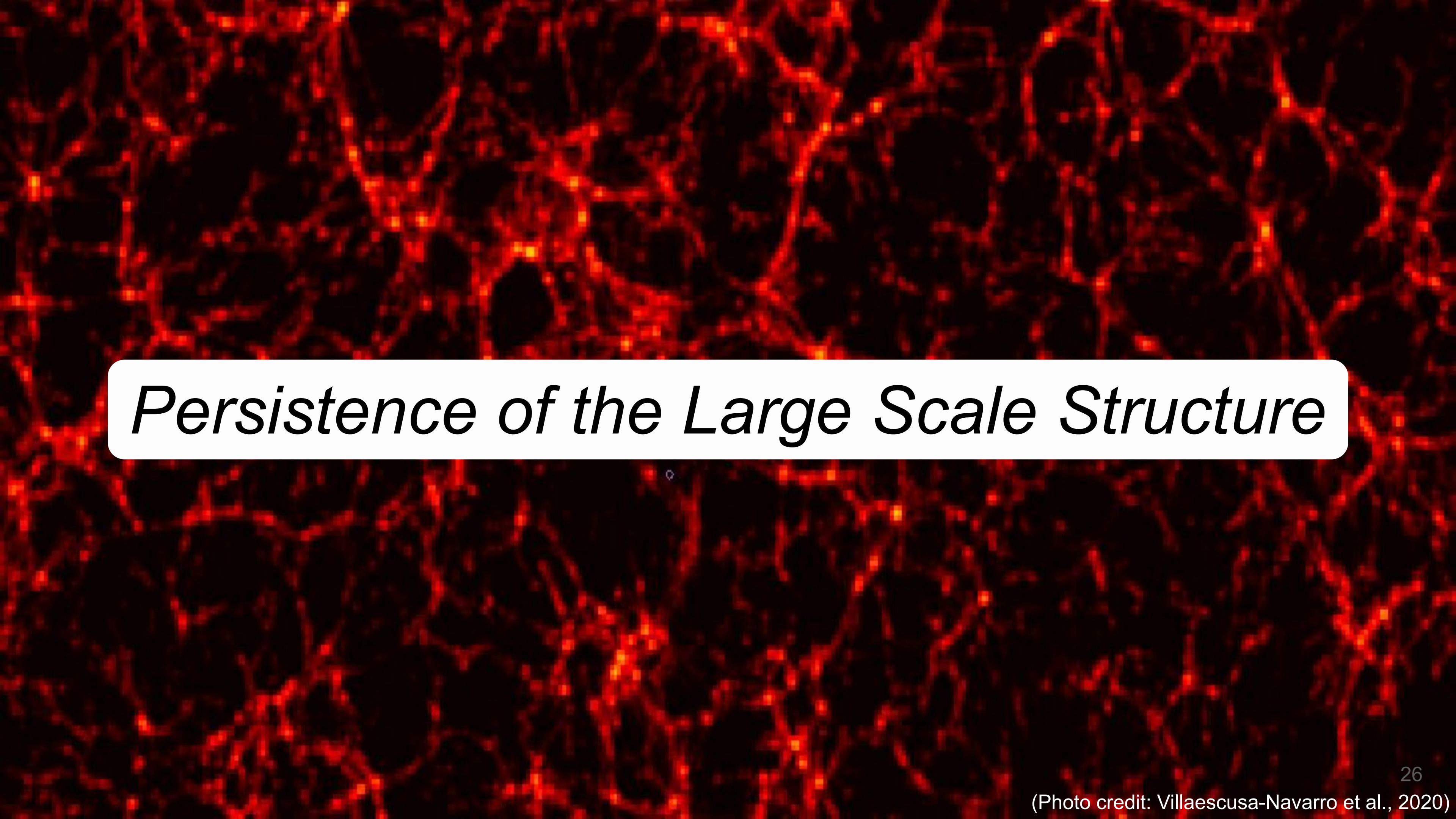


Applications of Persistent Homology in Other Fields

Field	*Example?*	*What is the point cloud?*
Computer vision	Shape characterization	Points sampled on the shape's surface
Neuroscience	Relating neuronal activities to image stimulations	Distribution of activities in the brain
Robotics	Mapping unknown environments by cockroaches	Positions of cyborg cockroaches 
Sensor networks	Signal coverage	Locations of sensors

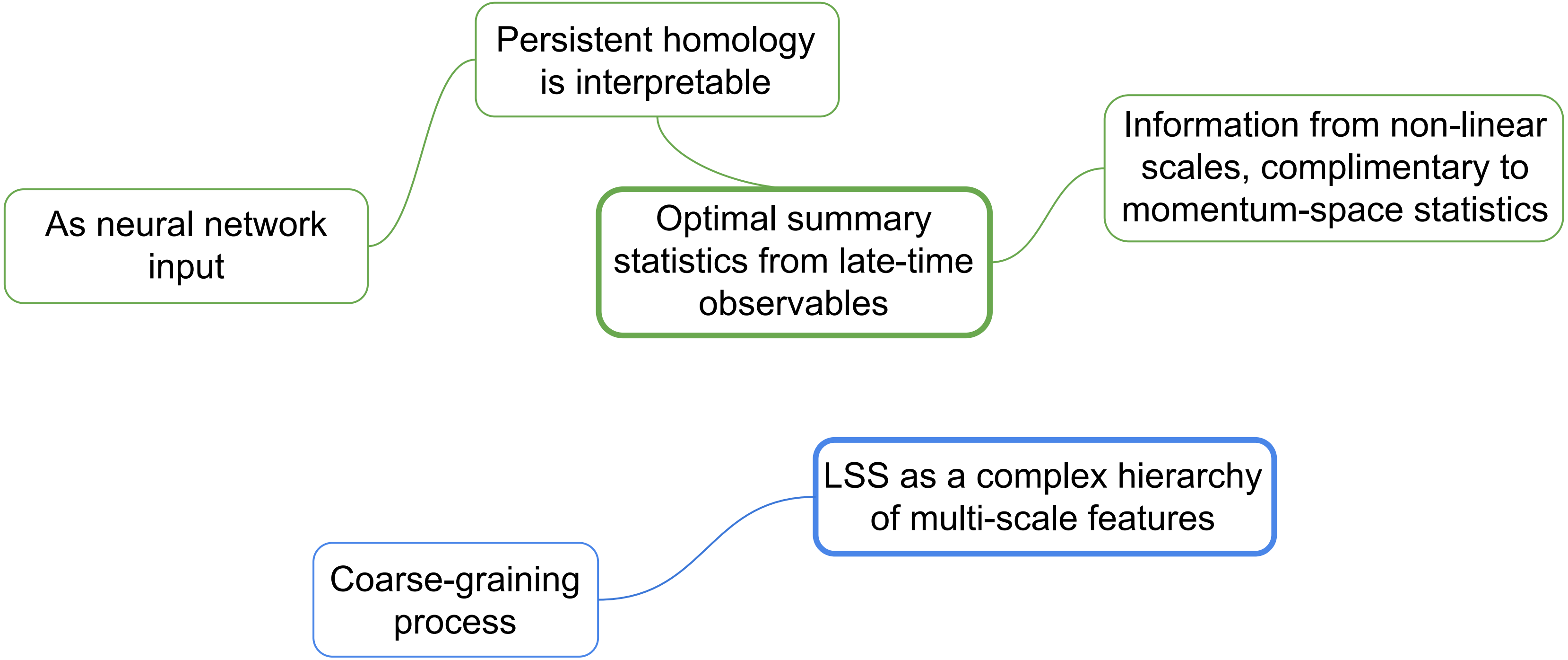
and bioinformation, signal analysis, genomics, language processing…

Persistence of the Large Scale Structure



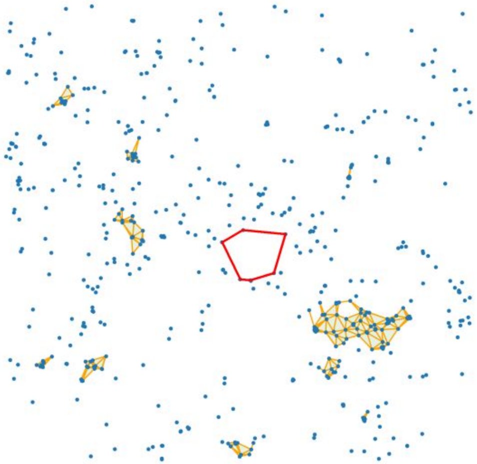
(Photo credit: Villaescusa-Navarro et al., 2020)

Applying Persistent Homology to the Large Scale Structure



Applying Persistent Homology to the Large Scale Structure

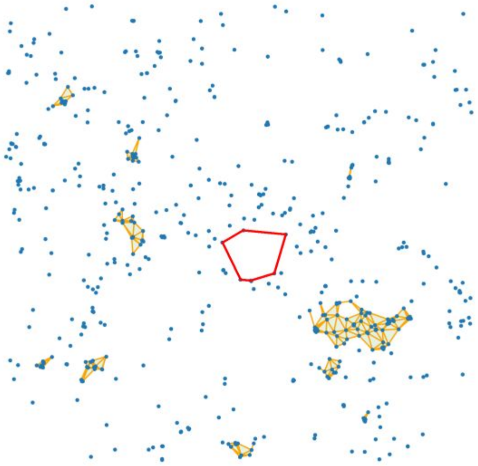
More on coarse-graining:



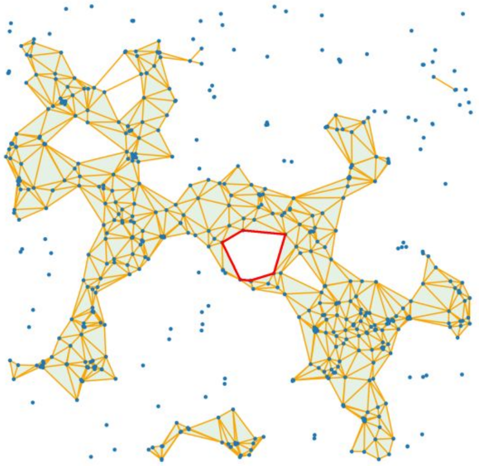
$\nu = 10$

Applying Persistent Homology to the Large Scale Structure

More on coarse-graining:



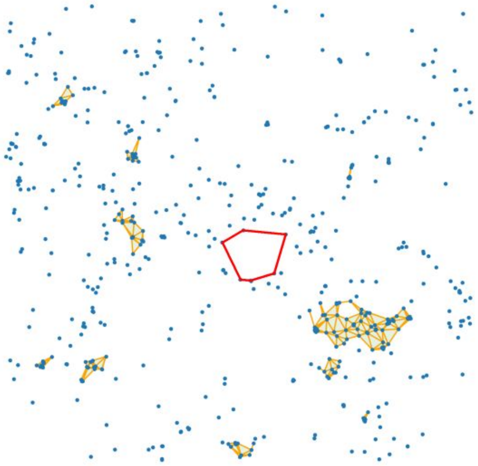
$\nu = 10$

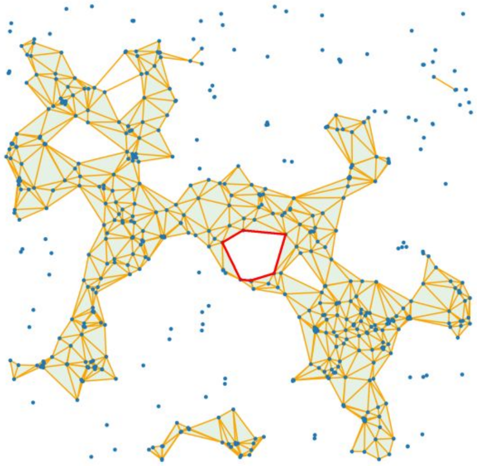


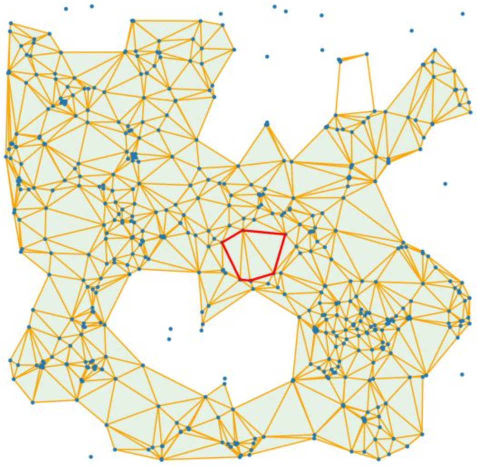
$\nu = 20$

Applying Persistent Homology to the Large Scale Structure

More on coarse-graining:





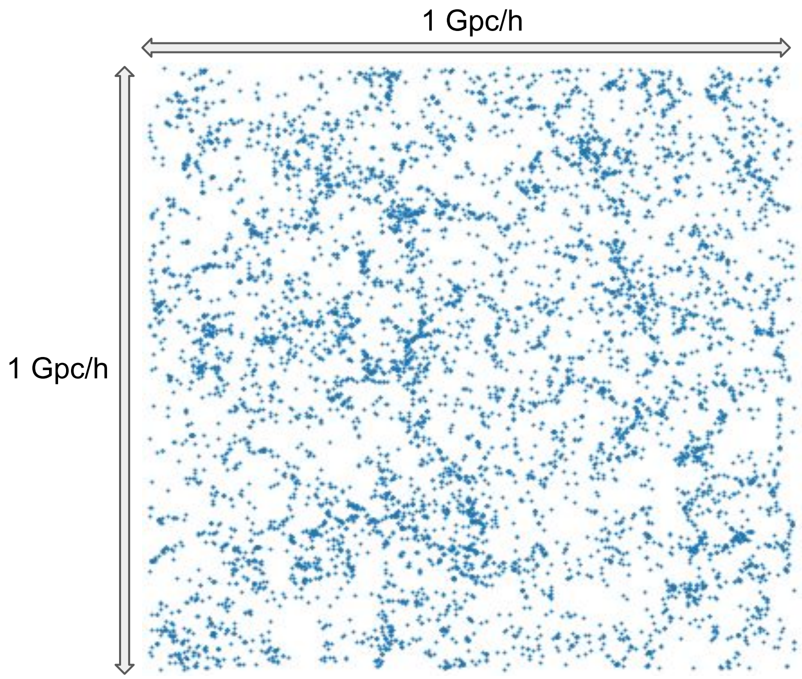


Defining the Filtration

1) The simplicial complex from which the n-simplices are taken

Defining the Filtration

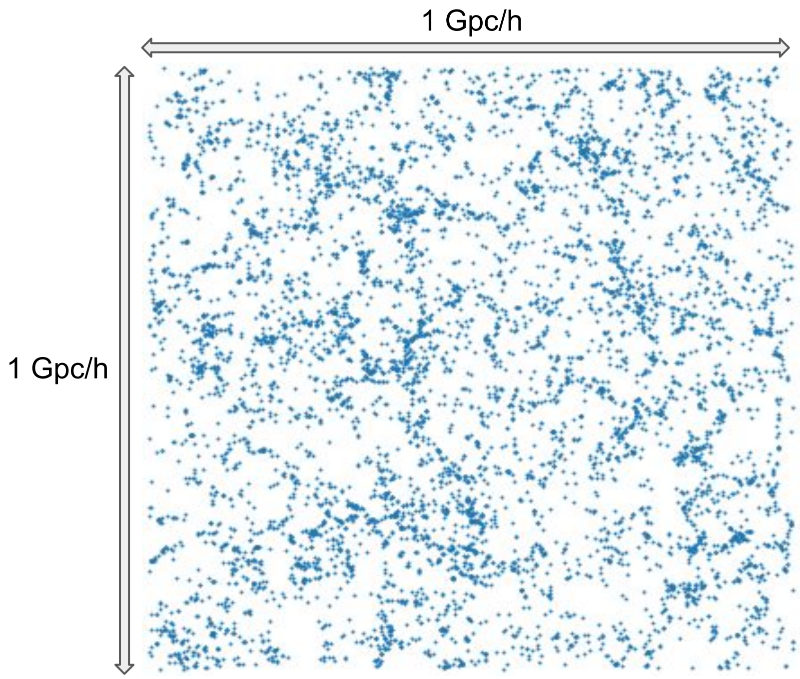
1) The simplicial complex from which the n-simplices are taken



Galaxy distribution (50 Mpc/h thick)

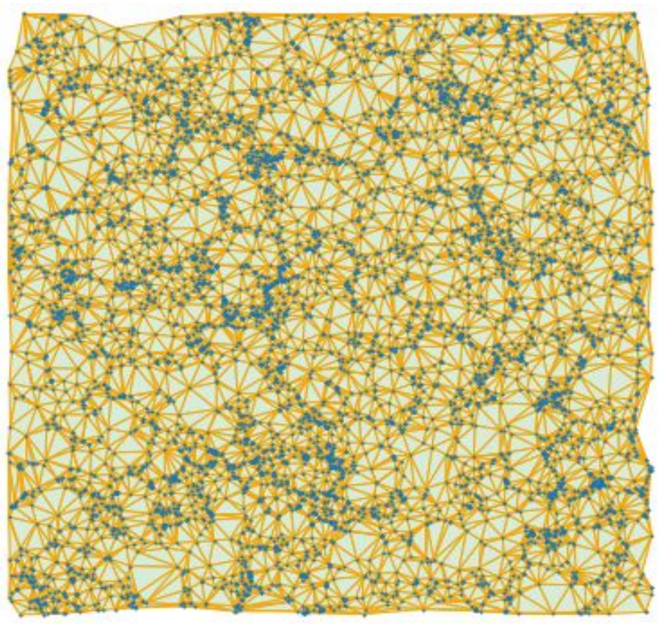
Defining the Filtration

1) The simplicial complex from which the n-simplices are taken



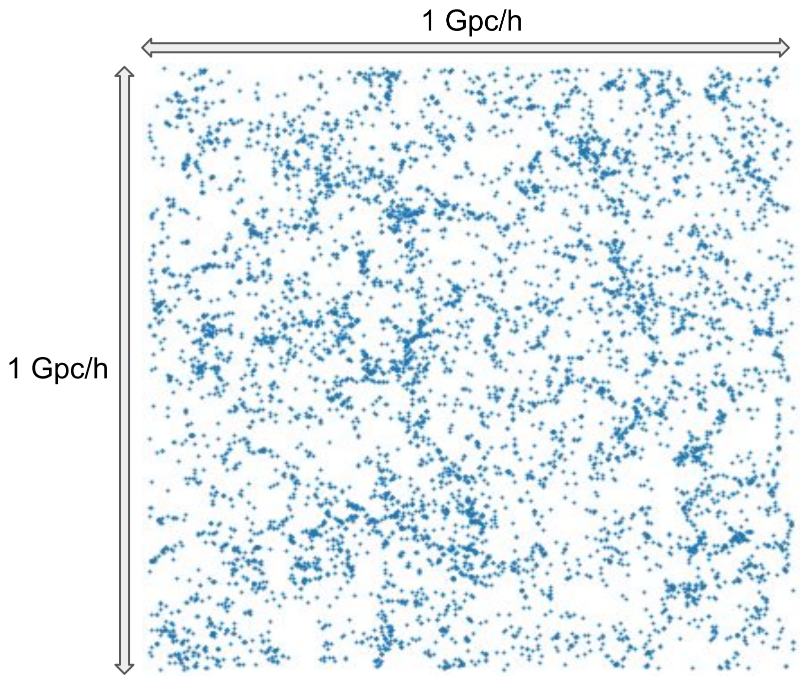
Galaxy distribution (50 Mpc/h thick)

Delaunay triangulation



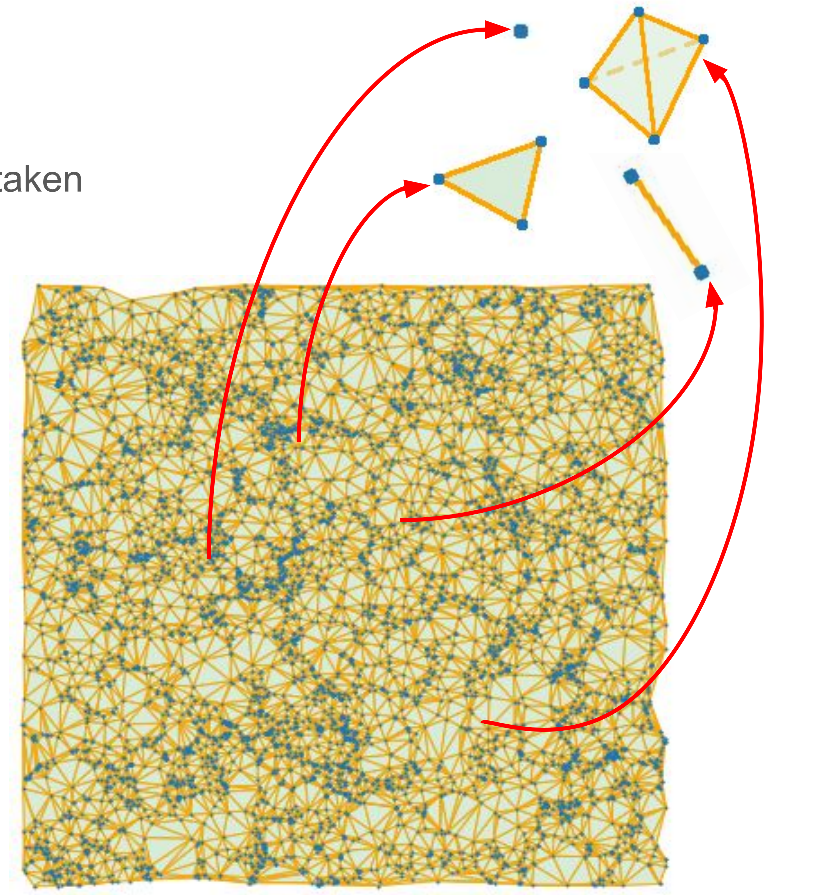
Defining the Filtration

1) The simplicial complex from which the n-simplices are taken



Galaxy distribution (50 Mpc/h thick)

Delaunay triangulation



Defining the Filtration

2) The rules for adding the n-simplices that relates to ν: the Vietoris-Rips filtration

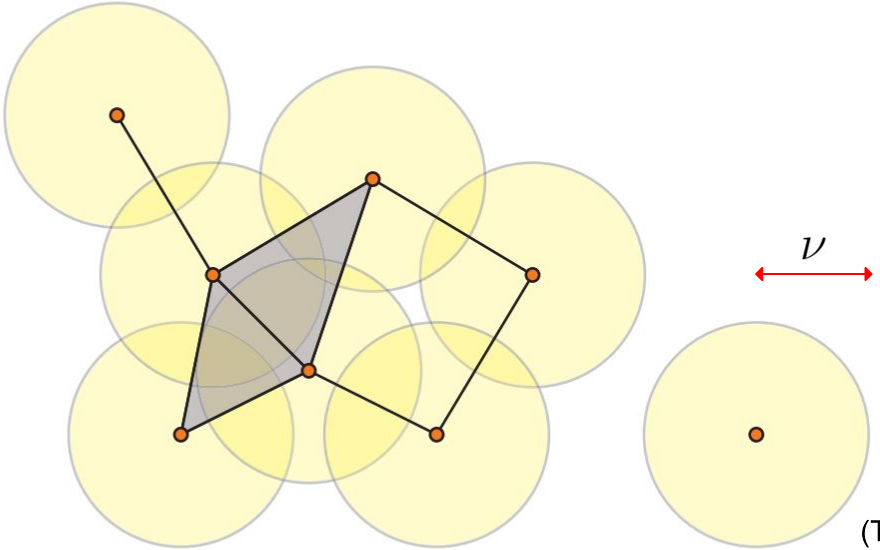
Defining the Filtration

2) The rules for adding the n-simplices that relates to ν: the Vietoris-Rips filtration
    a) All 0-simplices are added at $\nu = 0$.
    b) 1-simplices are added if the vertices are less than or equal to 2ν apart.
    c) Higher-dimensional simplices are added if all the faces are present.

Defining the Filtration

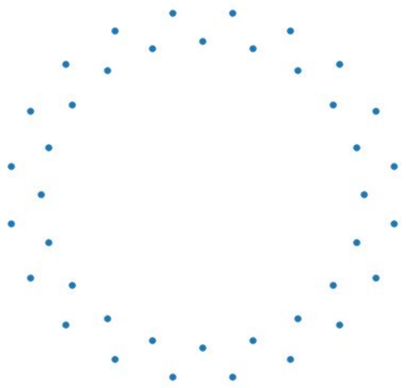
2) The rules for adding the n-simplices that relates to ν: the Vietoris-Rips filtration

   a) All 0-simplices are added at $\nu = 0$.
   b) 1-simplices are added if the vertices are less than or equal to 2ν apart.
   c) Higher-dimensional simplices are added if all the faces are present.



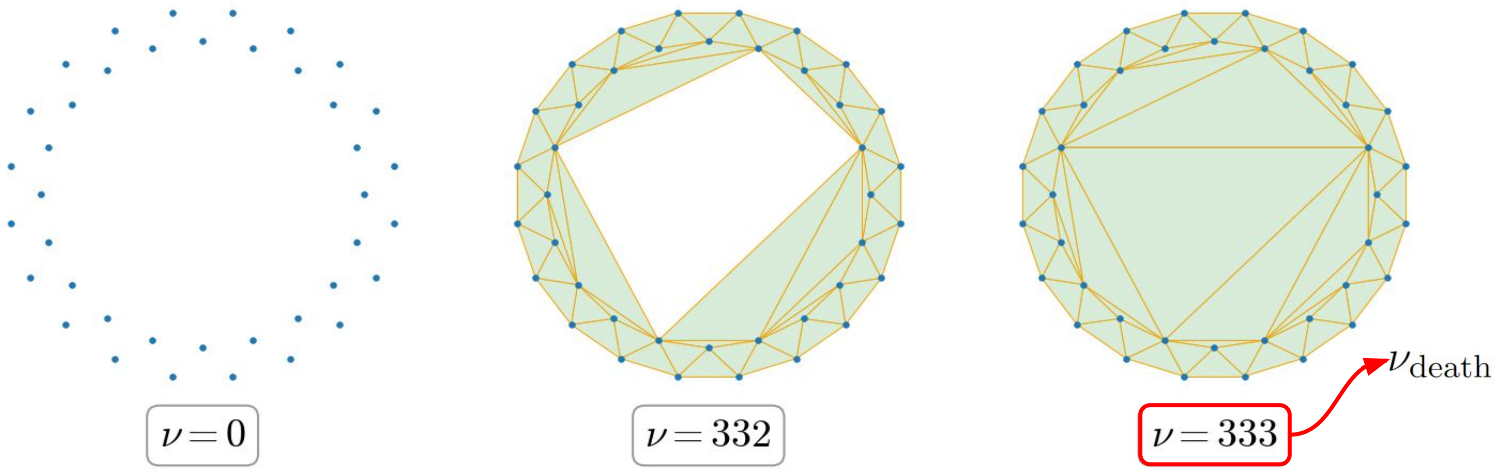
(Topaz et al., 2015)

Problem with Outliers

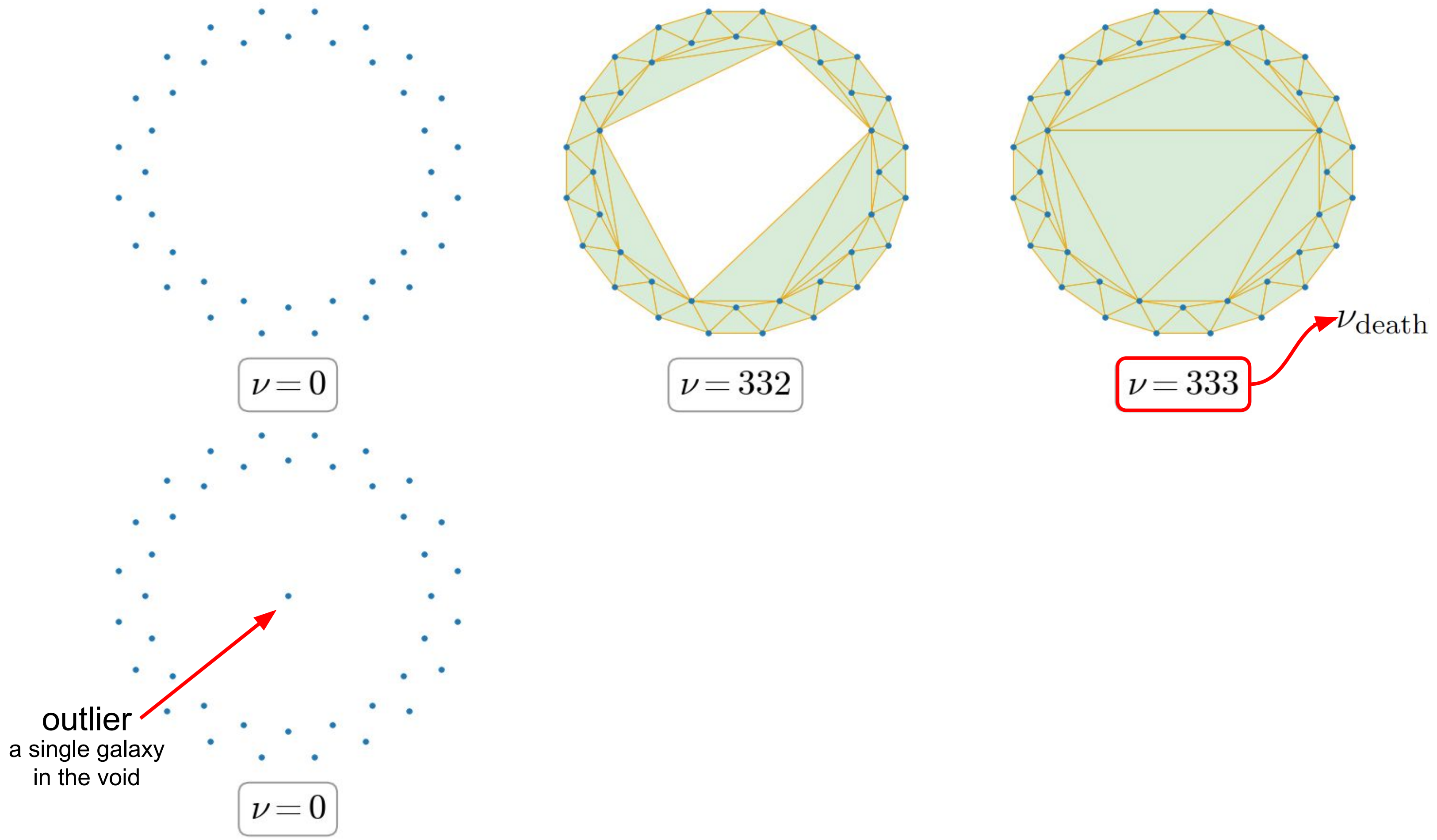


$\nu = 0$

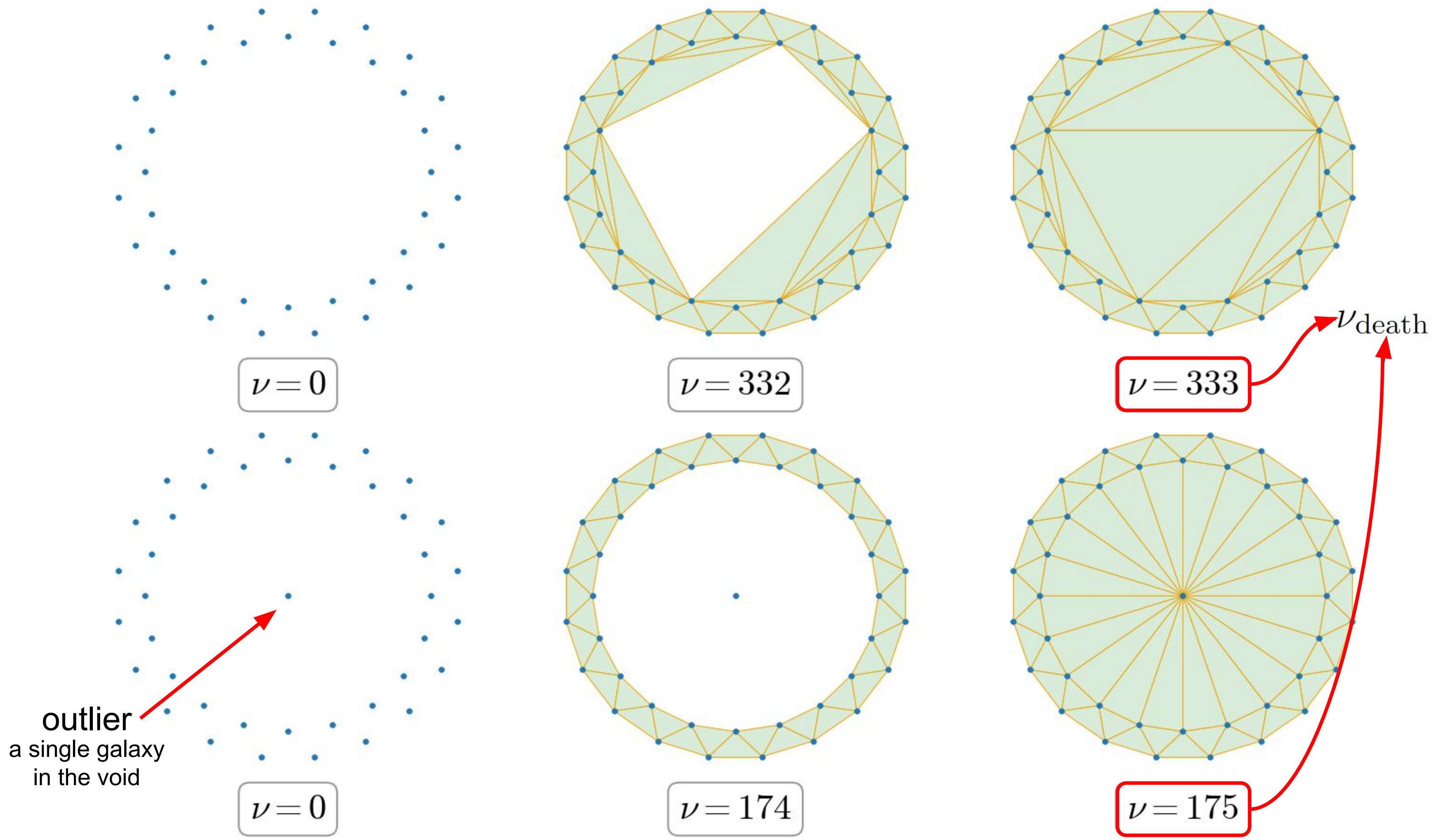
Problem with Outliers



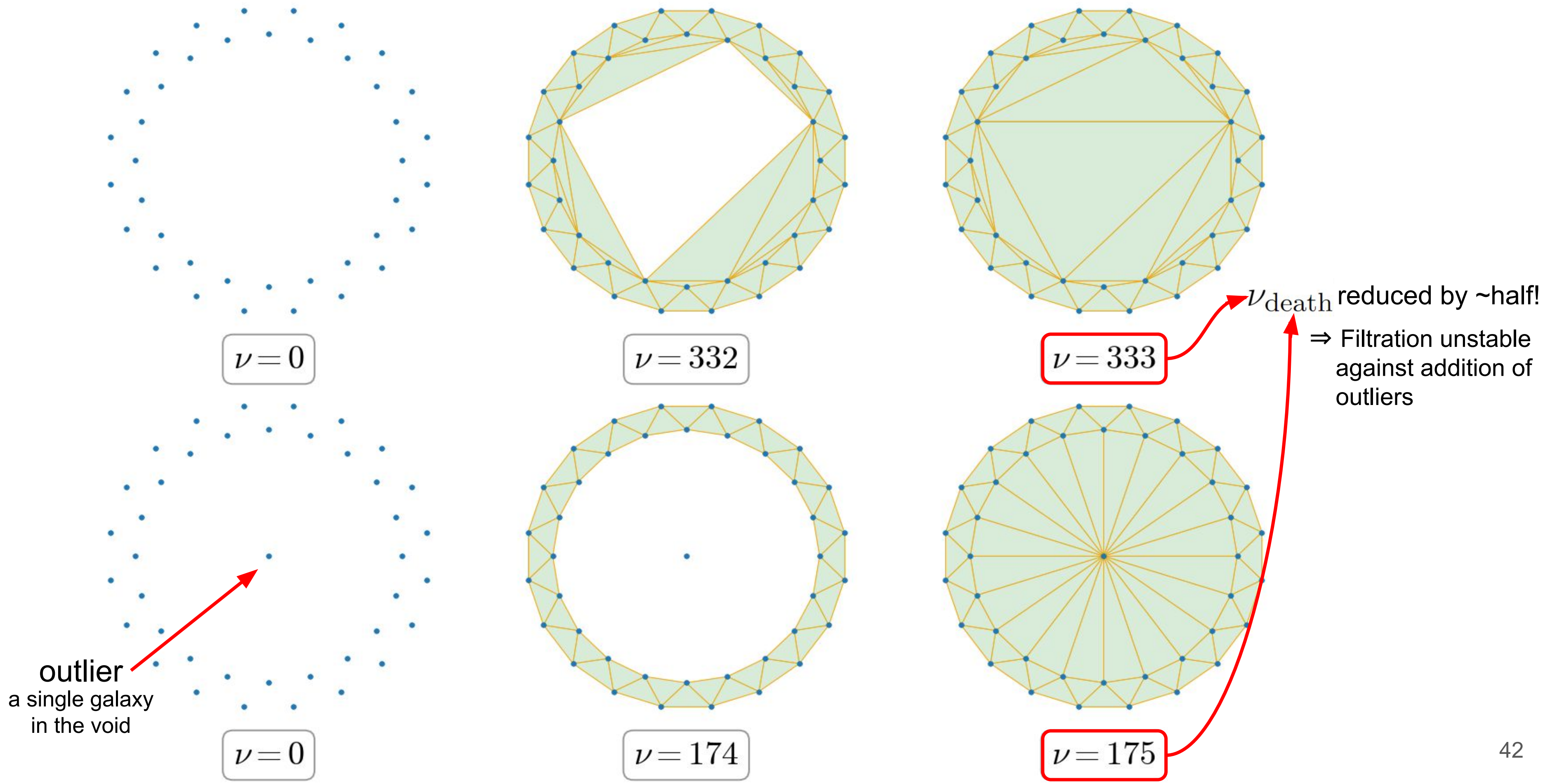
Problem with Outliers



Problem with Outliers



Problem with Outliers



Dealing with Outliers — The Radius Function

a) All 0-simplices are added at $\nu = 0$.
b) 1-simplices are added if the vertices are less than or equal to 2ν apart.
c) Higher-dimensional simplices are added if all the faces are present.

Dealing with Outliers — The Radius Function

a) All 0-simplices are added at $\nu = 0$.

~~b) 1-simplices are added if the vertices are less than or equal to 2ν apart.~~

c) Higher-dimensional simplices are added if all the faces are present.

b) 1-simplices are added if the two relevant balls with radius $r_x(\nu)$ touch or overlap.

$$r_x(\nu) = (\nu^q - f(x)^q)^{1/q}$$

$f(x)$: a function of point x that measures its “unimportance”; impedes the ball’s growth

q: a mixing parameter; $q = 2$

Dealing with Outliers — The Radius Function

a) All 0-simplices are added at $\nu = 0$.

~~b) 1-simplices are added if the vertices are less than or equal to 2ν apart.~~

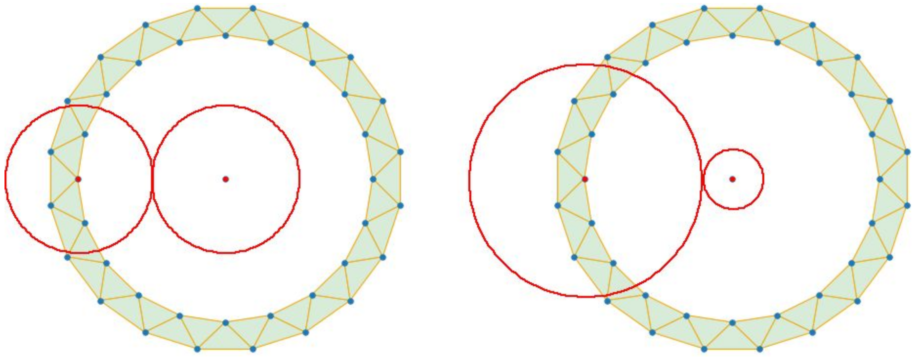
c) Higher-dimensional simplices are added if all the faces are present.

b) 1-simplices are added if the two relevant balls with radius $r_x(\nu)$ touch or overlap.

$$r_x(\nu) = (\nu^q - f(x)^q)^{1/q}$$

$f(x)$: a function of point x that measures its “unimportance”; impedes the ball’s growth

q: a mixing parameter; $q = 2$



uniform ν $\qquad$ $r_x(\nu)$

Dealing with Outliers — The Radius Function

a) All 0-simplices are added at $\nu = \cancel{0}$. $f(x)$

~~b) 1-simplices are added if the vertices are less than or equal to 2ν apart.~~

c) Higher-dimensional simplices are added if all the faces are present.

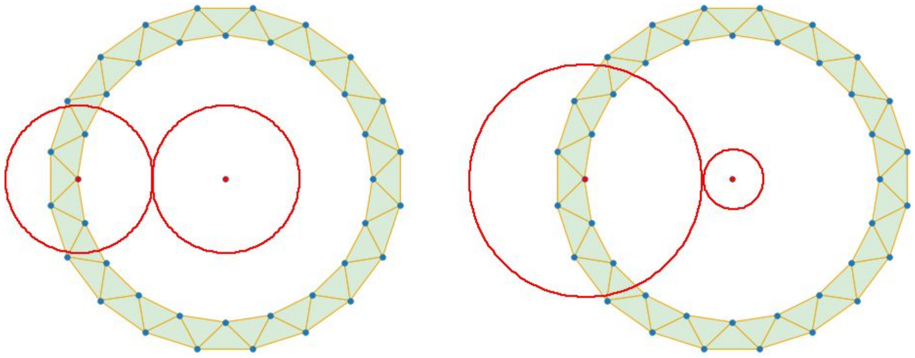
b) 1-simplices are added if the two relevant balls with radius $r_x(\nu)$ touch or overlap.

$$r_x(\nu) = (\nu^q - f(x)^q)^{1/q}$$

$f(x)$: a function of point x that measures its “unimportance”; impedes the ball’s growth

q: a mixing parameter; $q = 2$

uniform ν



$r_x(\nu)$

Dealing with Outliers — The DTM function

What should $f(x)$ be? It measures the “unimportance”, i.e., how much of an outlier x is.

Dealing with Outliers — The DTM function

What should $f(x)$ be? It measures the "unimportance", i.e., how much of an outlier x is.

Employ the Distance-To-Measure (DTM) function:

$$f(x) = \mathrm{DTM}(x) \equiv \left(\frac{1}{k} \sum_{X_i \in \mathcal{N}_k(x)} \|x - X_i\|^p \right)^{1/p}$$

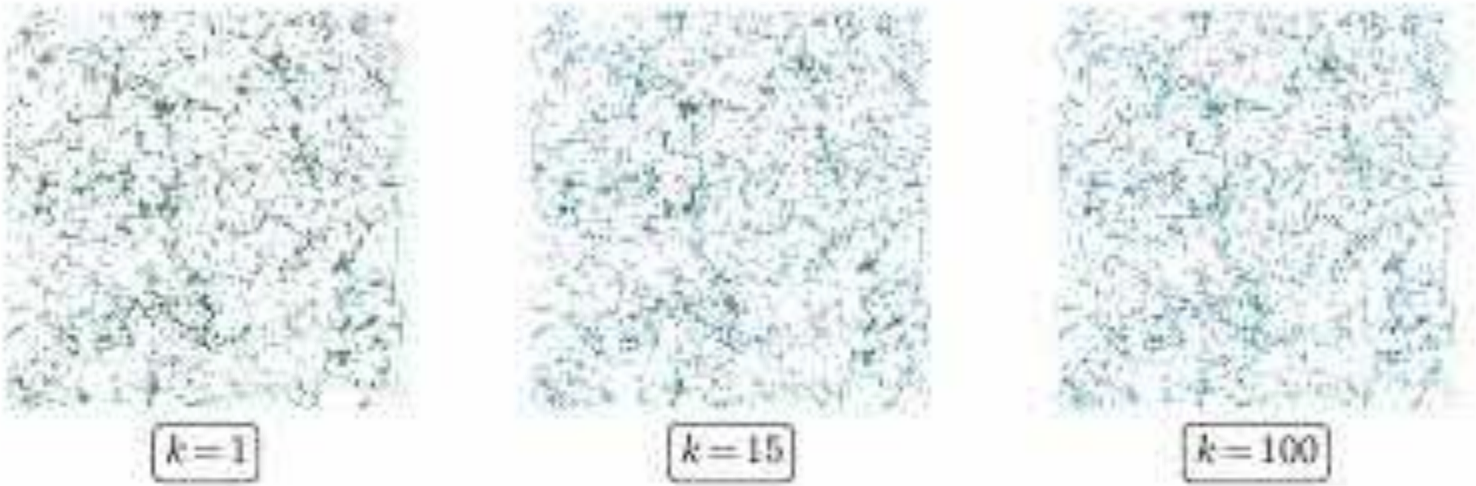
k : # of nearest neighbours
$\mathcal{N}_k(x)$: the set of k-nearest neighbours of x
p : a mixing parameter; $p = 2$

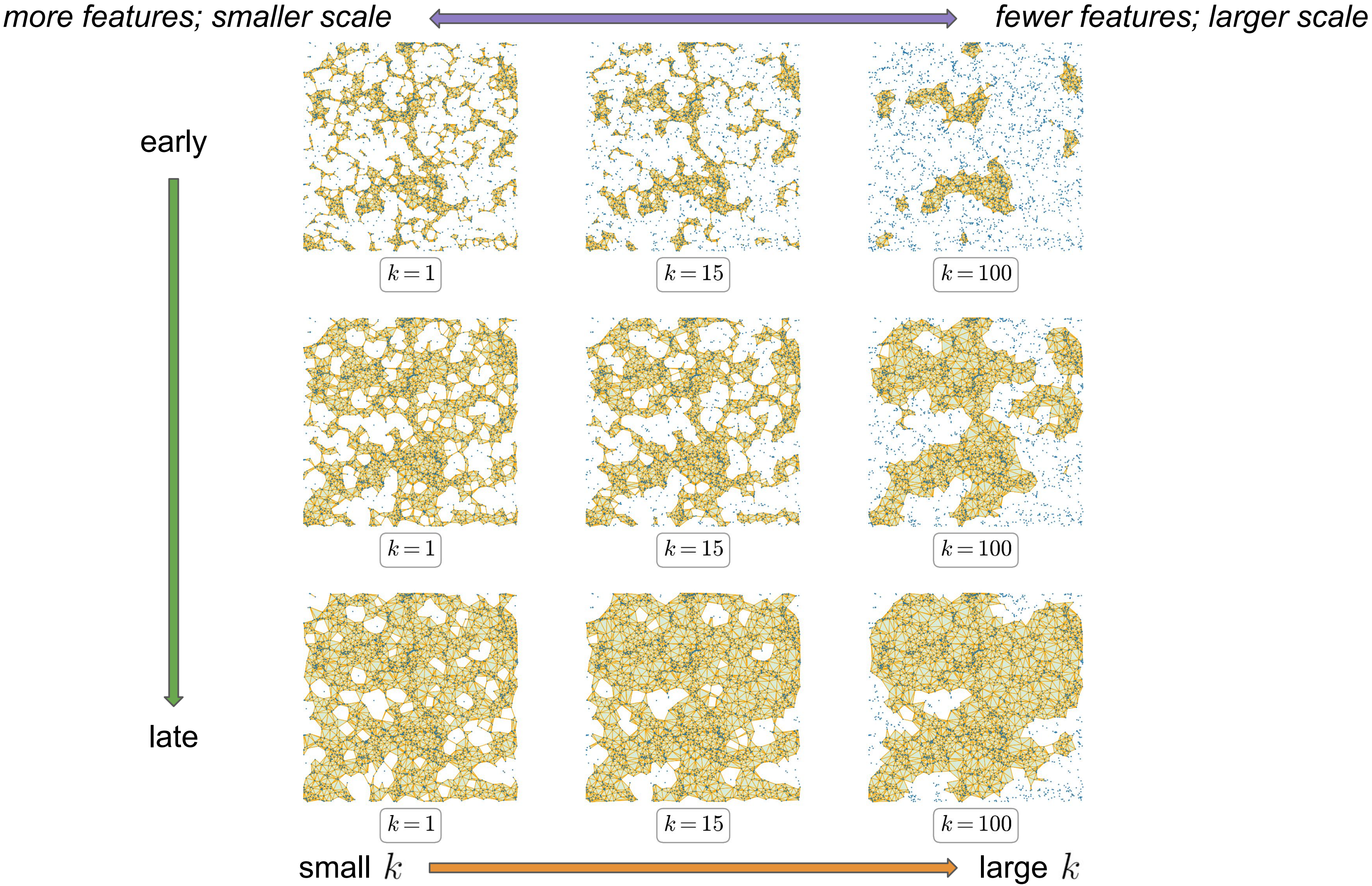
Hence, x in a sparsely populated region ⇒ x is an outlier ⇒ $f(x)$ is substantial.

The Nearest-Neighbour Parameter k

Note that we didn't fix k.

It turns out that k has an interesting effect on the filtration


$k = 1$
$k = 15$
$k = 100$



more features; smaller scale
fewer features; larger scale
early
late
$k=1$
$k=15$
$k=100$
$k=1$
$k=15$
$k=100$
$k=1$
$k=15$
$k=100$
small k
large k

The Nearest-Neighbour Parameter k

Note that we didn’t fix k.

It turns out that k has an interesting effect on the filtration

The Nearest-Neighbour Parameter k

Note that we didn't fix k.

It turns out that k has an interesting effect on the filtration $\Rightarrow$ k tracks features of the corresponding size

The Nearest-Neighbour Parameter k

Note that we didn't fix k.

It turns out that k has an interesting effect on the filtration ⇒ k tracks features of the corresponding size

⇓

k is a hyperparameter of the filtration

The Nearest-Neighbour Parameter k

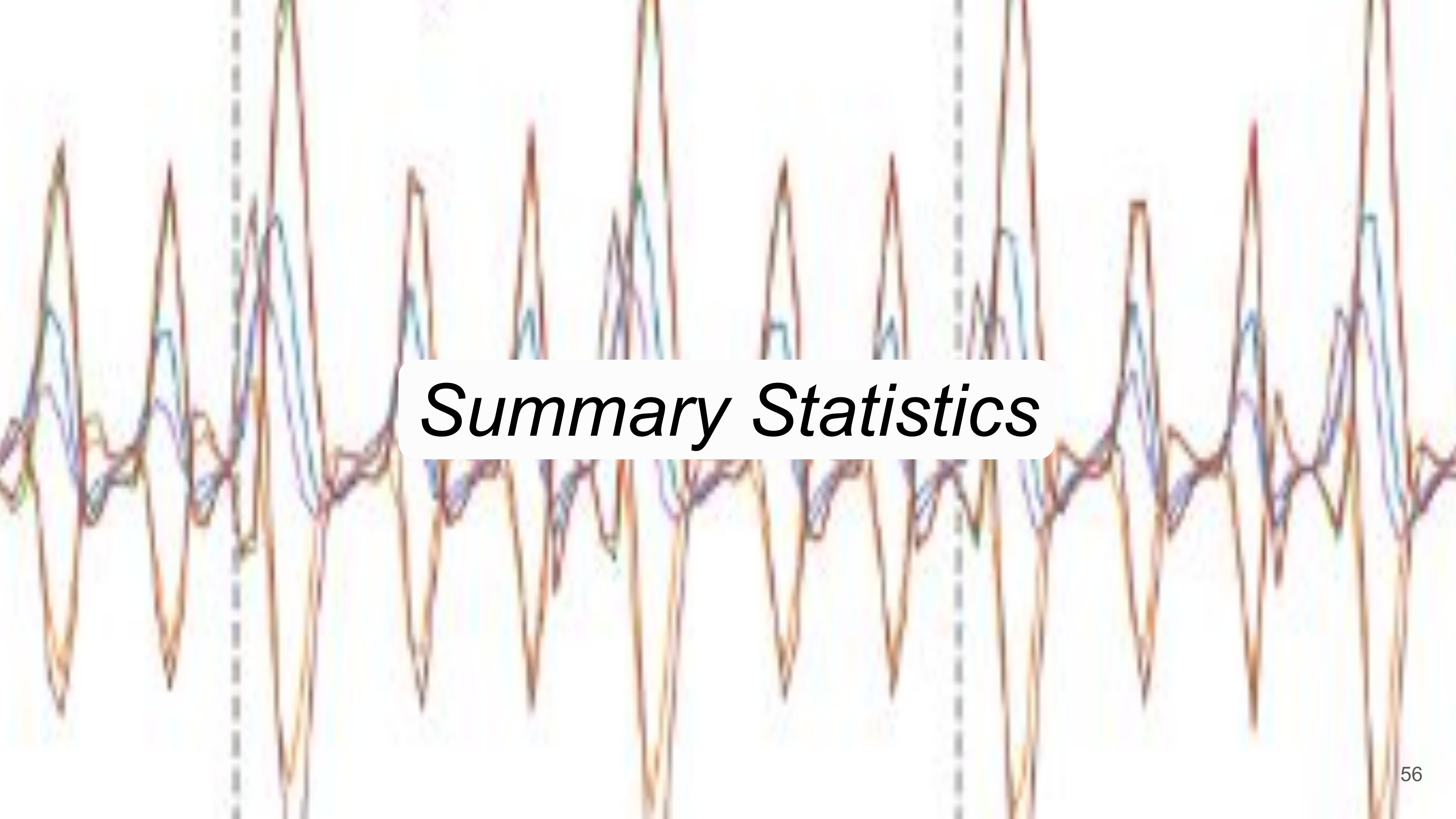
Note that we didn't fix k.

It turns out that k has an interesting effect on the filtration

k tracks features of the corresponding size

k is a hyperparameter of the filtration

Extract multi-scale information from all the filtrations from varying k



Summary Statistics

Dataset

Molino Suite
galaxy catalogs

(Hahn & Villaescusa-Navarro, 2020)

Dataset

- N-body simulations
- 5 cosmological parameters & neutrino mass
- 512^3 particles in 1 $(\mathrm{Gpc}/h)^3$ box

Quijote Simulations
halo catalogs
(Villaescusa-Navarro et al., 2020)

Molino Suite
galaxy catalogs
(Hahn & Villaescusa-Navarro, 2020)

Halo Occupation Distribution (HOD) model

- Places central & satellite galaxies based on mass of the host halo
- 5 HOD parameters

Dataset

- N-body simulations
- 5 cosmological parameters & neutrino mass
- 512^3 particles in 1 $(\mathrm{Gpc}/h)^3$ box

Quijote Simulations
$\boxed{\text{halo}}$ catalogs
(Villaescusa-Navarro et al., 2020)

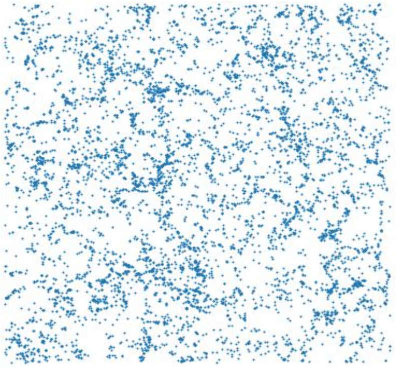
Halo Occupation Distribution (HOD) model
- Places central & satellite galaxies based on mass of the host halo
- 5 HOD parameters

molino

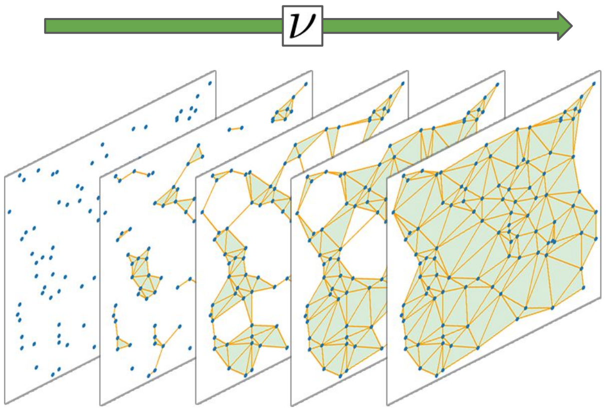
Molino Suite
$\boxed{\text{galaxy}}$ catalogs
(Hahn & Villaescusa-Navarro, 2020)

- 75000 catalogs in total
  - Covariance & derivative estimates
  - Convergence checks
- ~150000 galaxies at z = 0 in each catalog
  - ~60000 0-cycles *(clusters of galaxies)*
  - ~30000 1-cycles *(loops)*
  - ~20000 2-cycles *(voids)*
- Total 11 parameters
  - *Cosmological:* $\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8, M_\nu$
  - *HOD:* $\log M_{\rm min}, \sigma_{\log M}, \log M_0, \alpha, \log M_1$
  - Catalogs for each varied one step above and below
- Redshift space displacement applied along each axis

Data Vector for a single k



Point cloud of a galaxy mock



Given k, compute the filtration

ID	Dimension	ν_{birth}	ν_{death}
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

A list of birth and death times

Data Vector for a single k

ID	Dimension	ν_{birth}	ν_{death}
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

A list of birth and death times

Data Vector for a single k

ID	Dimension	ν_{birth}	ν_{death}
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

A list of birth and death times

ν_{birth}

a distribution for 0-cycles, B_0

a distribution for 1-cycles, B_1

a distribution for 2-cycles, B_2

ν_{death}

a distribution for 0-cycles, D_0

a distribution for 1-cycles, D_1

a distribution for 2-cycles, D_2

Data Vector for a single k

ID	Dimension	ν_{birth}	ν_{death}
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

A list of birth and death times

ν_{birth}

ν_{death}

a distribution for 0-cycles, B_0

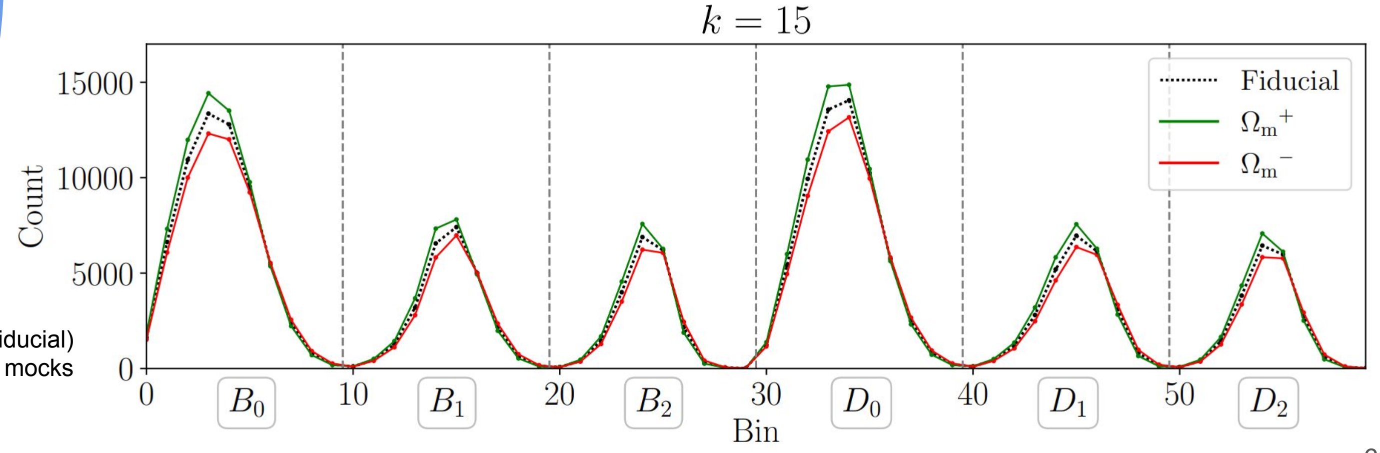
a distribution for 1-cycles, B_1

a distribution for 2-cycles, B_2

a distribution for 0-cycles, D_0

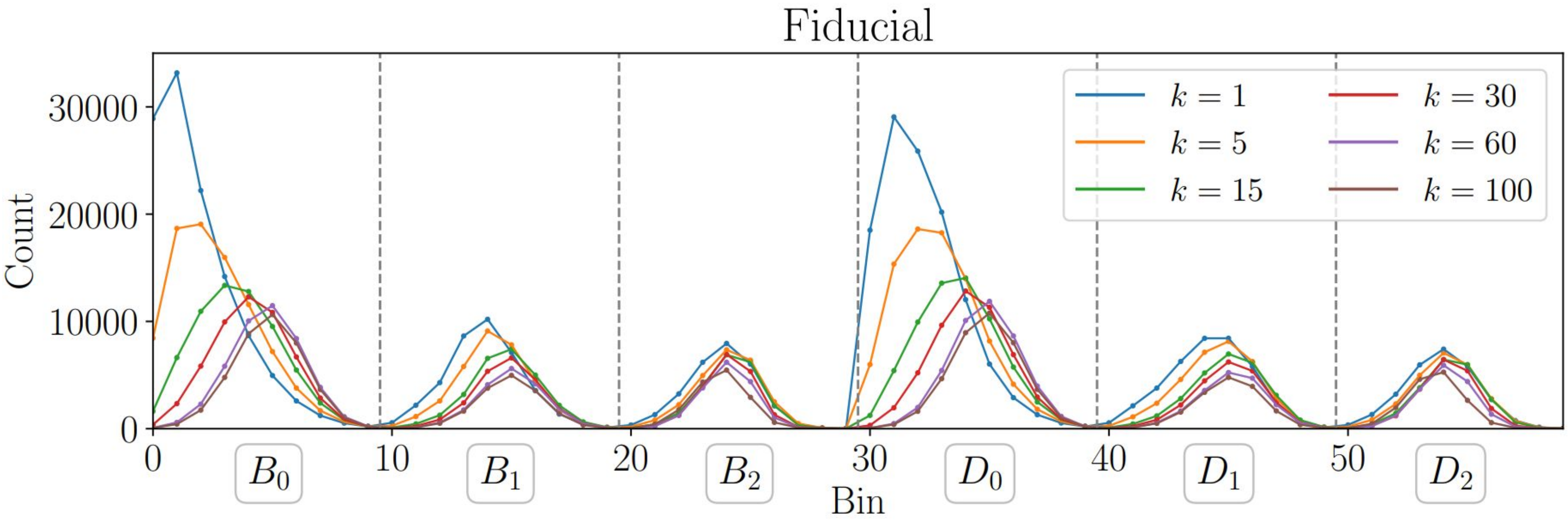
a distribution for 1-cycles, D_1

a distribution for 2-cycles, D_2



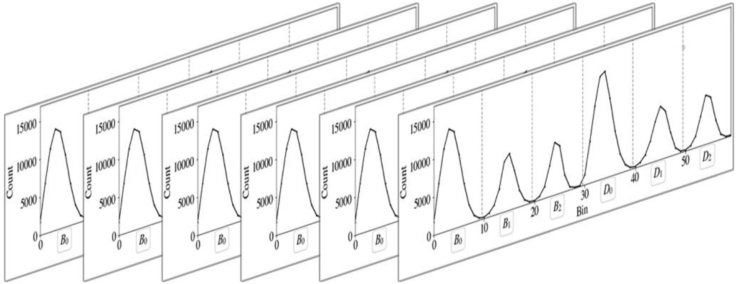
averaged over 9000 (fiducial) and 1500 ($\Omega_m^\pm$) mocks

Varying k



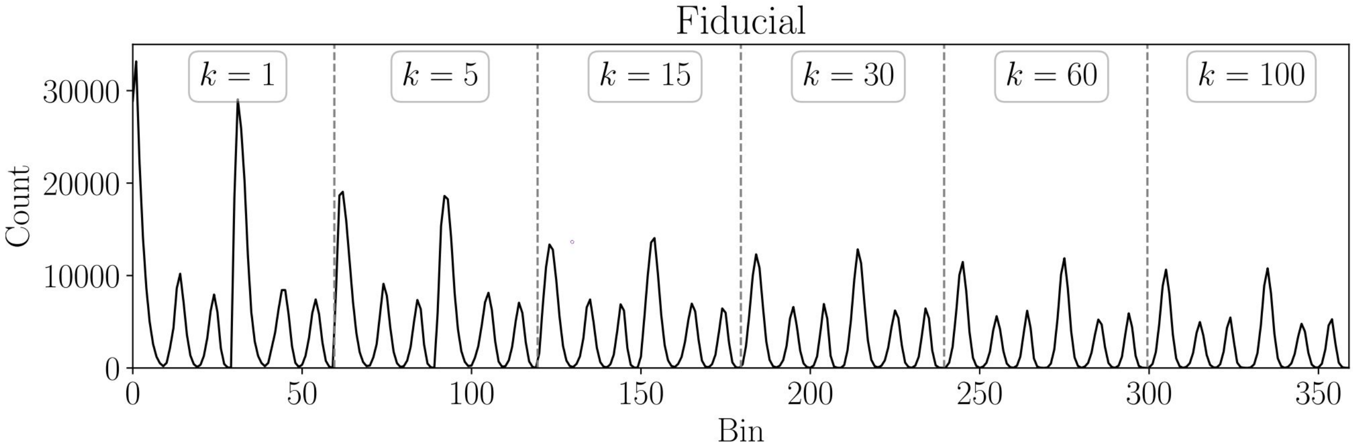
Fiducial
Count
30000
20000
10000
0
0
10
20
30
40
50
Bin
B_0
B_1
B_2
D_0
D_1
D_2
$k = 1$
$k = 5$
$k = 15$
$k = 30$
$k = 60$
$k = 100$

Full Data Vector

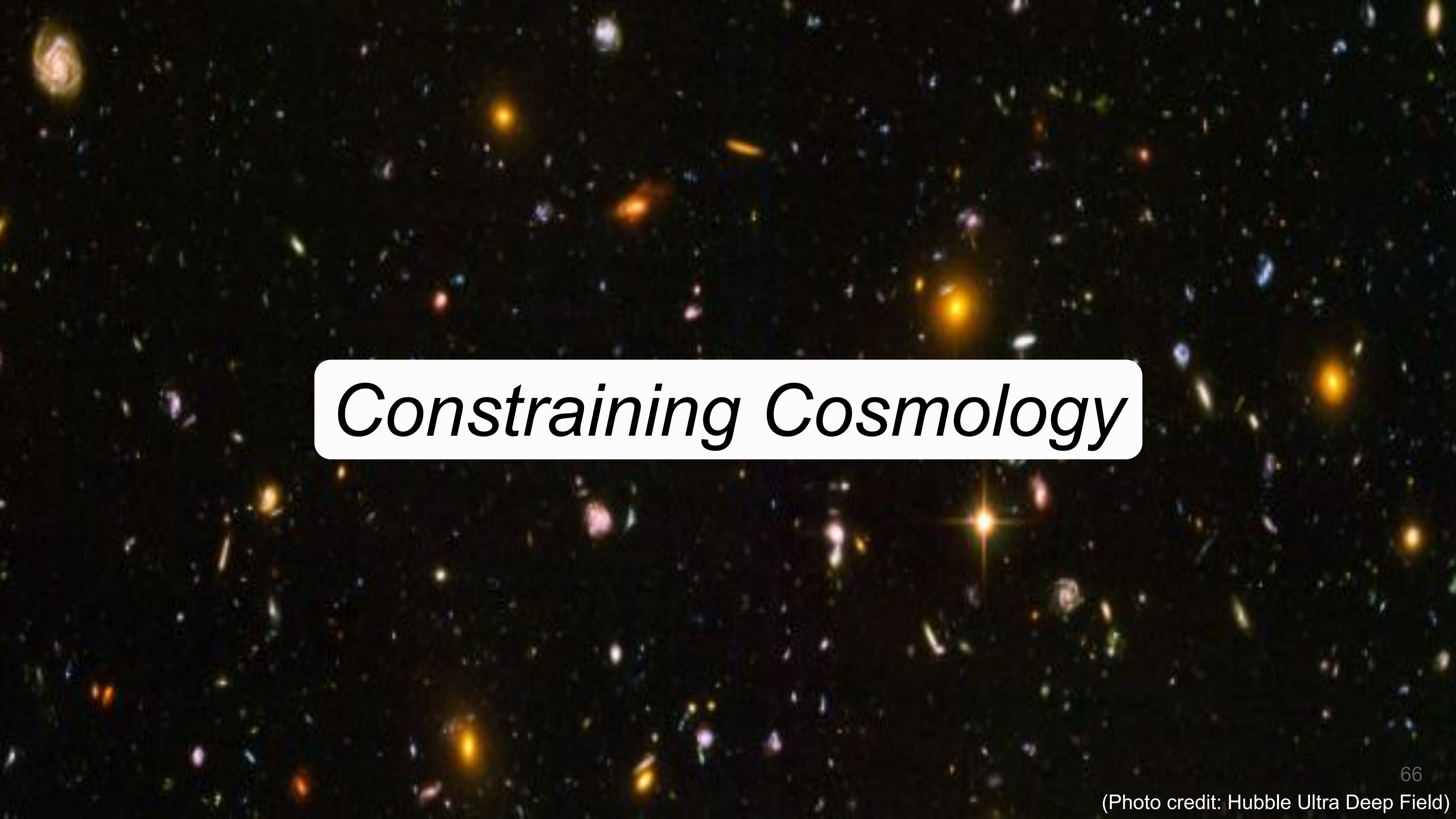


Data vectors from $k = 1, 5, 15, 30, 60, 100$

concatenate



Constraining Cosmology



(Photo credit: Hubble Ultra Deep Field)

Fisher Information Matrix

$$F_{ij} = D_{,i}^T C^{-1} D_{,j} + \frac{1}{2} \mathrm{Tr} \left[\left(C^{-1} C_{,i} \right) \left(C^{-1} C_{,j} \right) \right]$$

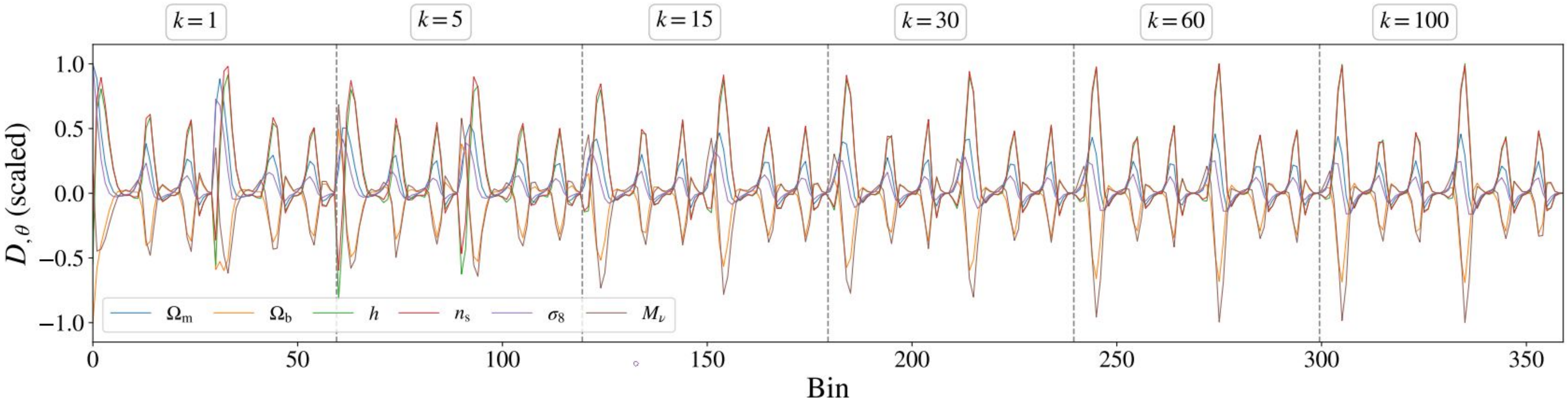
Covariance matrix: F^{-1}

1-σ constraint for parameter θ_i*:* $\sigma_i = \sqrt{(F^{-1})_{ii}}$

Fisher Information Matrix

$$F_{ij} = \underline{D^T_{,i}} C^{-1} \underline{D_{,j}} + \frac{1}{2}\mathrm{Tr}\left[\left(C^{-1}C_{,i}\right)\left(C^{-1}C_{,j}\right)\right]$$

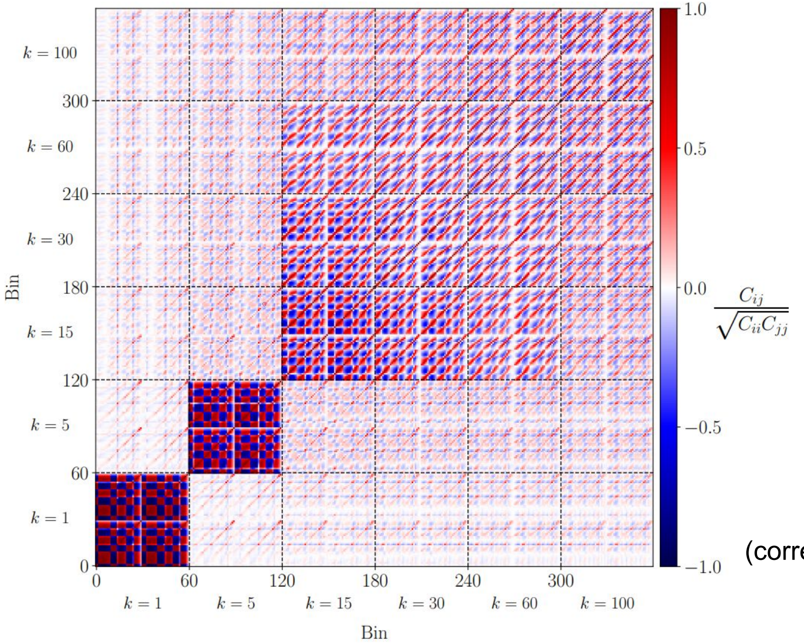
Data vector derivative for θ_i: $D_{,i} = (D(\theta_i^+) - D(\theta_i^-))/2\theta_i$



Fisher Information Matrix

$$F_{ij} = D_{,i}^T \underline{C^{-1}} D_{,j} + \frac{1}{2} \mathrm{Tr}\left[\left(C^{-1} C_{,i}\right) \left(C^{-1} C_{,j}\right) \right]$$

Covariance of the data vector at fiducial cosmology: $C_{ij} = \langle (D_i - D_{\mathrm{mean}})(D_j - D_{\mathrm{mean}}) \rangle$



(correlation matrix here)

Fisher Information Matrix

$$F_{ij} = D_{,i}^{T} C^{-1} D_{,j} \underline{+ \frac{1}{2} \mathrm{Tr} \left[\left(C^{-1} C_{,i} \right) \left(C^{-1} C_{,j} \right) \right]}$$

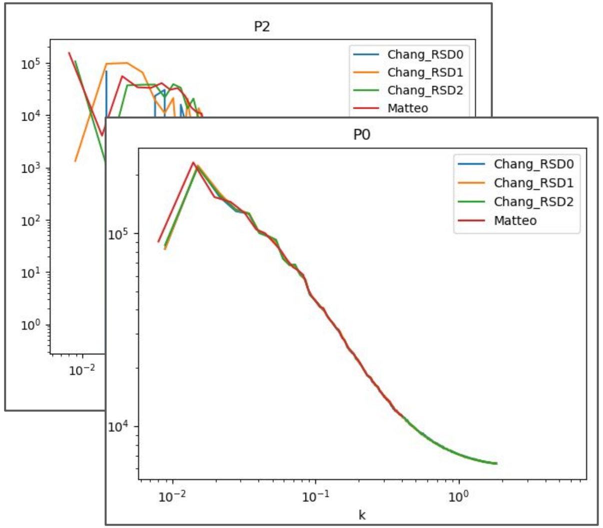
Covariance derivative term: conservatively omitted; information overlap

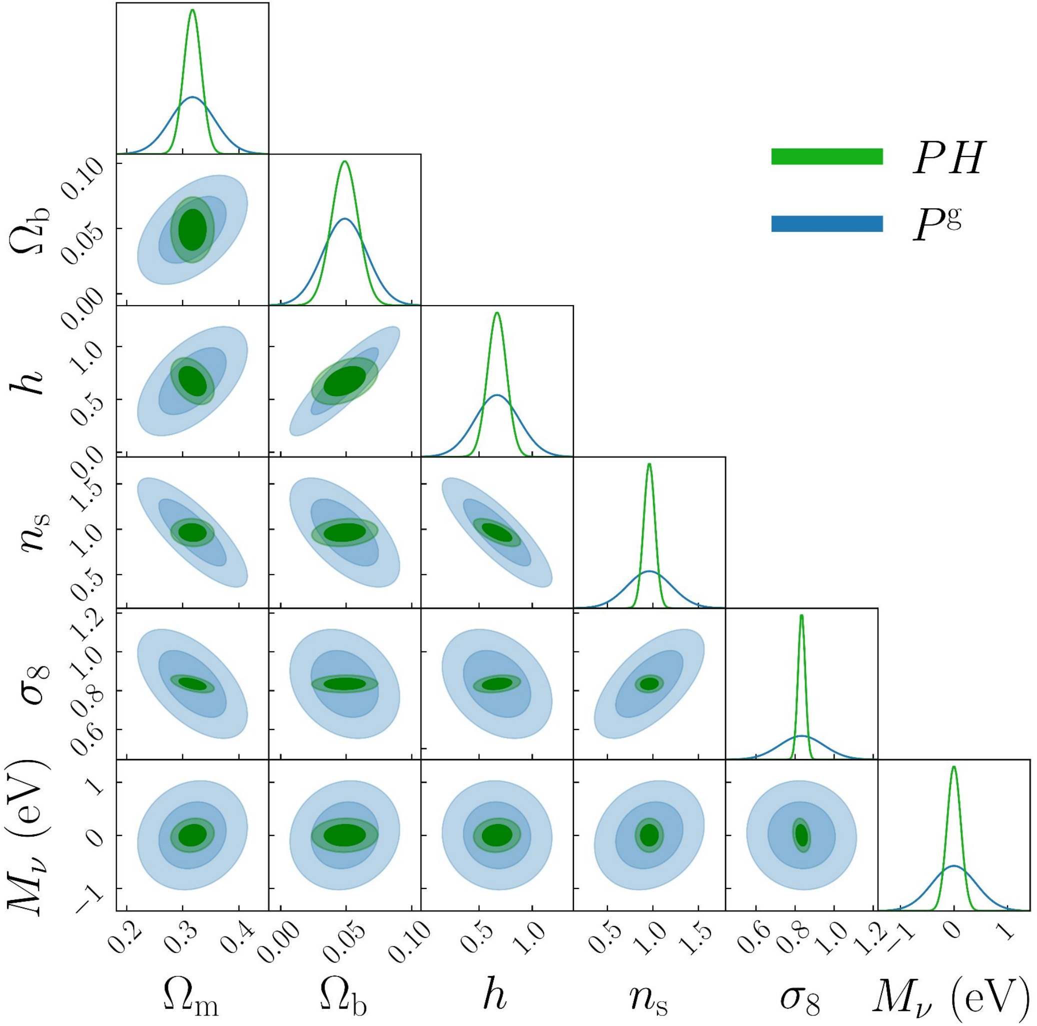
Other checks:

- Invertible covariance matrix
- Gaussianity of summary statistics
- Convergence of components and constraints

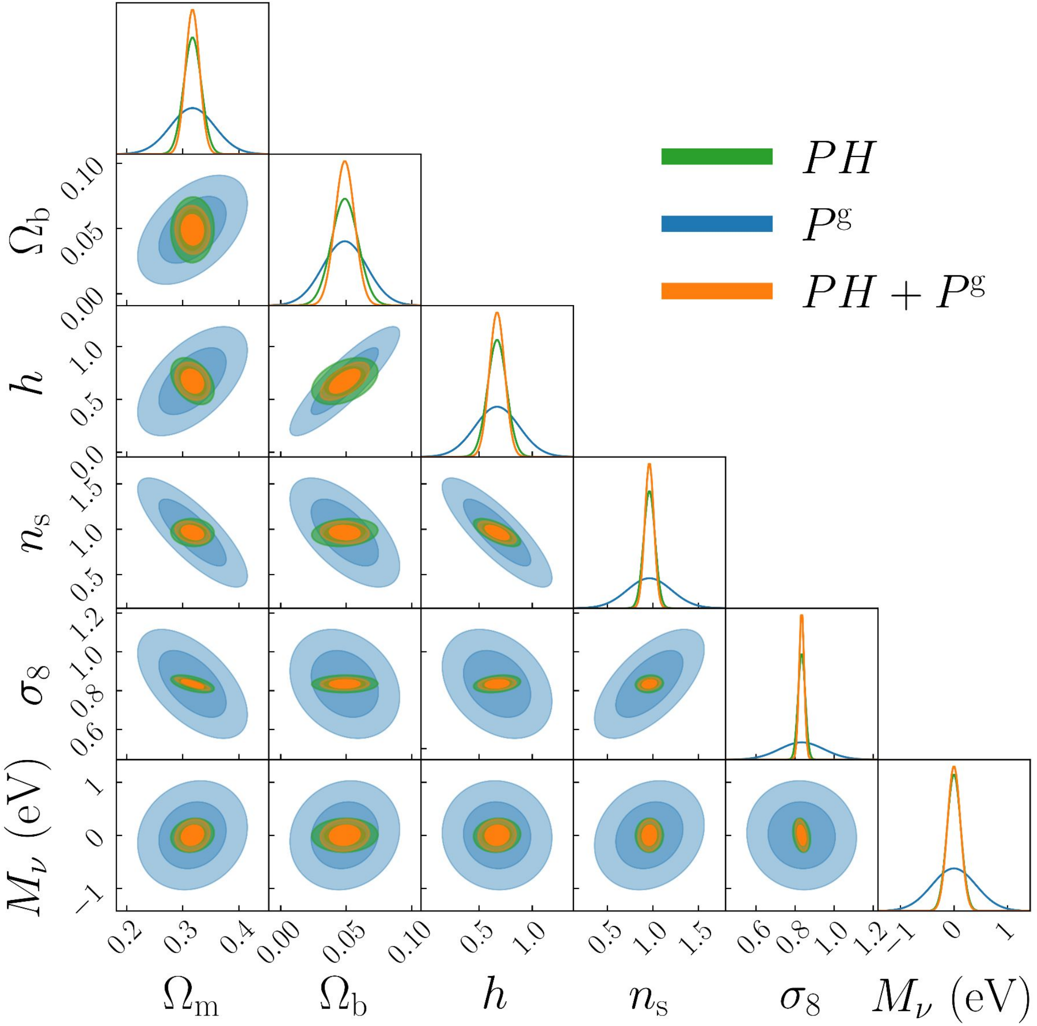
Power Spectrum Measurements

We also compute the
galaxy power spectra *monopole* and *quadrupole*
up to $k_{\max} = 0.4\ h/\mathrm{Mpc}$,
as the baseline for comparison.





(marginalized over
HOD parameters)



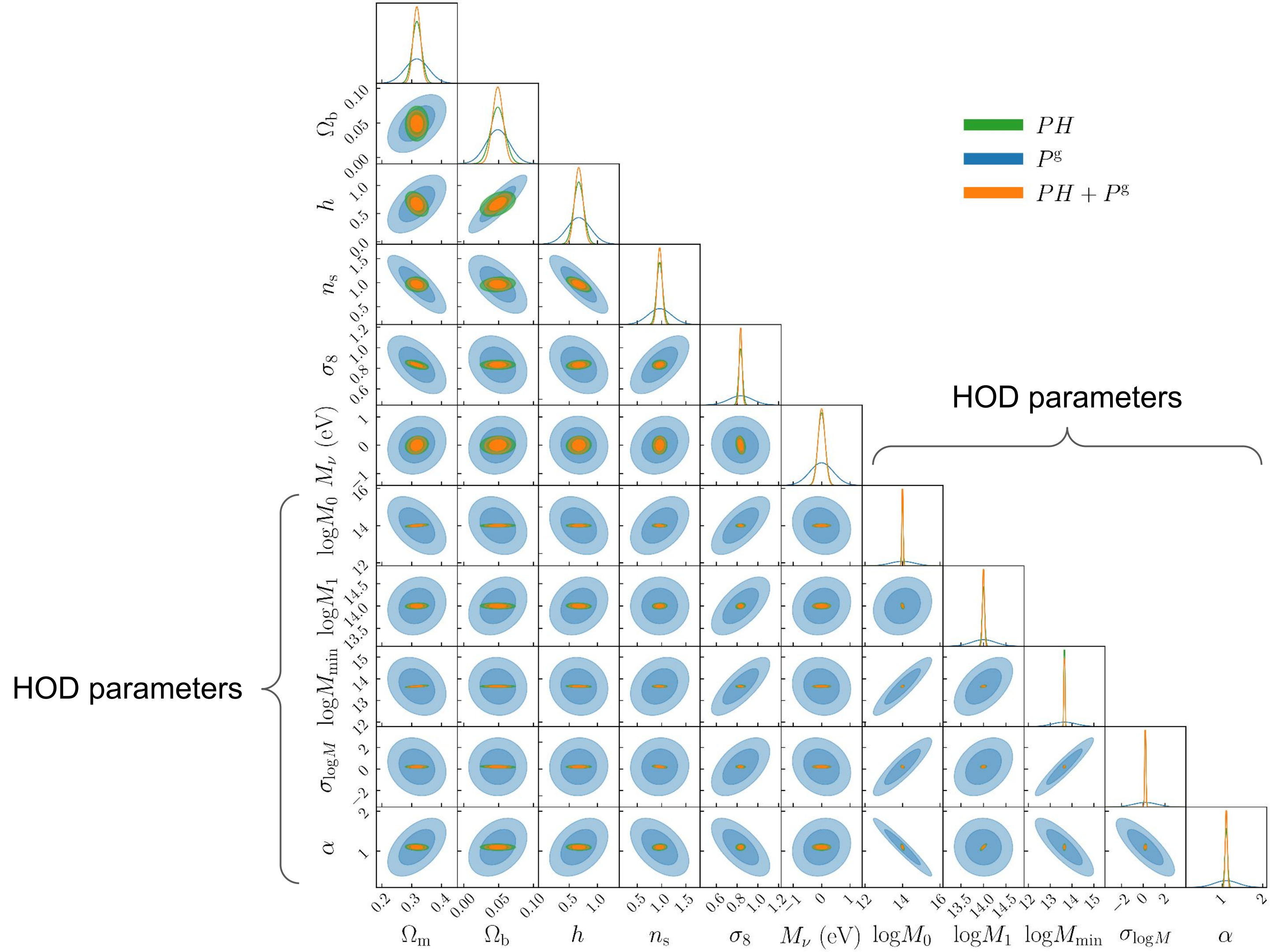
(marginalized over HOD parameters)

1-σ constraints

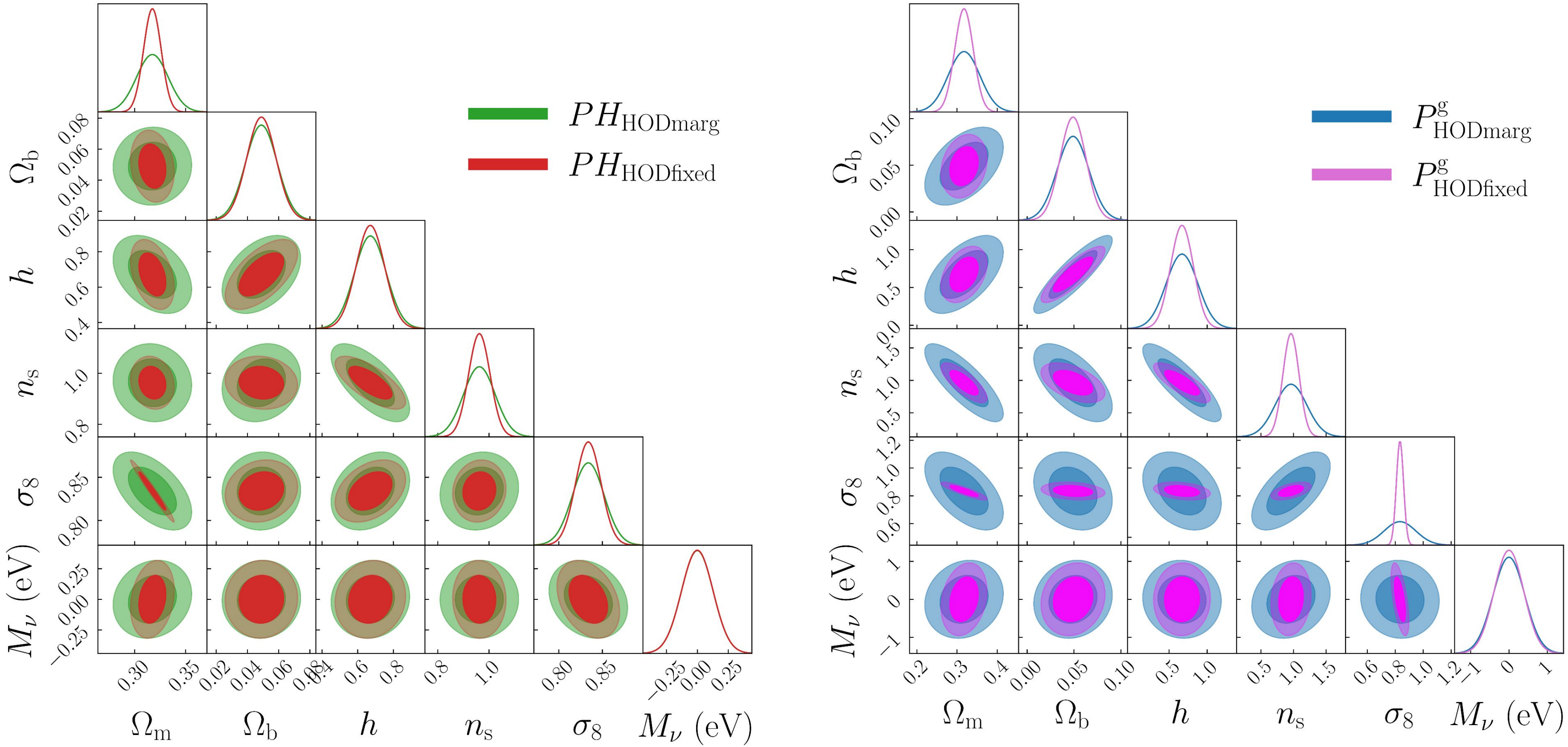
galaxy bispectrum monopole

(Hahn & Villaescusa-Navarro, 2020)

	PH	$P_0^{\rm g} + P_2^{\rm g}$ ($k_{\rm max} = 0.4\ h/{\rm Mpc}$)	$P_0^{\rm g} + P_2^{\rm g} + B_0^{\rm g}$ ($k_{\rm max} = 0.2\ h/{\rm Mpc}$)
$\Omega_{\rm m}$	0.016	0.039	0.030
$\Omega_{\rm b}$	0.010	0.017	0.013
h	0.089	0.208	0.157
$n_{\rm s}$	0.062	0.243	0.165
σ_8	0.018	0.113	0.053
M_ν (eV)	0.129	0.414	0.282


PH
P^{g}
$PH + P^{\mathrm{g}}$
HOD parameters
HOD parameters
Ω_{m}
Ω_{b}
h
n_{s}
σ_8
M_ν (eV)
$\log M_0$
$\log M_1$
$\log M_{\mathrm{min}}$
$\sigma_{\log M}$
α

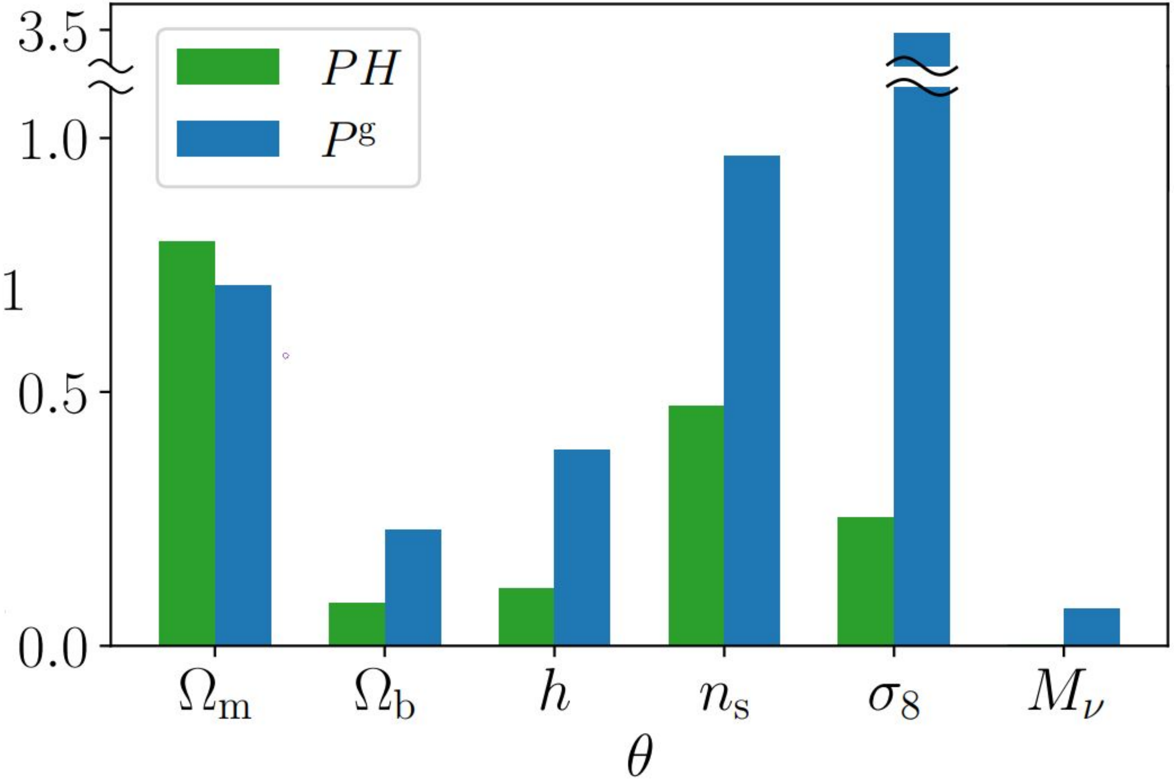
Robustness Against HOD Parameters



Robustness Against HOD Parameters

"Cost of marginalization":

$$(\sigma_\theta^{\mathrm{HODmarg}}/\sigma_\theta^{\mathrm{HODfixed}}) - 1$$



Concluding Remarks

- Promising results and robustness over momentum-space statistics
  - Against marginalization over nuisance parameters
  - Despite the use of a very basic data vector

Concluding Remarks

- Promising results and robustness over momentum-space statistics
  - Against marginalization over nuisance parameters
  - Despite the use of a very basic data vector

- Data vector excludes much information on the level of individual features
  - Limited by dataset, convergence requirements
  - Persistence diagrams, persistence images…

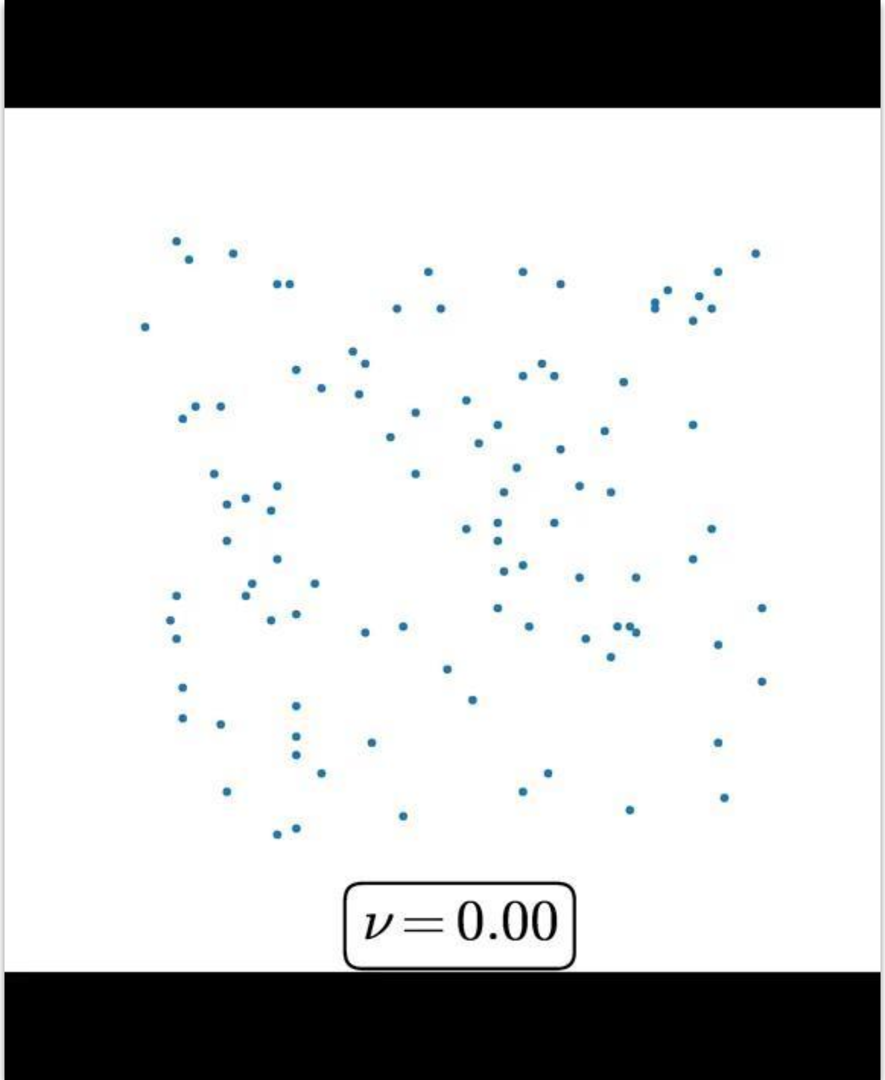
Concluding Remarks

- Promising results and robustness over momentum-space statistics
  - Against marginalization over nuisance parameters
  - Despite the use of a very basic data vector

- Data vector excludes much information on the level of individual features
  - Limited by dataset, convergence requirements
  - Persistence diagrams, persistence images…

- Output of persistence computation as input of machine learning strategies

Grazie mille !


$\nu = 0.00$

Persistence of the Large Scale Structures