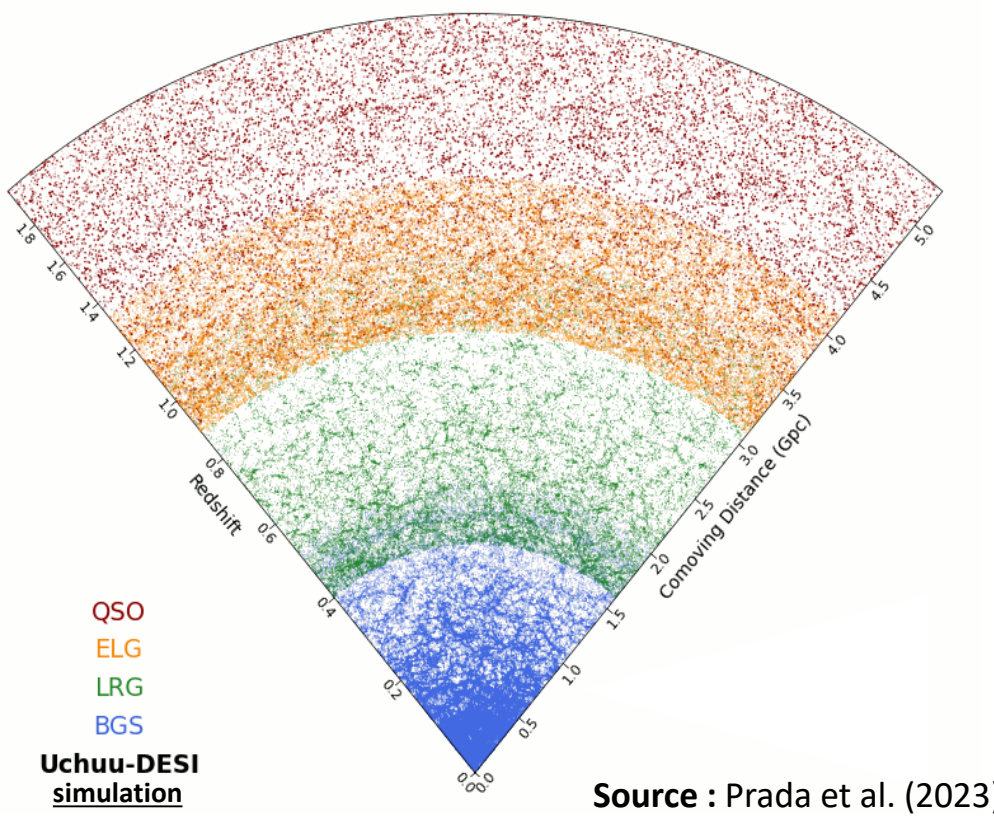
Beyond standard galaxy clustering for DESI BGS

Simon BOUCHARD,
under Pauline ZARROUK supervision

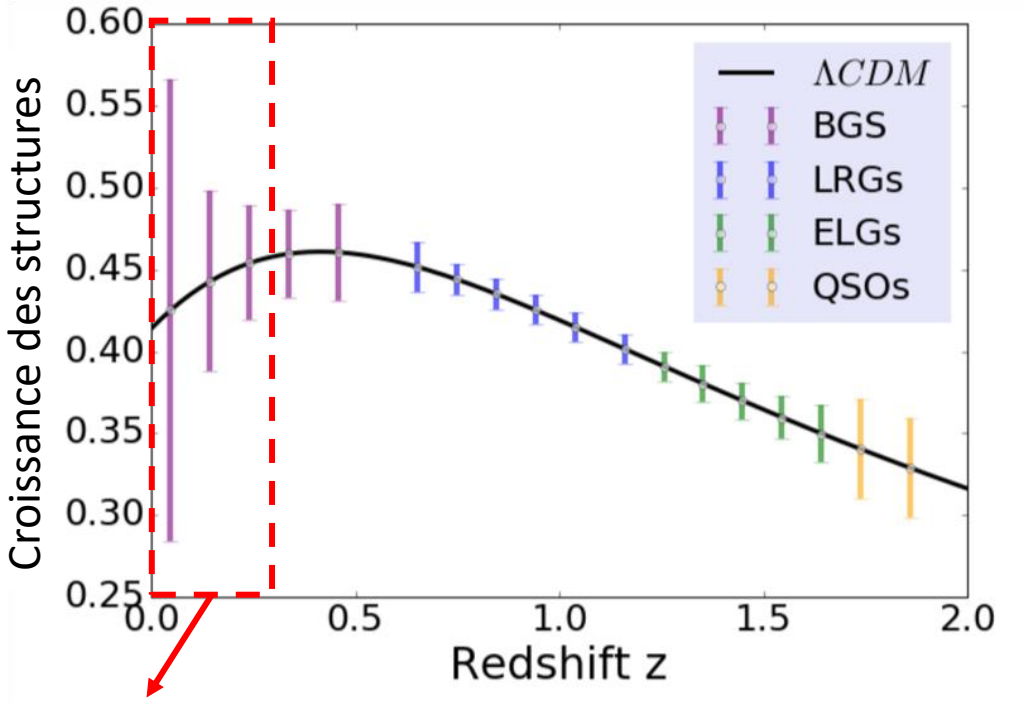
DESI : BGS

Bright Galaxy Survey

- Densest tracer ($n \approx 2.10^{-2}$)
- Closest tracer ($z < 0,5$)



Source : Prada et al. (2023)

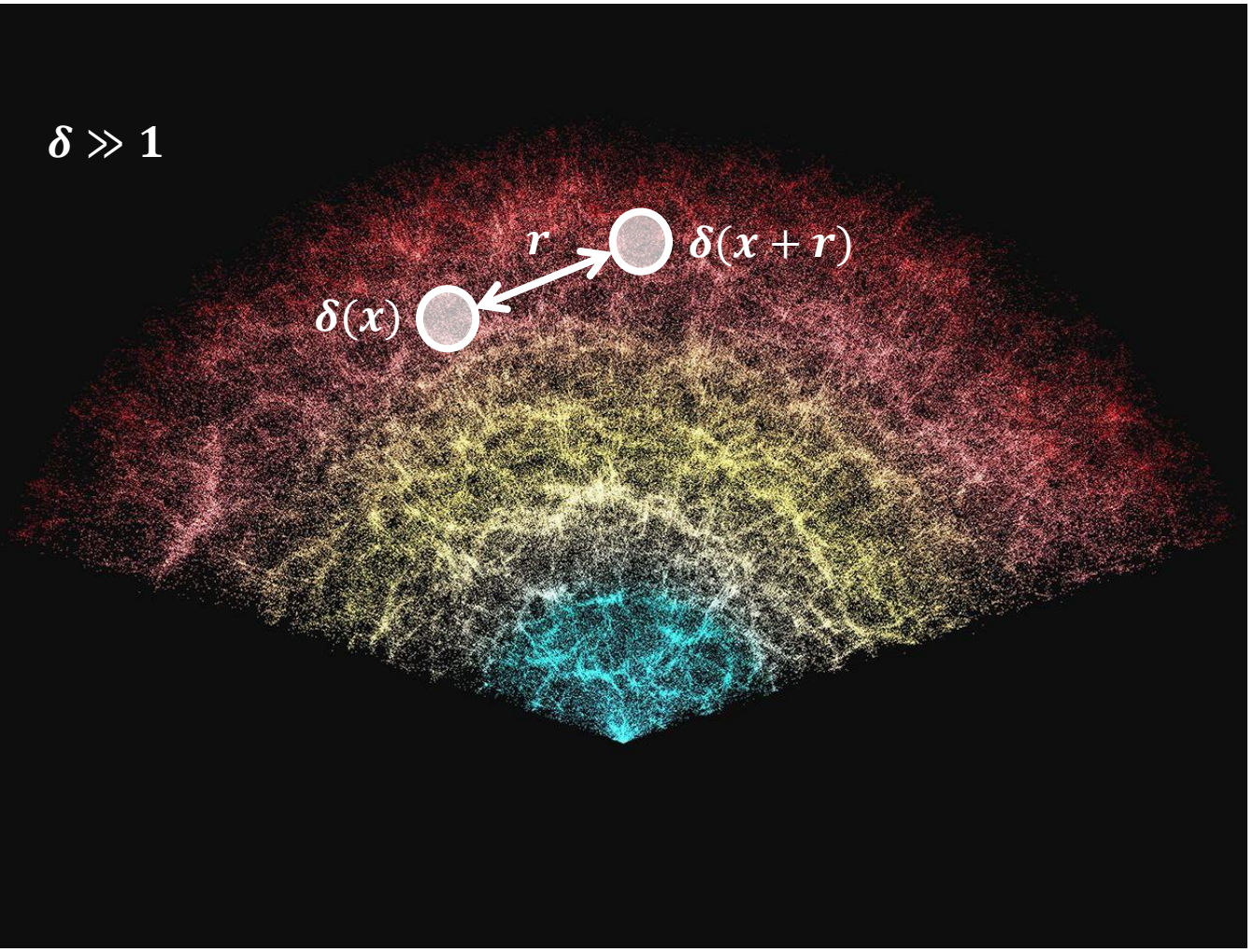


Dominated by
« Cosmic Variance »
(volume sample variance)

Source : DESI Collaboration (2016a)

Cosmological probe

Galaxy clustering



Source : NOIRLab | DESI Collaboration

Density contrast

$$\delta(x) = \frac{\rho(x) - <\rho>}{<\rho>}$$

« Standard » analysis

Two-point correlation function (« 2PCF »)

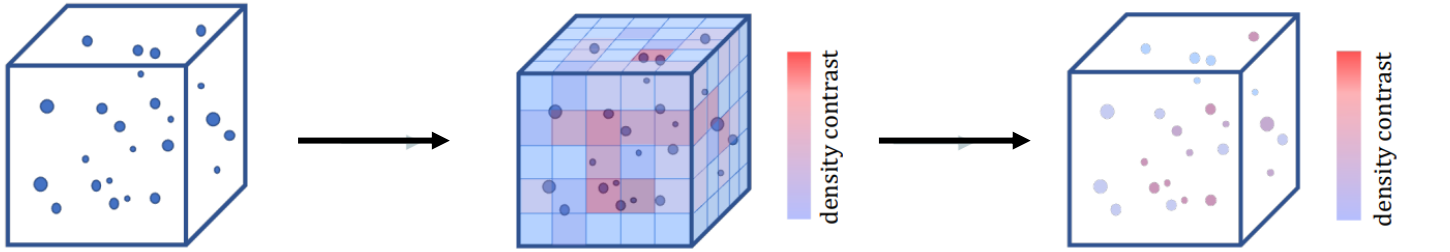
$$\xi(r) = <\delta(x)\delta(x+r)>_x$$

- Statistical tool
- Describes matter structuration
- Only captures gaussian information

Densitysplit

Paillas et al. (2020)

Source : Mathilde Pinon →



1. Catalog of tracers

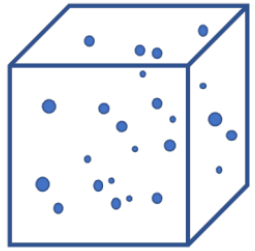
2. Density mesh with cell size R

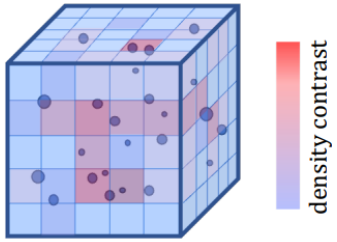
density at each particle location

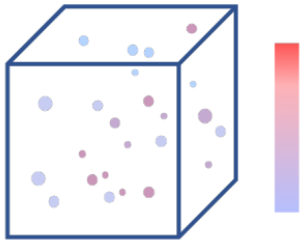
Densitysplit

Paillas et al. (2020)

Source : Mathilde Pinon →







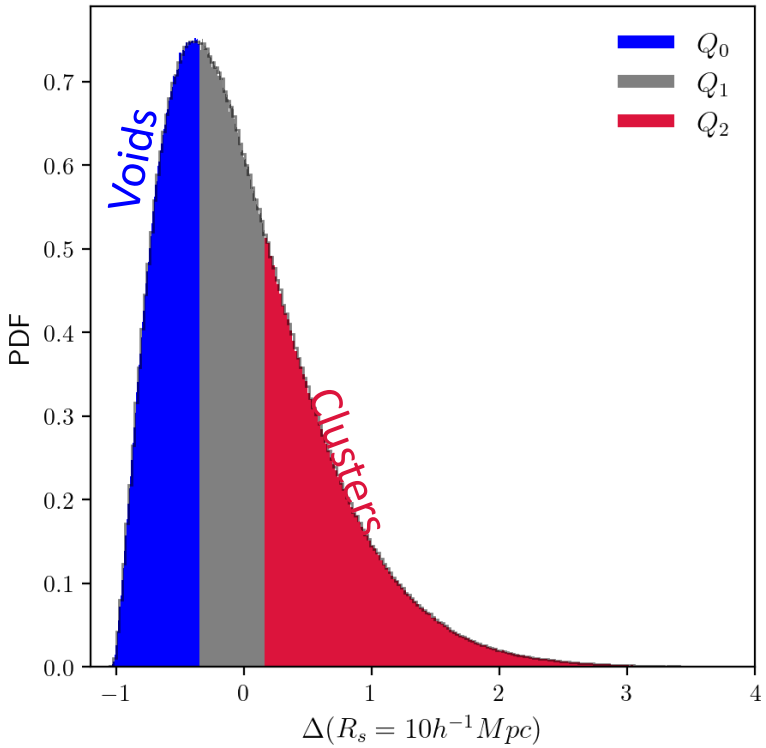
density at each particle location

→ Non-standard analysis !

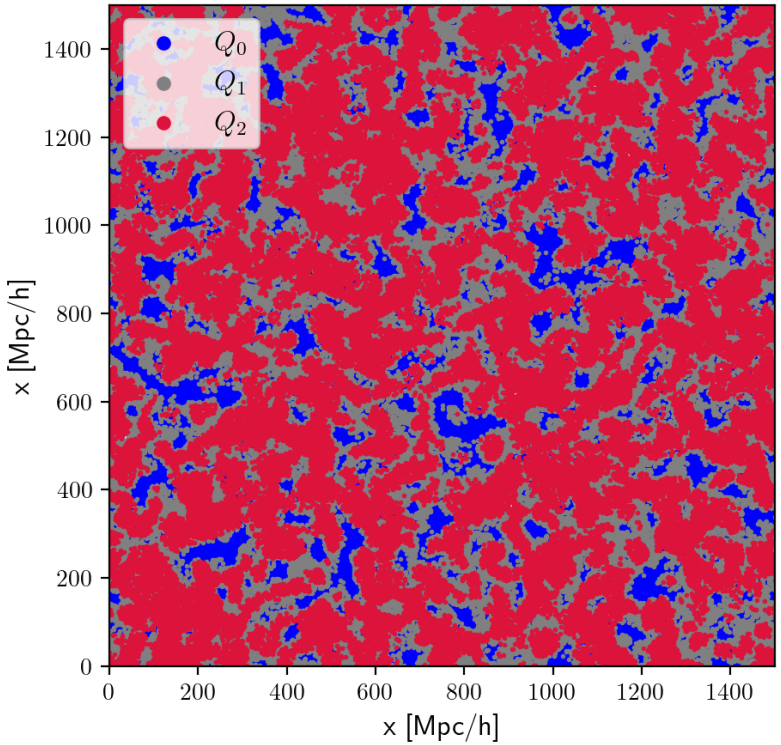
1. Catalog of tracers

2. Density mesh with cell size R

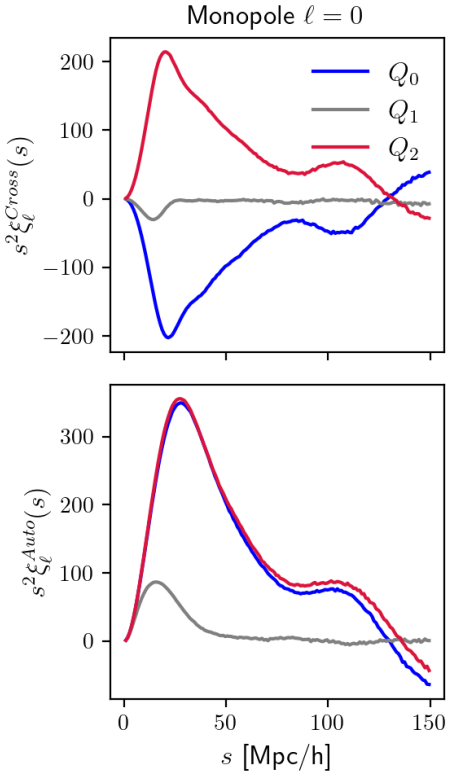
4. Density histogram



5. Split in N quantiles

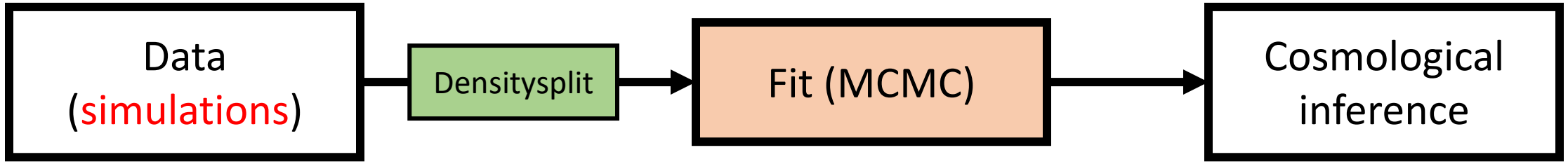


6. Correlations



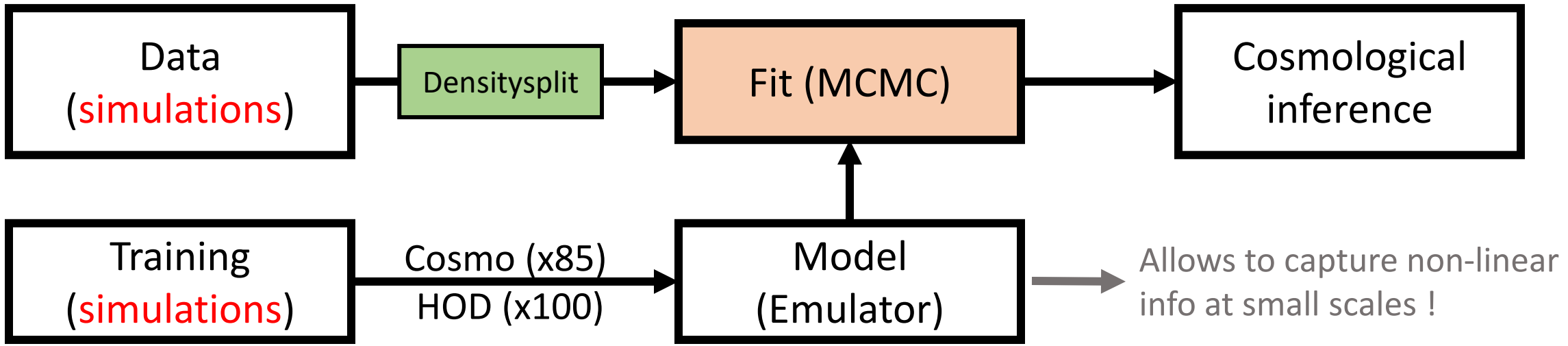
Inference pipeline

→ Goal : Constrain the cosmological parameters

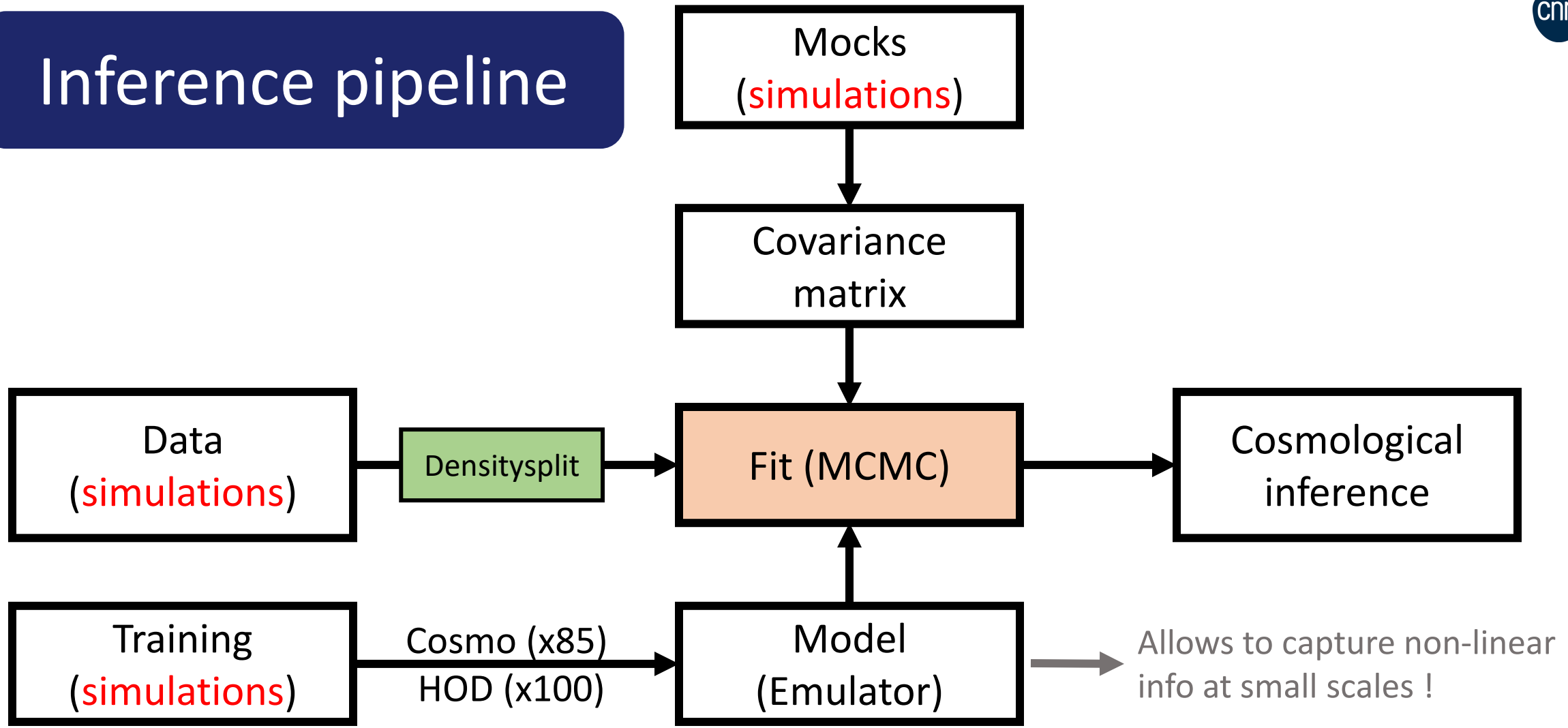


Inference pipeline

→ Goal : Constrain the cosmological parameters



Inference pipeline

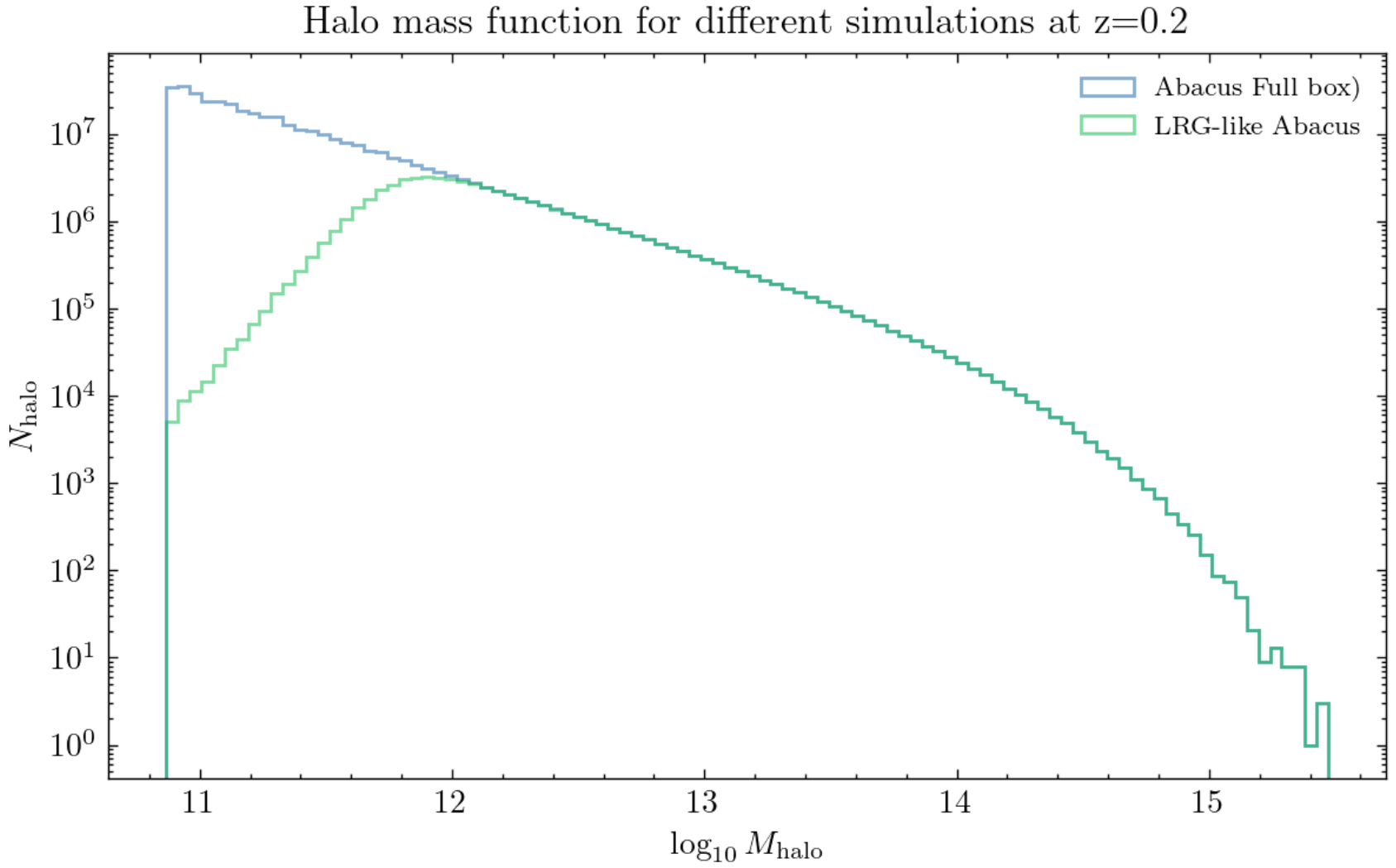


→ Developing the interface between the codes (acm)

→ Generation of BGS simulations

N-body simulations

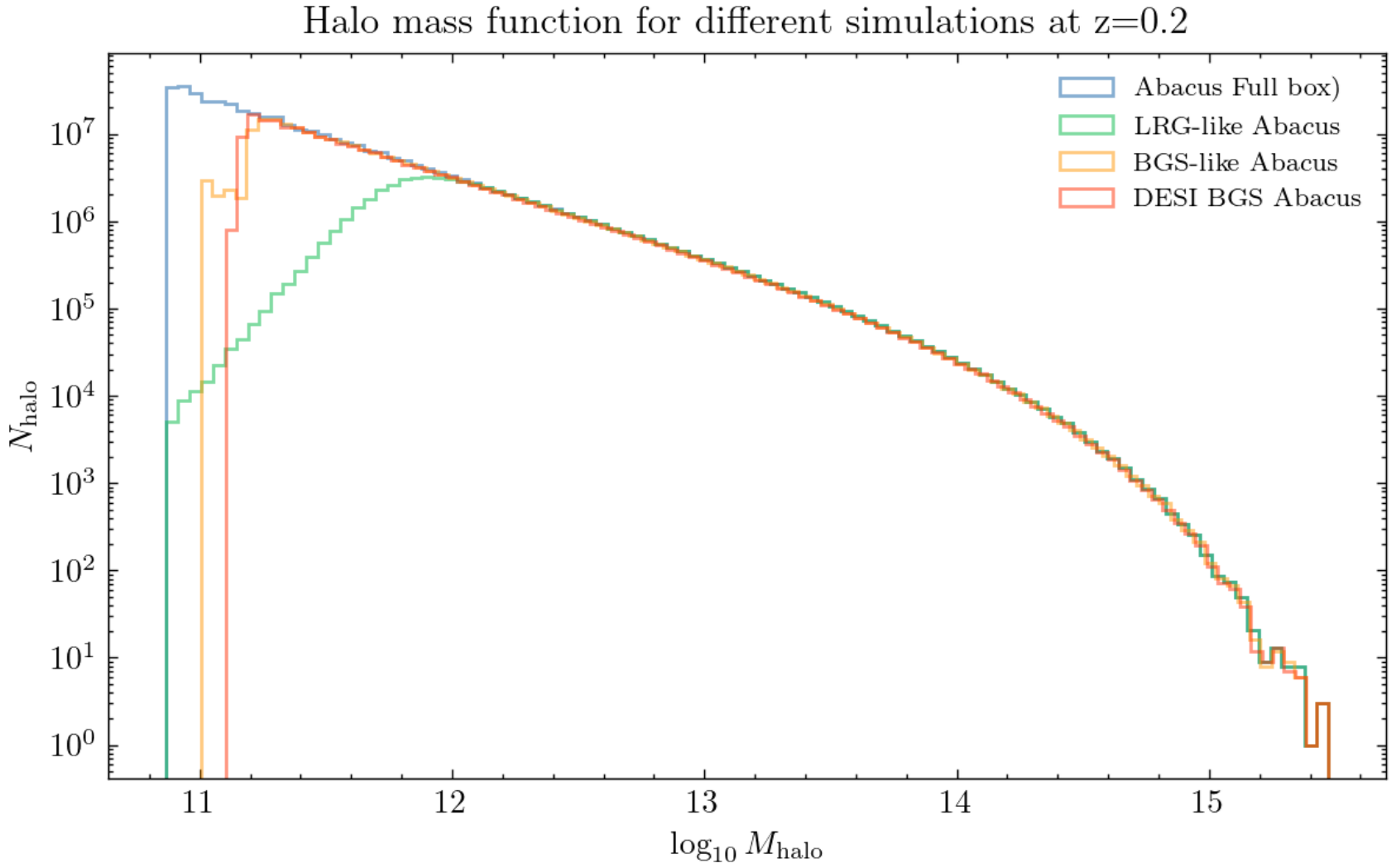
Halo mass function



Simulations : AbacusSummit (Maksimova et al. 2021)

N-body simulations

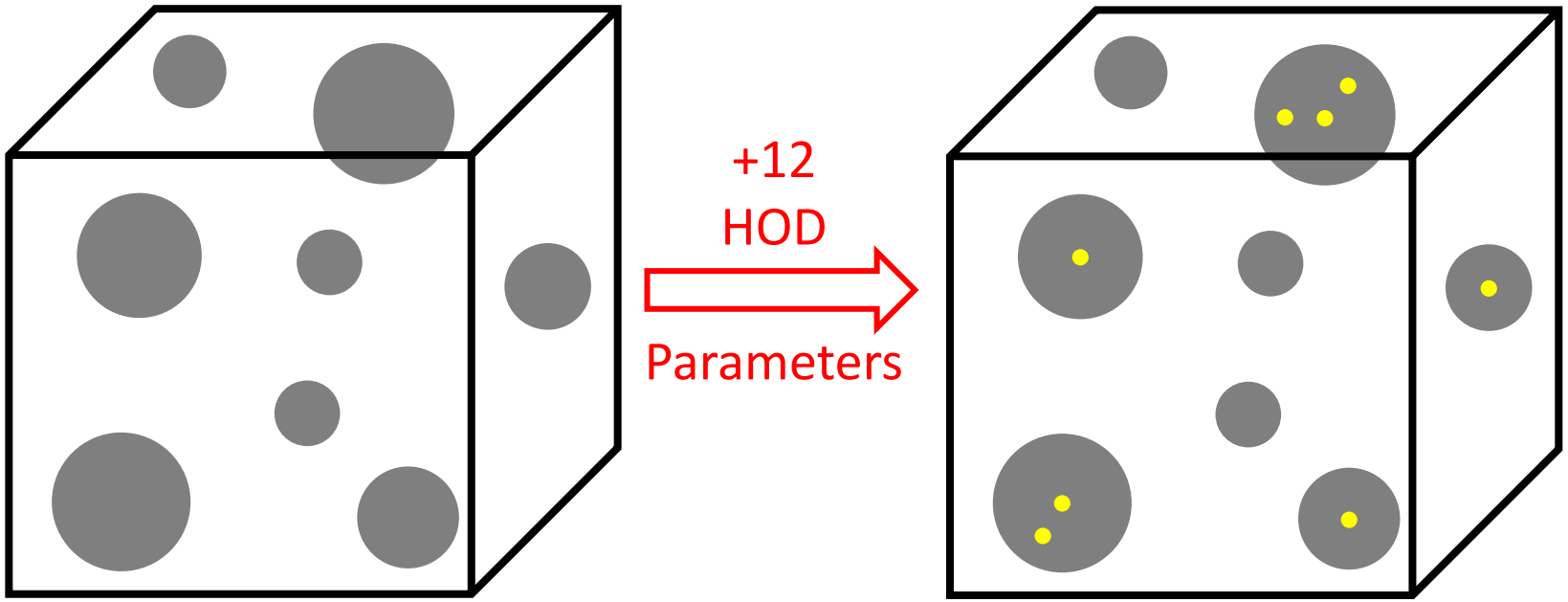
Halo mass function



Simulations : AbacusSummit (Maksimova et al. 2021)

Simulations

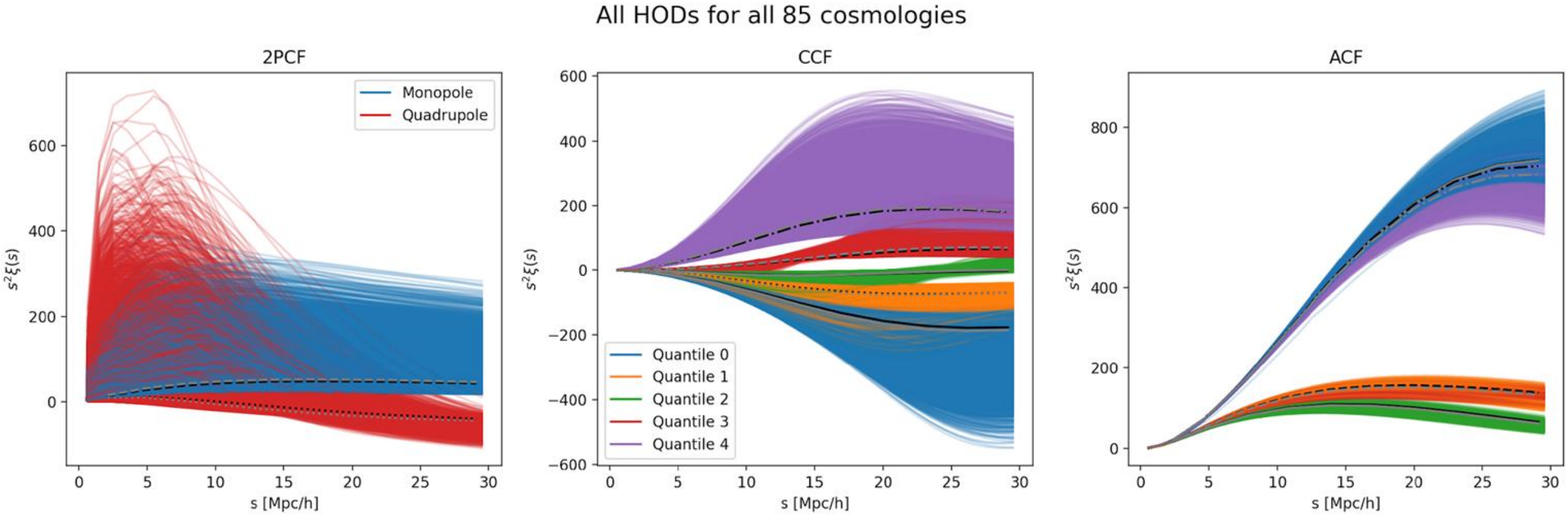
HOD Model



→ $\{M_{min}, \sigma_{\log M}, M_1, \alpha, \kappa, \alpha_c, \alpha_s, B_{cen}, B_{sat}, A_{cen}, A_{sat}, s\}$ to fit with cosmological parameters

Simulations

Statistics

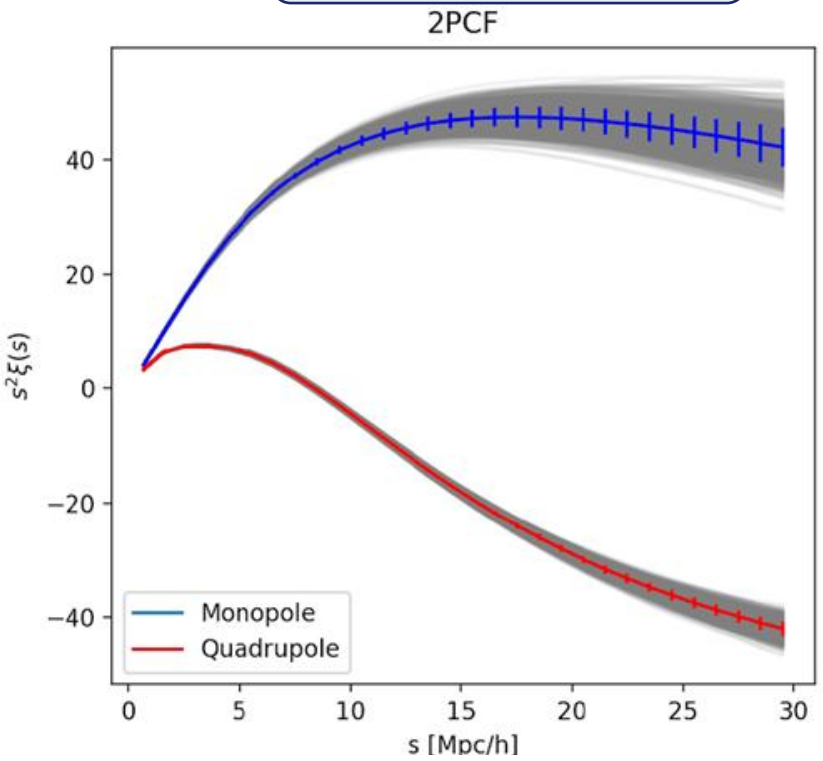


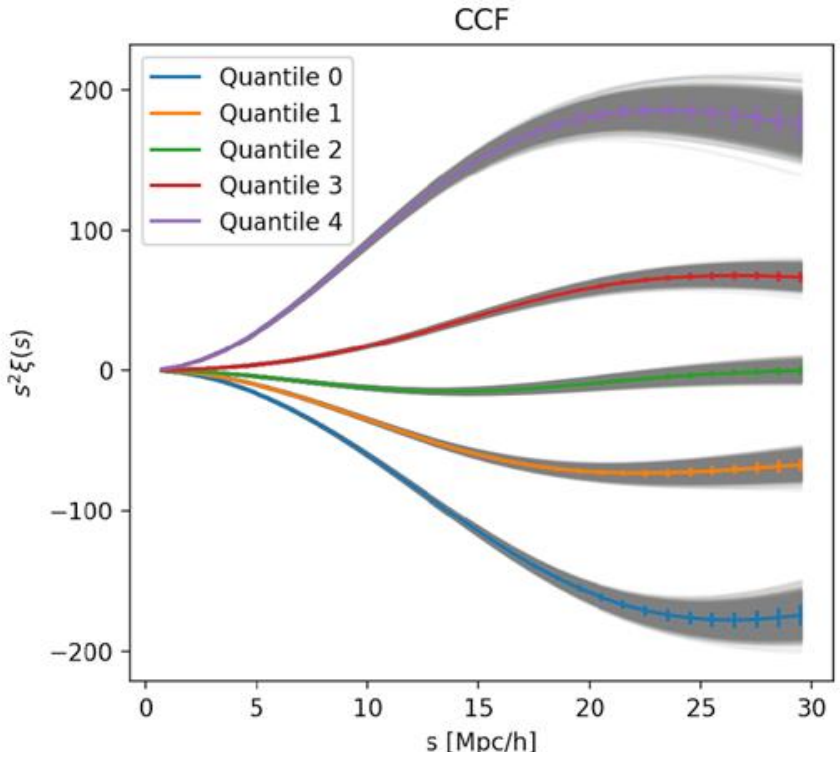
Simulations

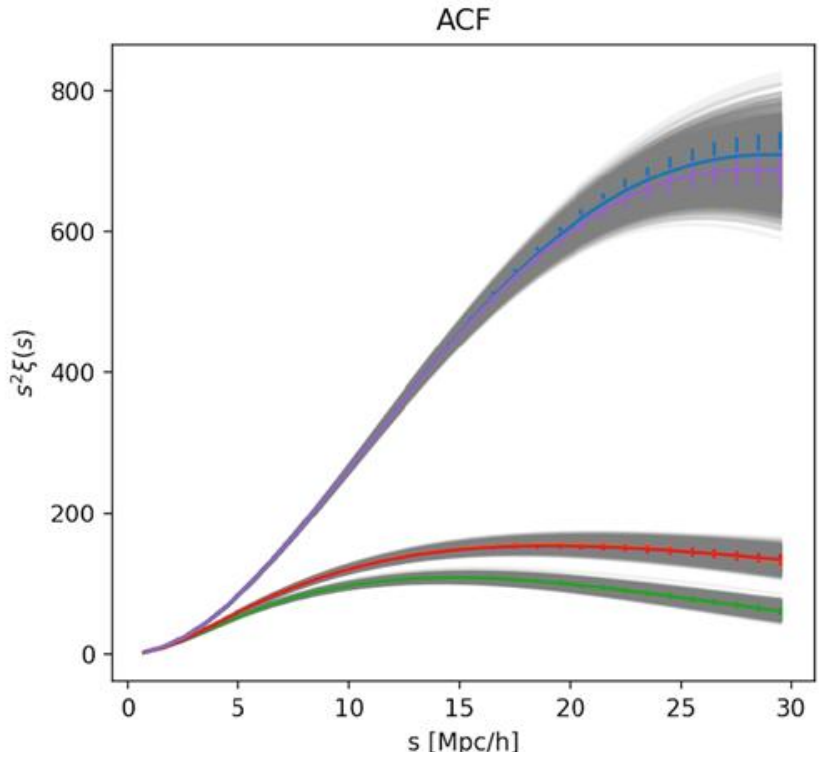
Covariance

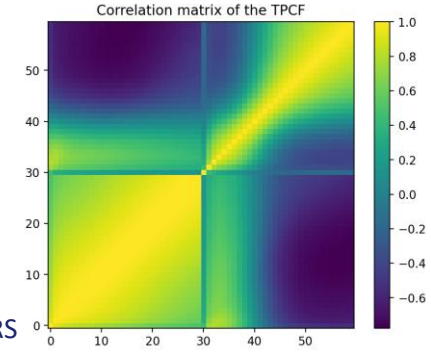
1 cosmo, 1 HOD, 1629 initial conditions

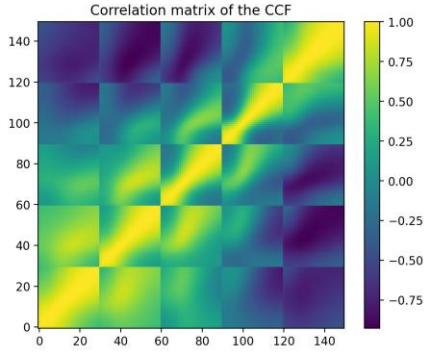
Error bars for 1629 small boxes

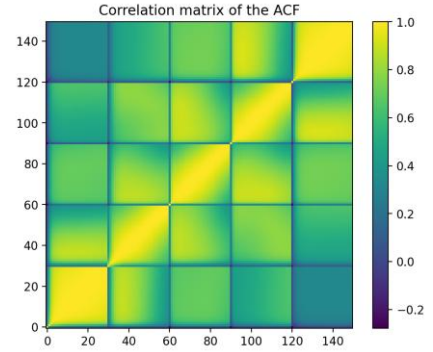












Emulator

Sunbird : Cuesta-Lazaro et al. 2023

Training set

Data

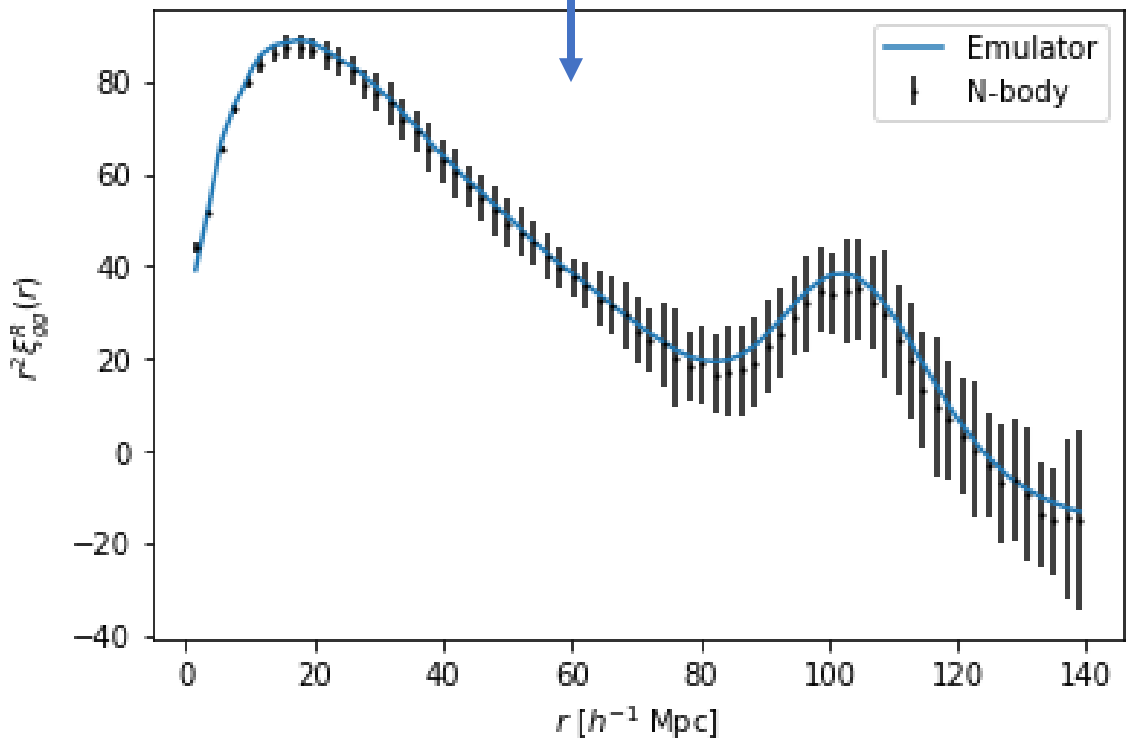
Parameters*

Emulator

Parameters*

Validation set

Simulation

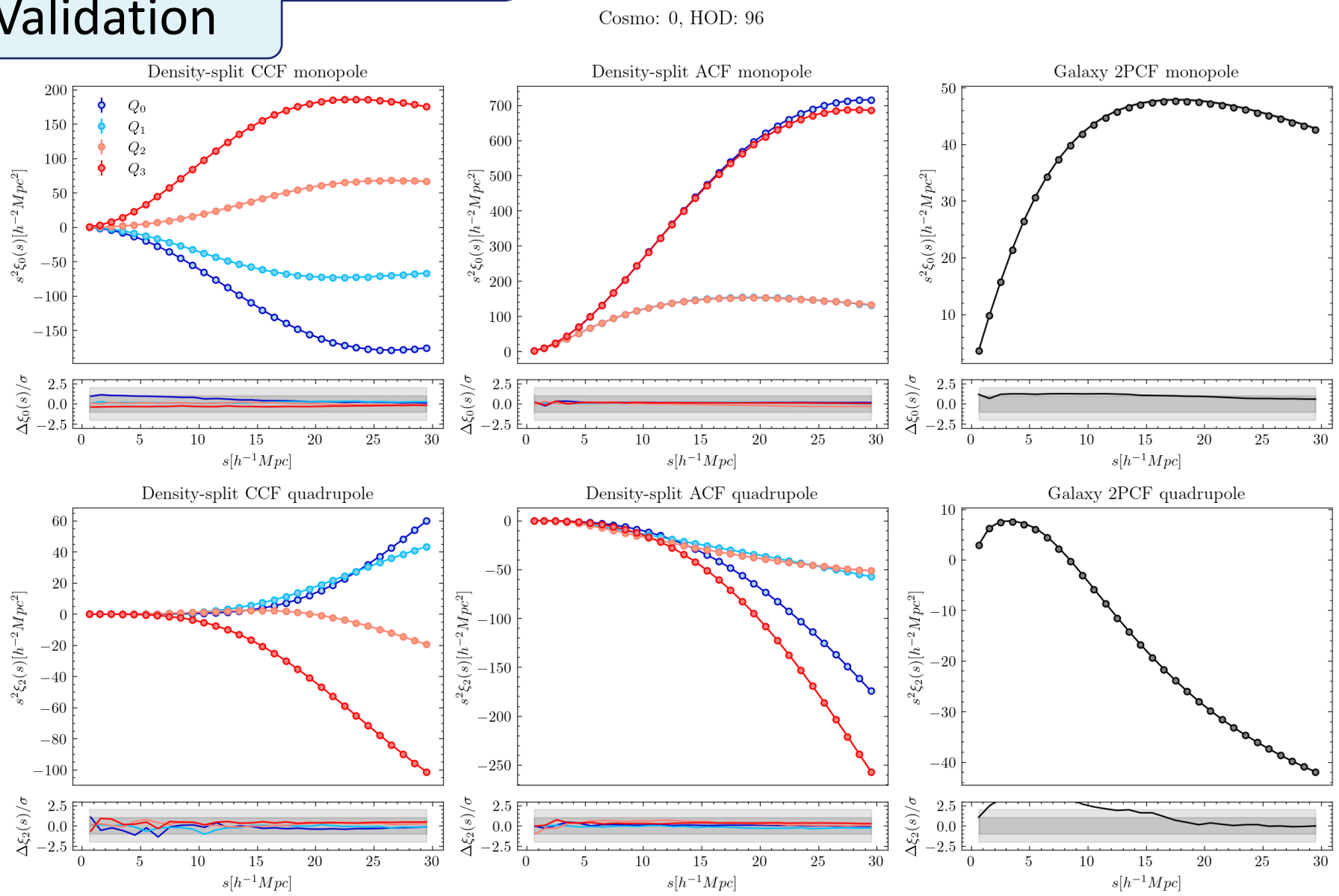


Data

*Cosmo, HOD

Emulator

Validation

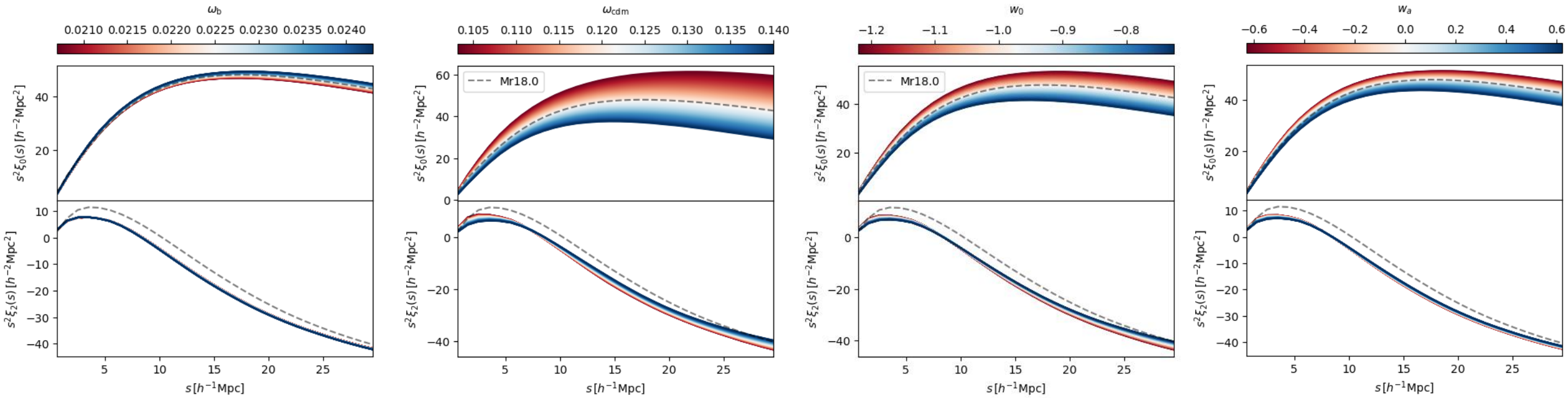


Emulator

Sensitivity

→ 2PCF only here

→ Variation of each parameter around bestfit mock



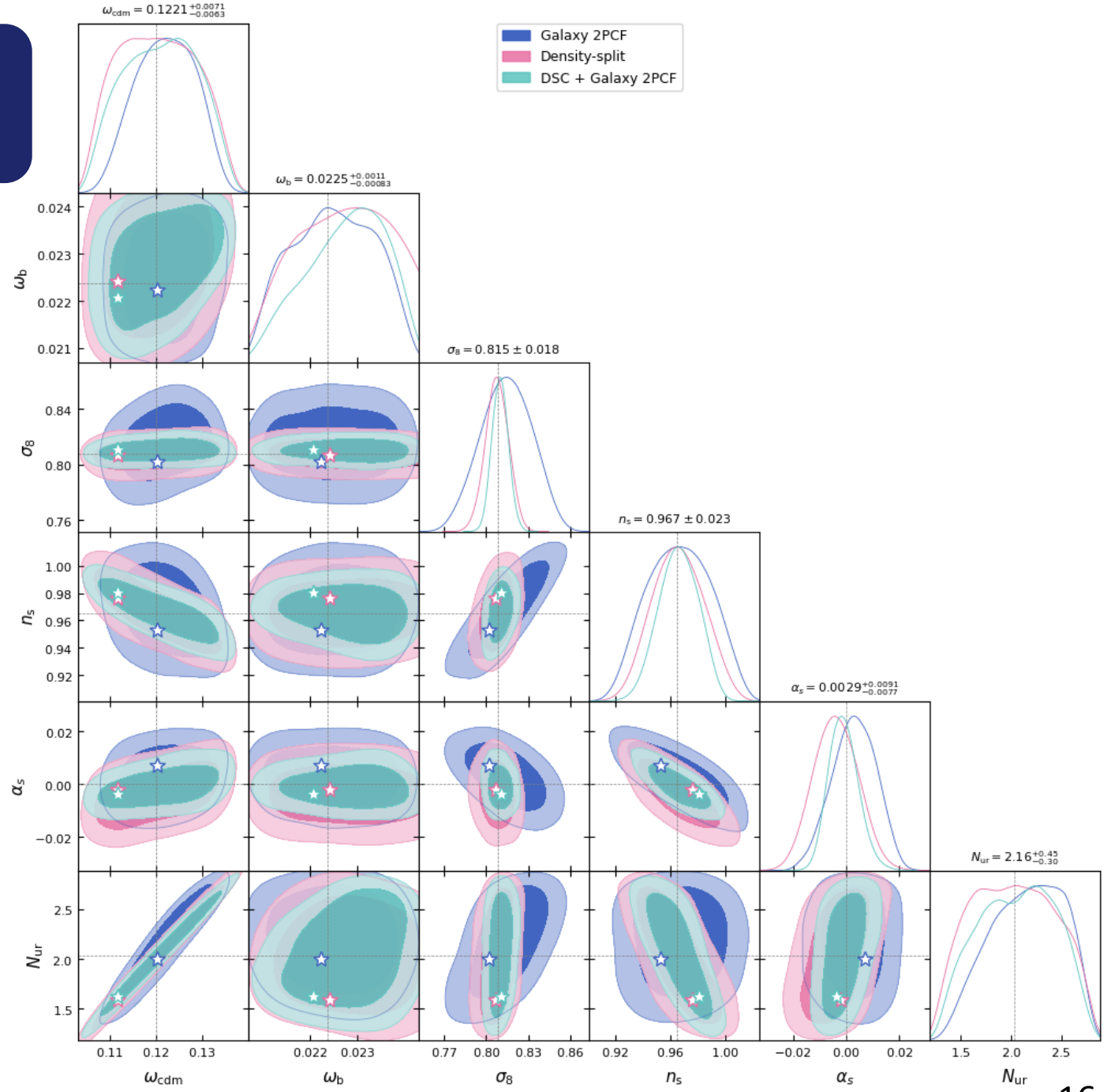
Cosmological inference

BGS validation set

- Fitting Cubic Validation BGS mocks
- Fitting Λ-CDM + « vanilla » HOD parameters
- Separation range : $0 < s < 30$ Mpc/h
- Marginalized over HOD parameters

→ 1σ recovery of the cosmology !

→ TPCF+DS gives tighter constraints !

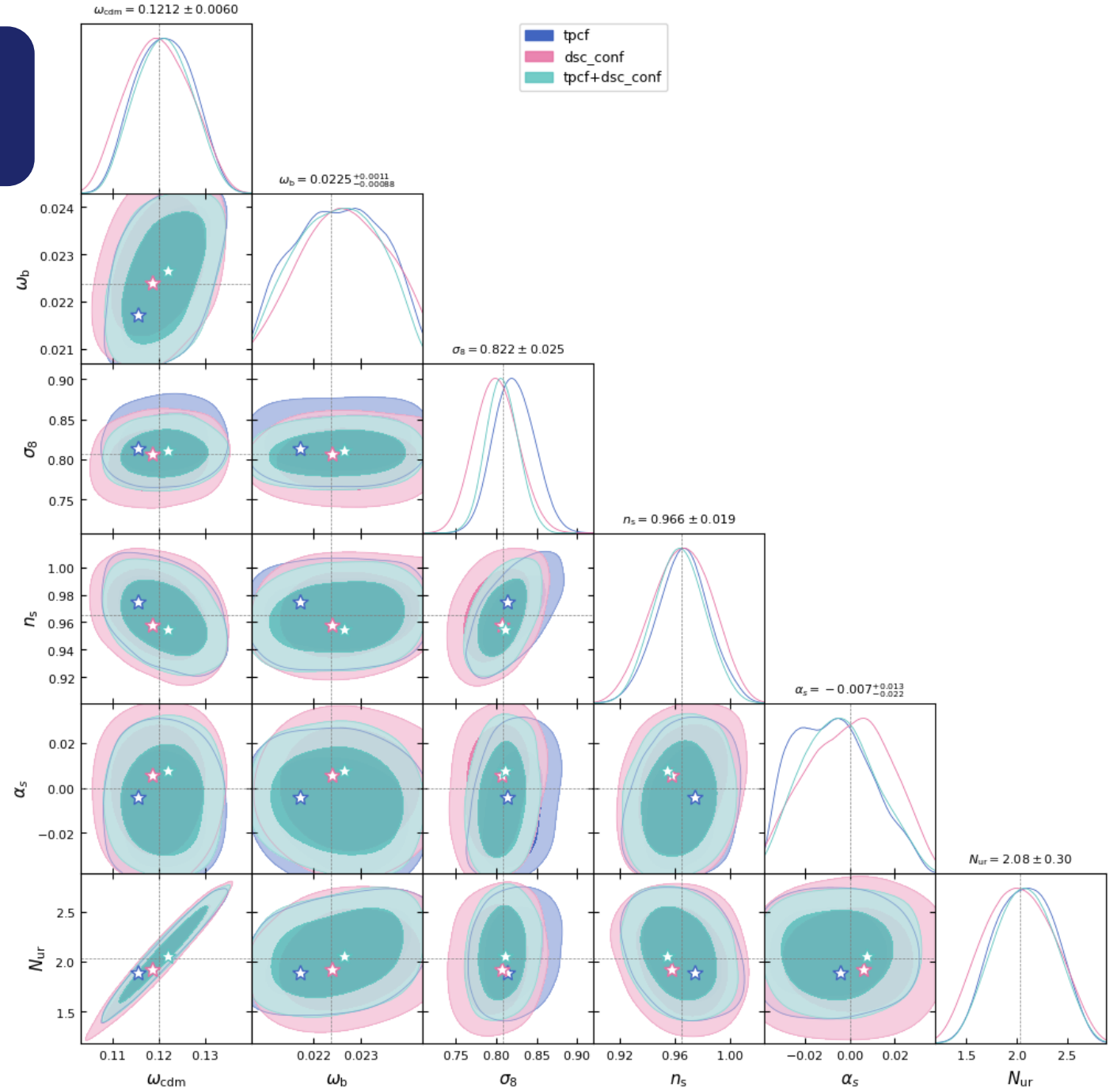


Cosmological inference

LRG DESI mocks

- Fitting Cubic Abacus LRG DESI DR1-mocks
- Fitting Λ-CDM + all HOD parameters
- Separation range : $0 < s < 150$ Mpc/h
- Marginalized over HOD parameters

→ 1σ recovery of the cosmology !

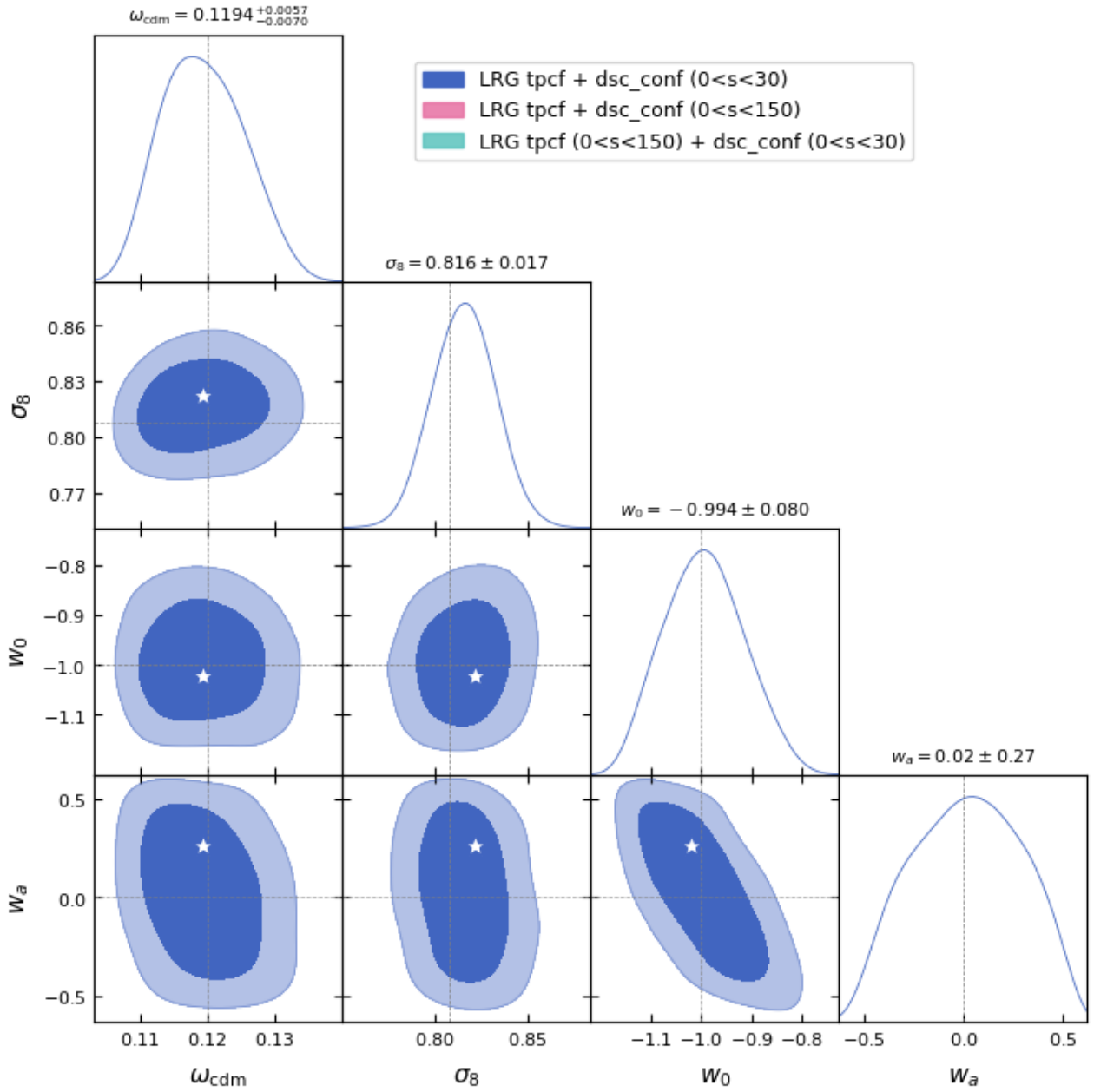


Cosmological inference

s scale choice

TPCF + DS : $0 < s < 30$ Mpc/h

- No larges scales
- Low computing time (*ct*)
- 1σ recovery : ✓

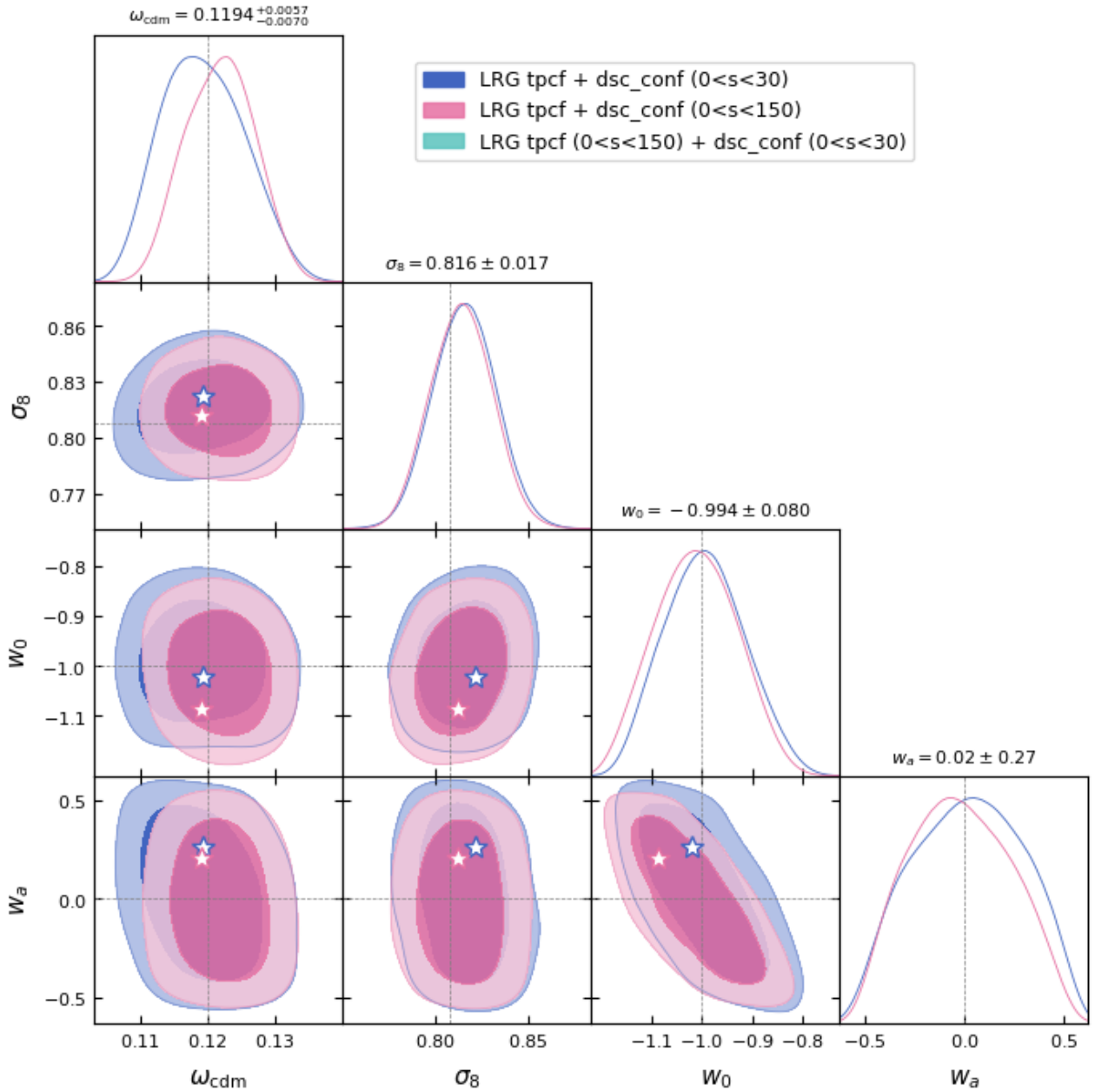


Cosmological inference

s scale choice

TPCF + DS : $0 < s < 150$ Mpc/h

- Large scales included
- High computing time (11 CFs to compute !)
- Slightly better constraints
- 1σ recovery : ✓



Cosmological inference

s scale choice

TPCF : $0 < s < 150$ Mpc/h

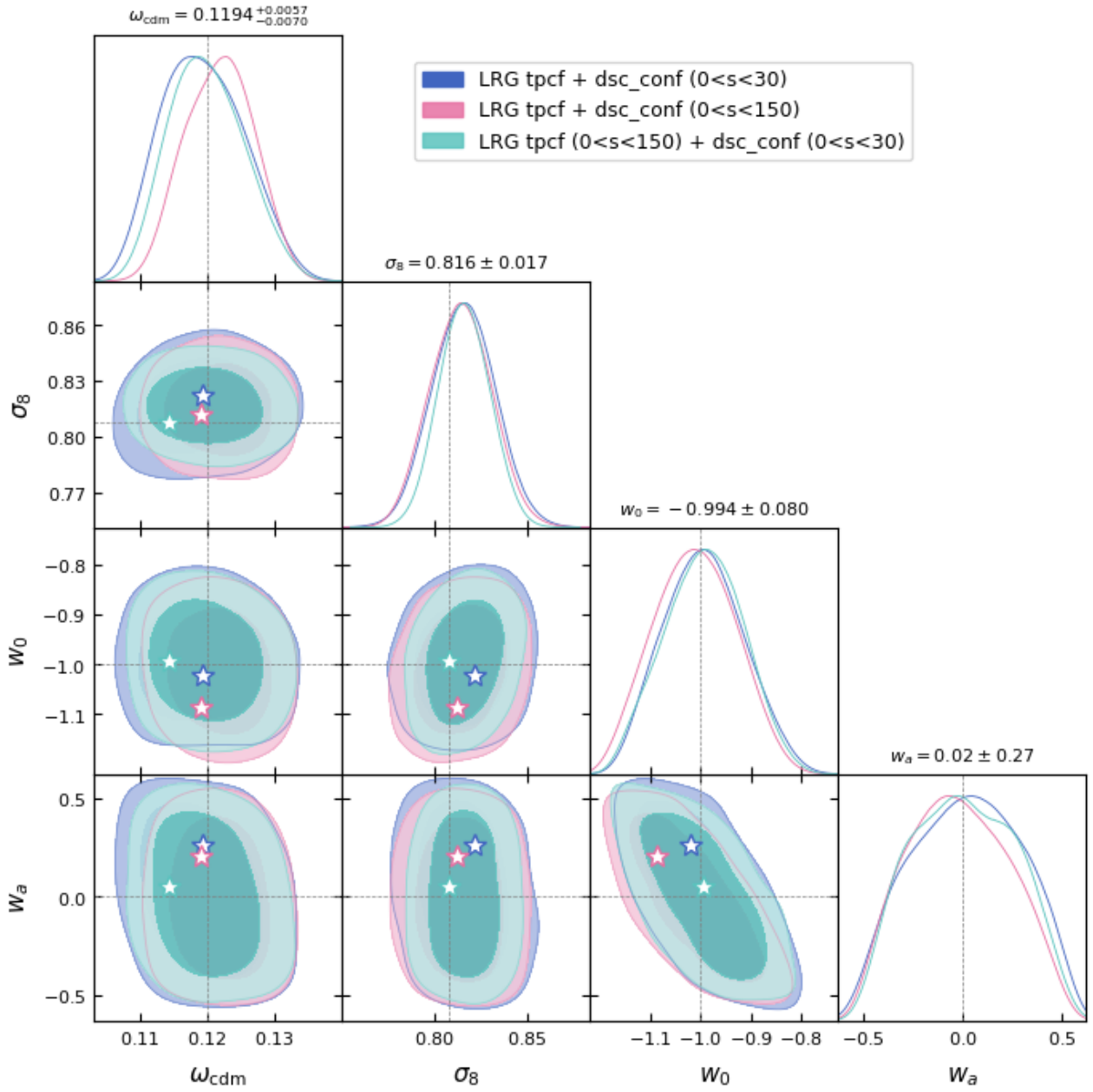
- Large scales included
- Only one per mock (low *ct* impact)

DS : $0 < s < 30$ Mpc/h

- No larges scales
  - DS info mostly comes from small scales
- 10 CFs per mock

Also :

- Good compromise : Time vs constraints
- 1σ recovery : ✓

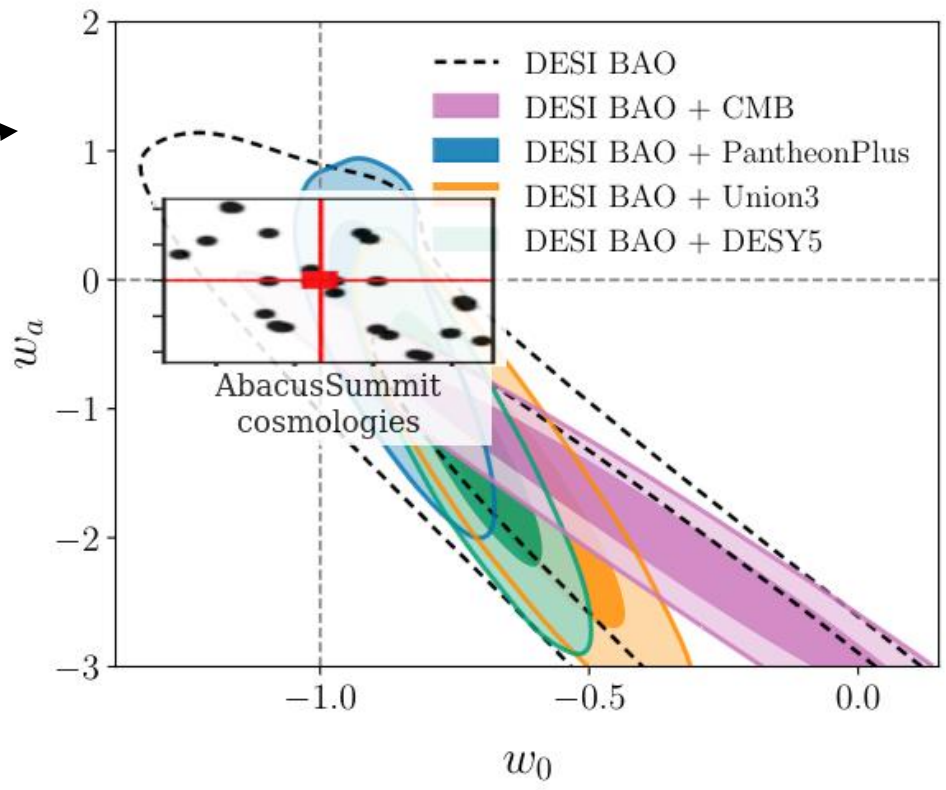


What's next ?

The next steps

- Finalize pipeline for cubic mocks
  - Prior space HOD → Physical motivations ?
  - Prior space cosmo → Need more simulations around DESI DR2 results !
- Towards application to real data !
  - Emulator trained on cutsky mocks
  - Include systematics (Fiber assignment, …)

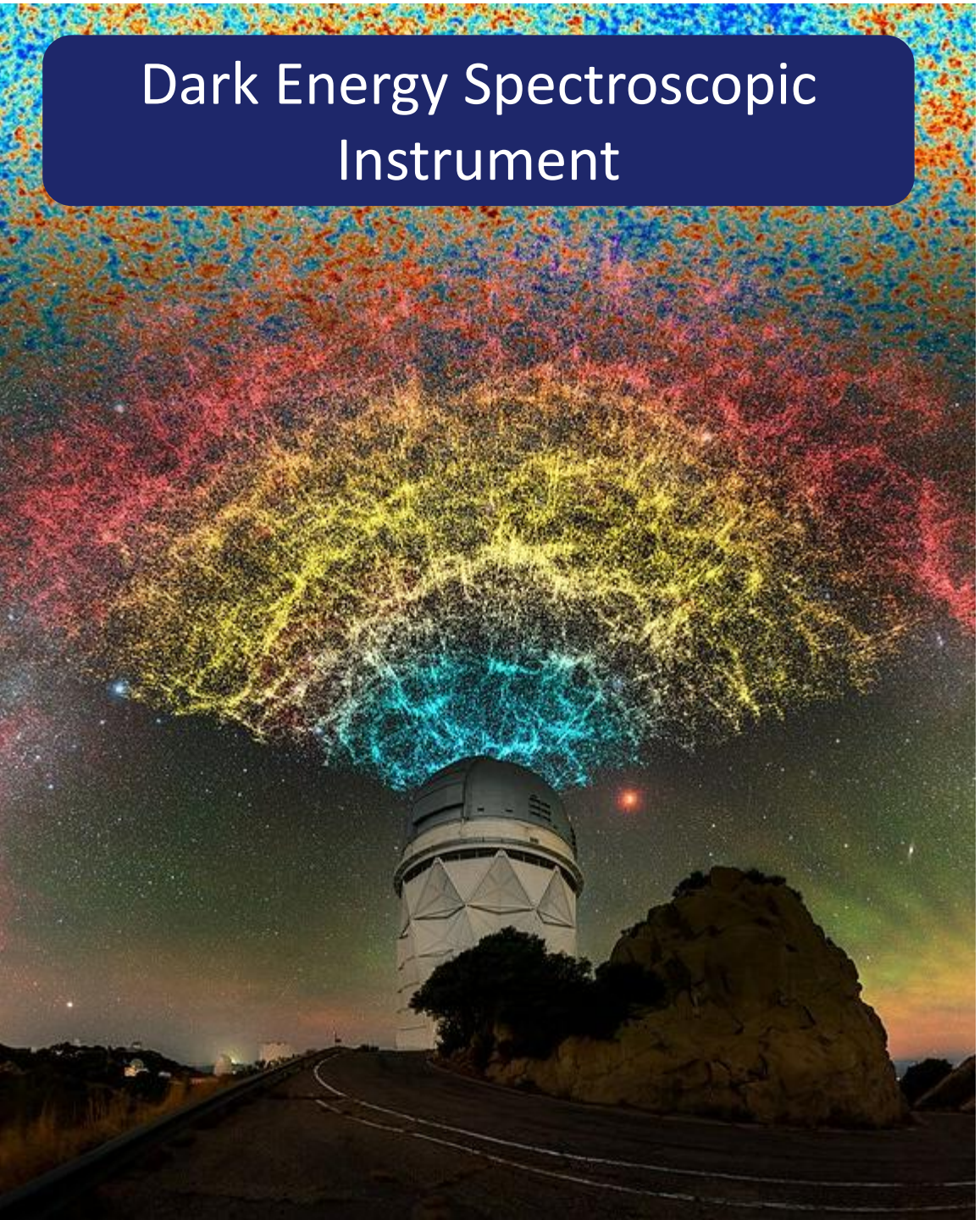
Source: Georgios Valogiannis



Thank you for your attention !

Backup slides

Dark Energy Spectroscopic Instrument



The instrument

- 1st new generation spectroscopic instrument
- 4m Mayall Telescope at Kitt Peak National Observatory
- 5000 spectra measured in one observation of approx. 20’ (previous gen : 1000/1h30)

The goal

- Constrain the dark energy parameters & gravity

Different galaxy samples

$[z > 1,5]$ **Quasars (3M)**

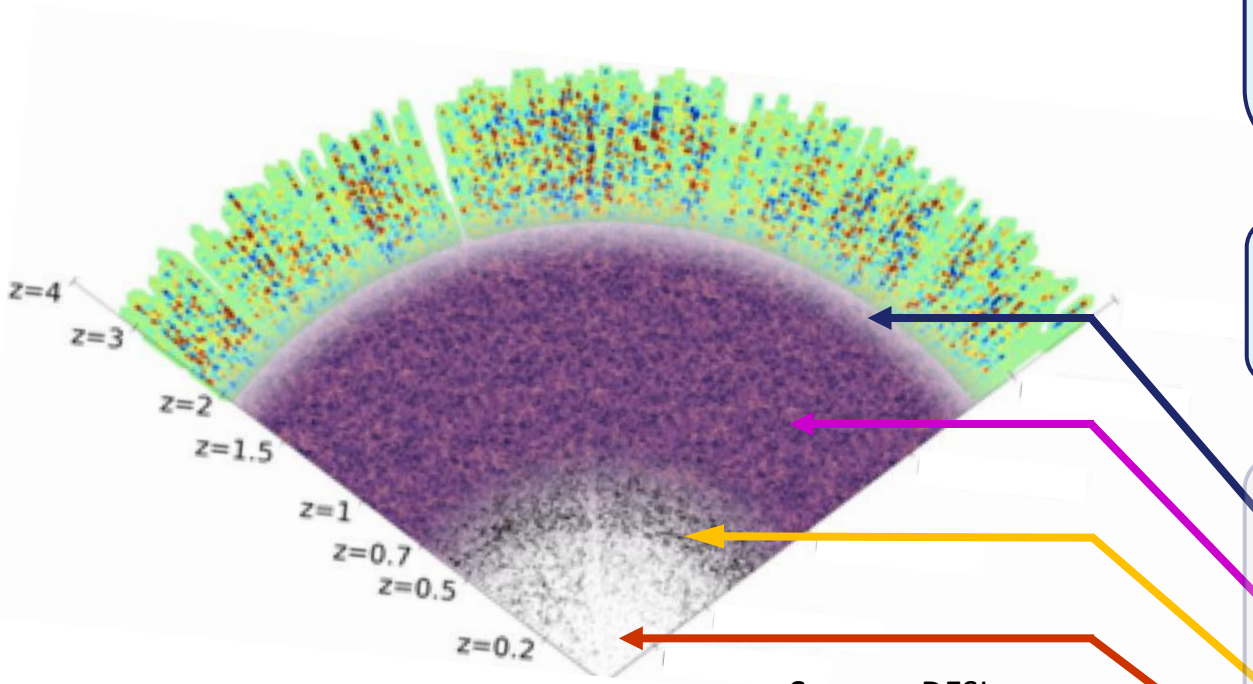
$[0,6 < z < 1,6]$ **ELG (Emission Line Galaxies, 16M)**

$[0,5 < z < 1,0]$ **LRG (Luminous Red Galaxies, 8M)**

$[z < 0,5]$ **BGS (Bright Galaxy Survey, 13M)**

cnrs LPNHE PARIS

Dark Energy Spectroscopic Instrument



Source : DESI

The instrument

- 1st new generation spectroscopic instrument
- 4m Mayall Telescope at Kitt Peak National Observatory
- 5000 spectra measured in one observation of approx. 20’ (previous gen : 1000/1h30)

The goal

- Constrain the dark energy parameters & gravity

Different galaxy samples

$[z > 1,5]$ **Quasars (3M)**

$[0,6 < z < 1,6]$ **ELG (Emission Line Galaxies, 16M)**

$[0,5 < z < 1,0]$ **LRG (Luminous Red Galaxies, 8M)**

$[z < 0,5]$ **BGS (Bright Galaxy Survey, 13M)**

Total : 60M galaxy spectra over 7 years

Data acquisition started in May 2021

First paper (DR1) published in April

Theory

HOD Model

Halo Occupation Distribution

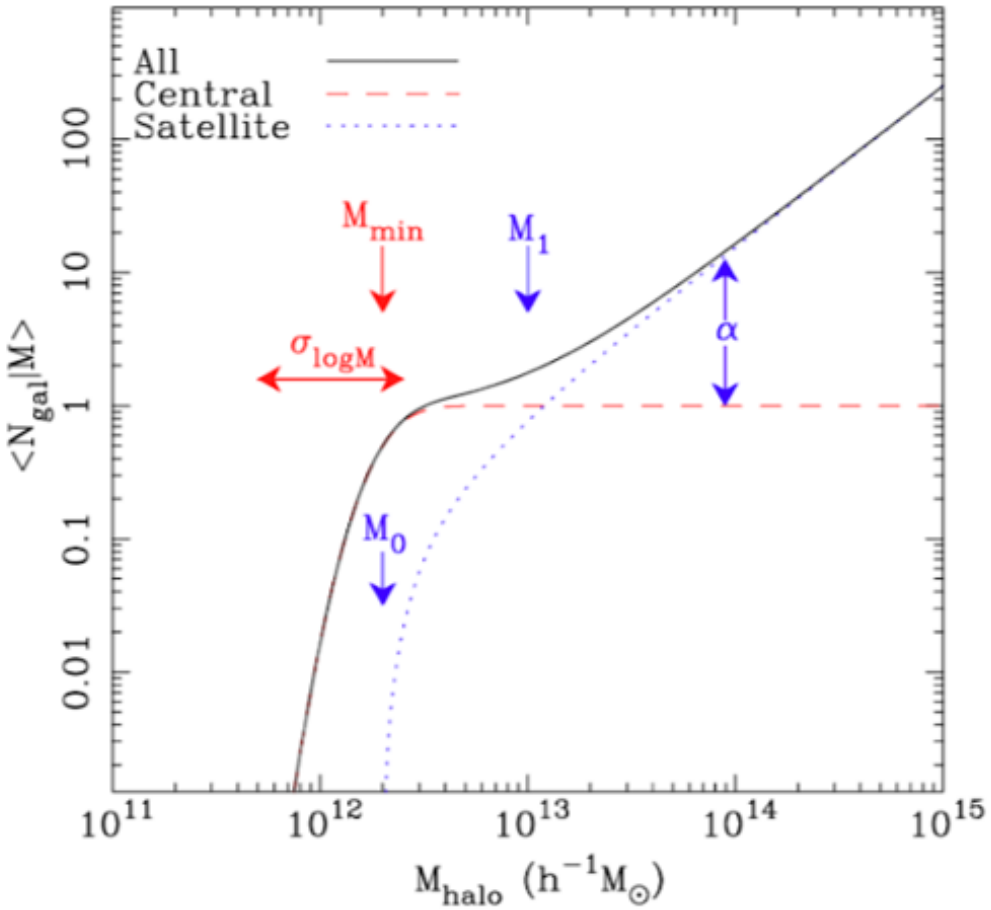
$$\langle N_c \rangle (M) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\log M - \log M_{\min}}{\sigma_{\log M}}\right)\right)$$

$$\langle N_s \rangle (M) = \langle N_c \rangle (M).\left(\frac{M - M_0}{M_1}\right)^{\alpha}$$

Notes

$\{M_{min}, \sigma_{\log M}, M_0, M_1, \alpha\}$
are parameters we need to adjust

Source : Wake et al. (2011)



Covariance matrix

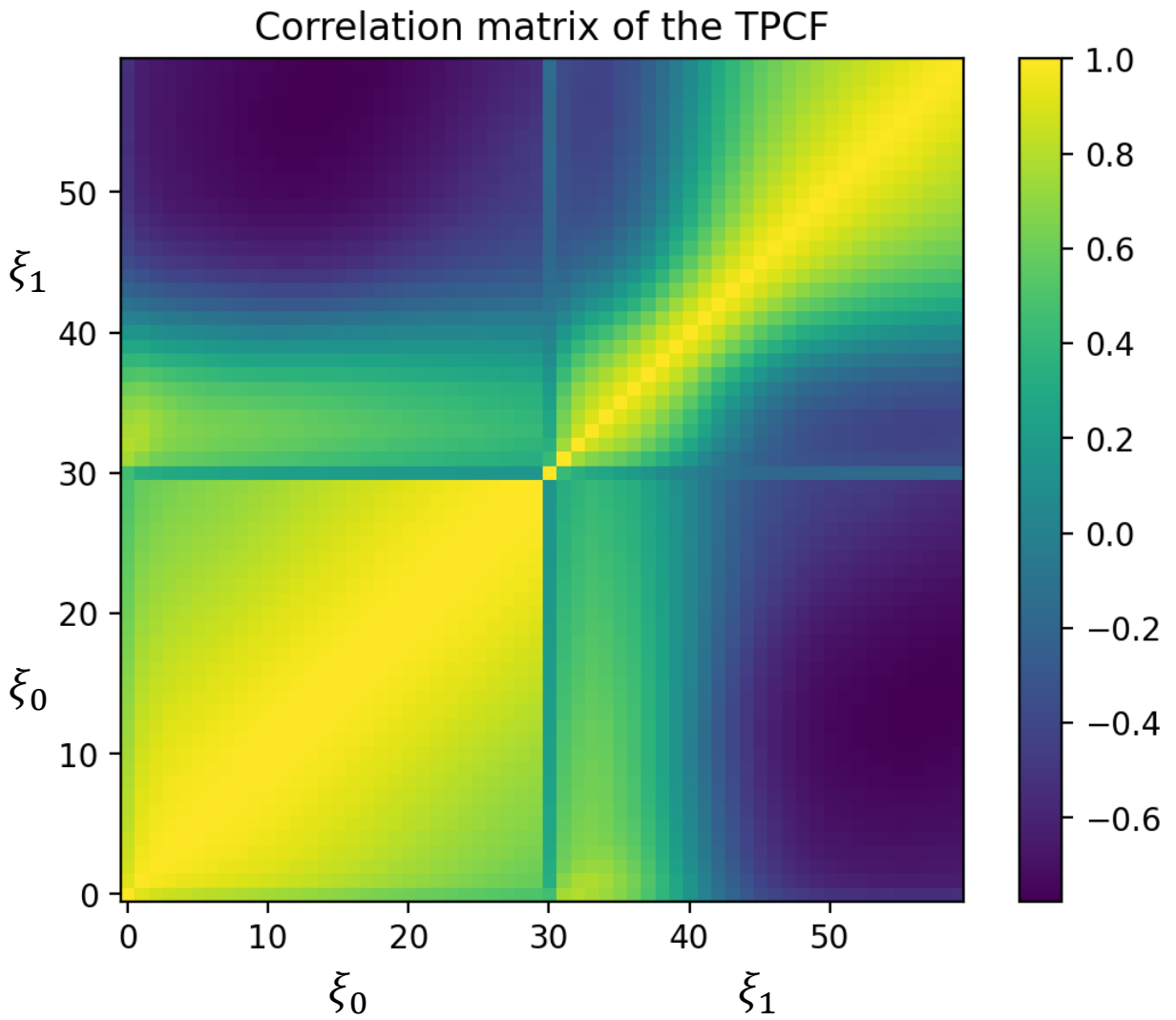
Covariance matrix

$$Cov_{ij} = < x_i x_j > - < x_i >< x_j >$$

Correlation matrix

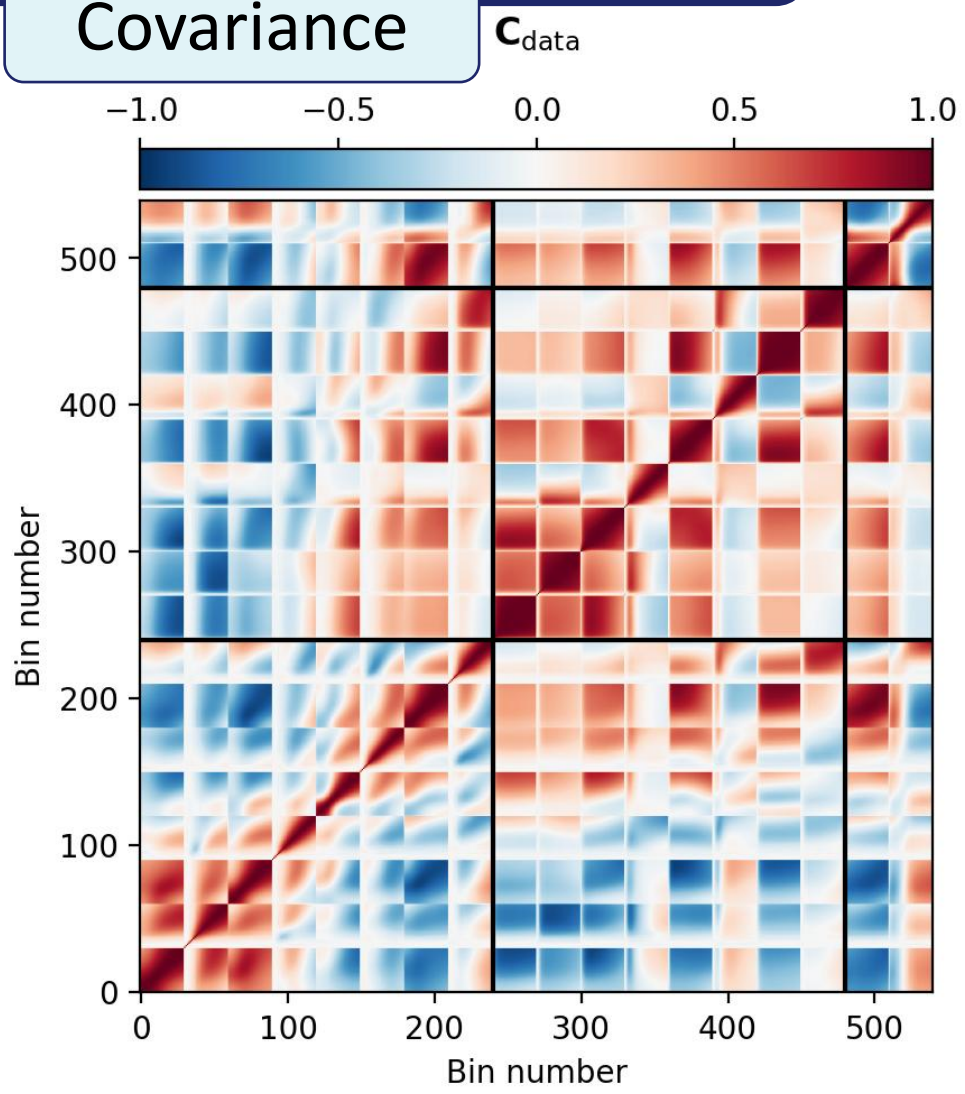
$$Corr_{jj} = \frac{Cov_{ij}}{\sigma_i \sigma_j}$$

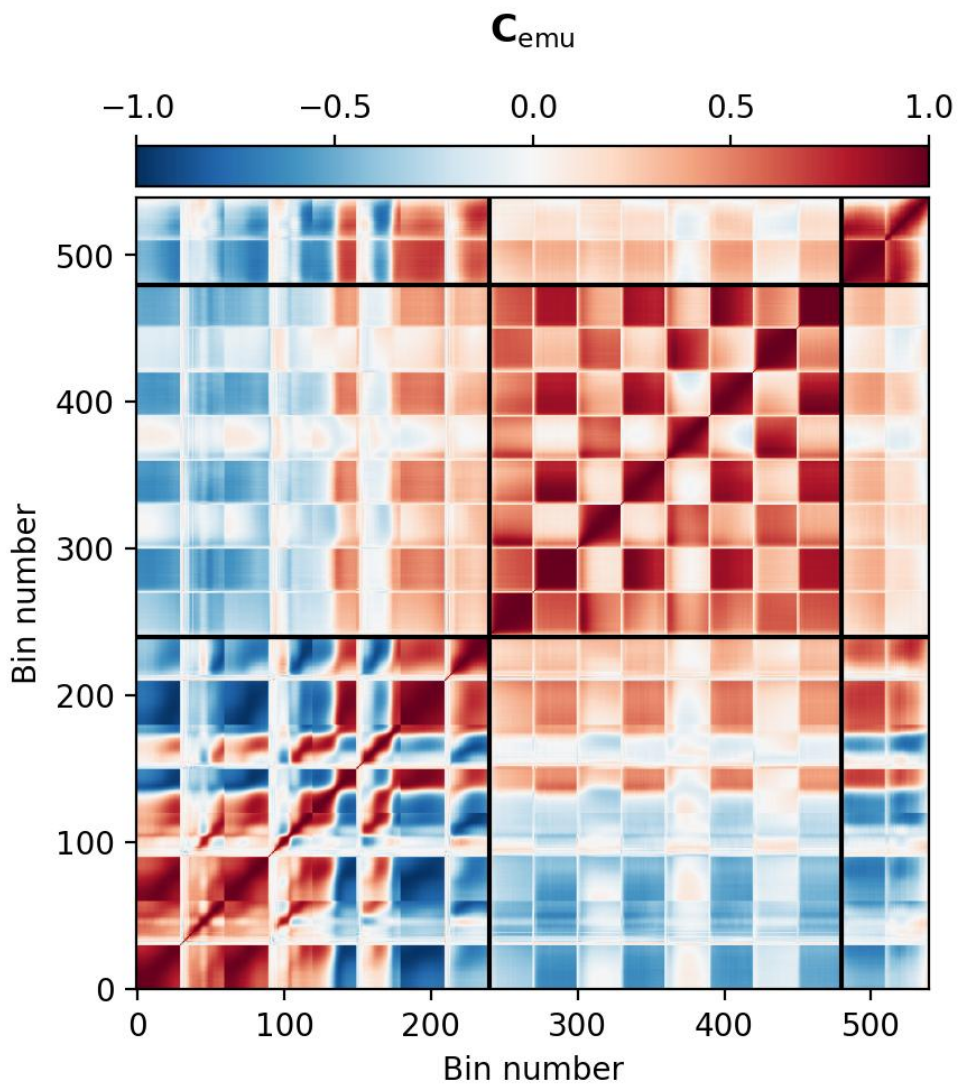
$$\sigma_i = \sqrt{< x_i^2 > - < x_i >^2}$$



Emulator

Covariance





Simulations

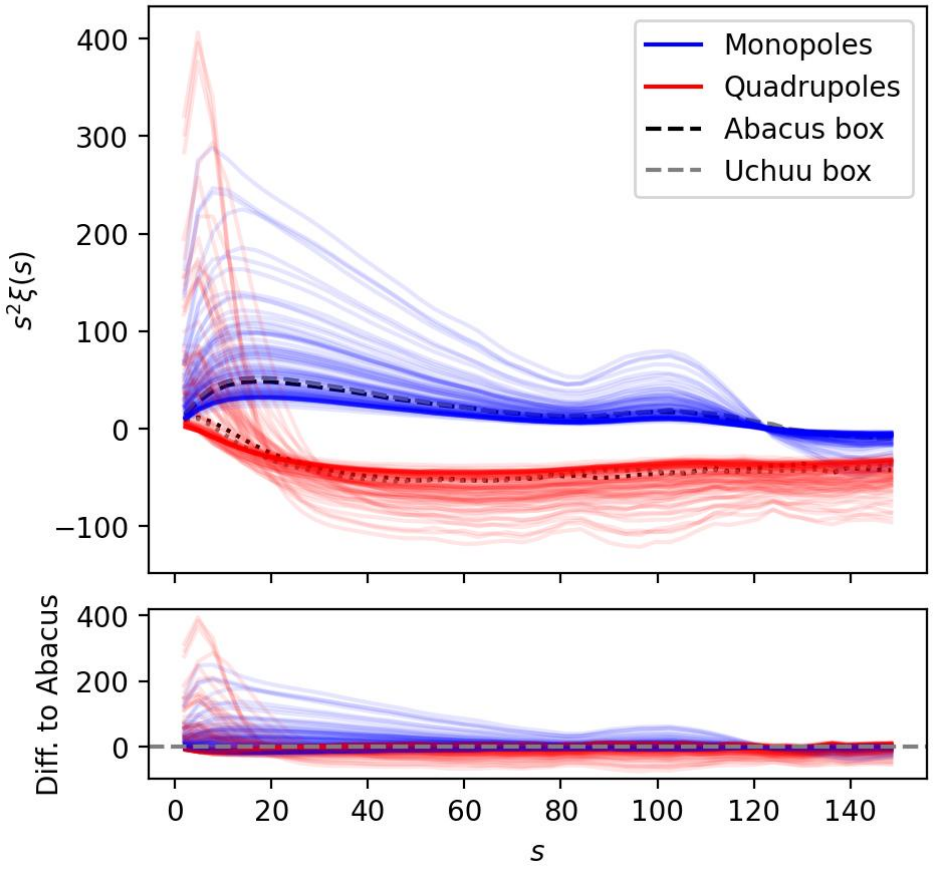
BGS challenges

Nb of cosmo	85			
Nb of HOD (per cosmo)	100			
Nb of los	3			
	Time (s) for 1	Total (s)	Total (min)	Total (h)
Time for loading cosmo	340	28900	481,666667	8,02777778
Time for populating HOD	31	263500	4391,66667	73,1944444
Computing density	27	229500	3825	63,75
Computing quantiles	130	39000	650	10,8333333
Time for 2PCF (s<150)	1485	37867500	631125	10518,75
Total estimated time (h)	115862,0556			
2PCF for s<50	80	2040000	34000	566,666667
Total estimated time (h)	6389,138889			
2PCF for s<30	25	637500	10625	177,083333
Total estimated time (h)	2103,722222			

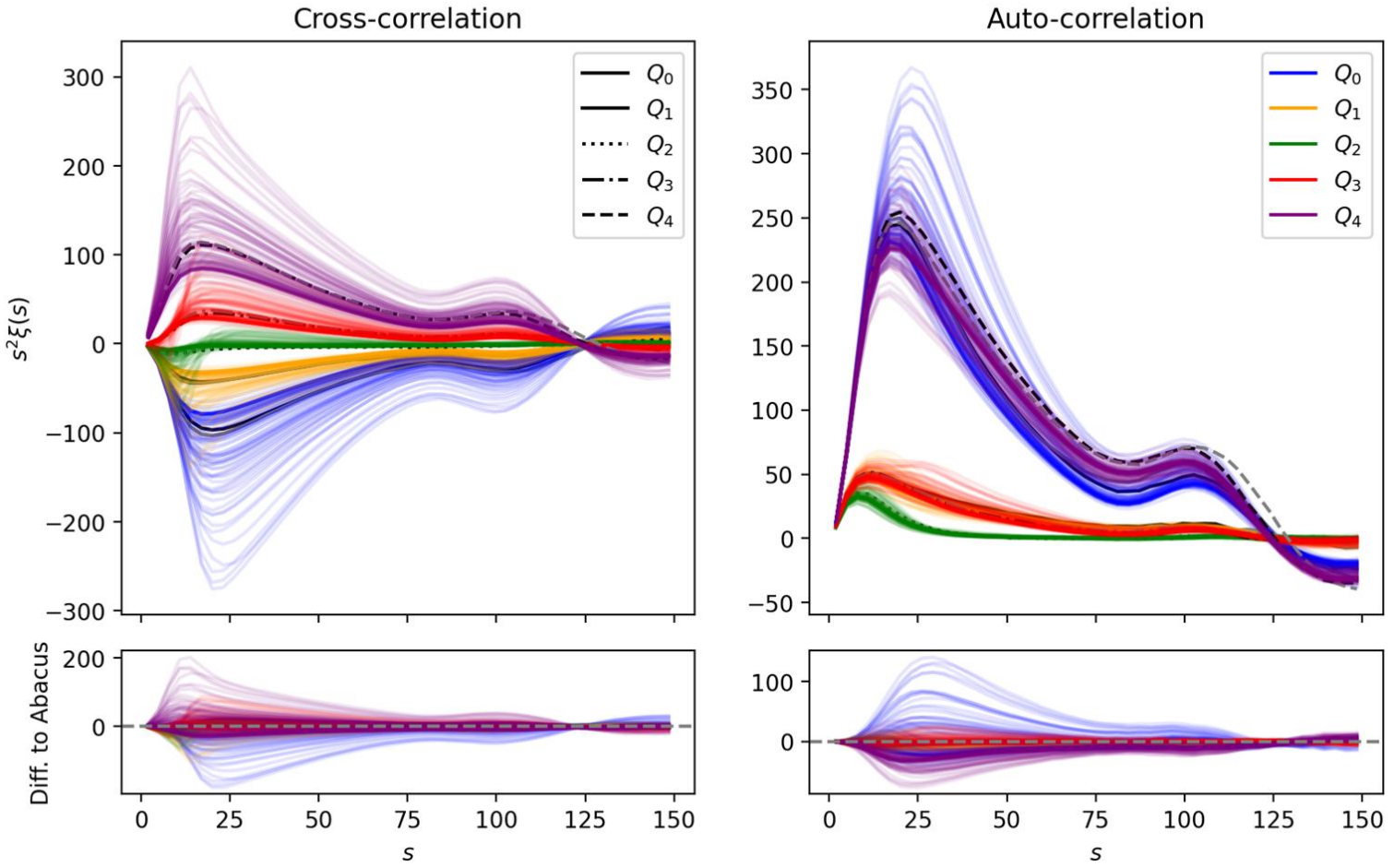
Simulations

Statistics

2PCF of 93 HOD catalogs

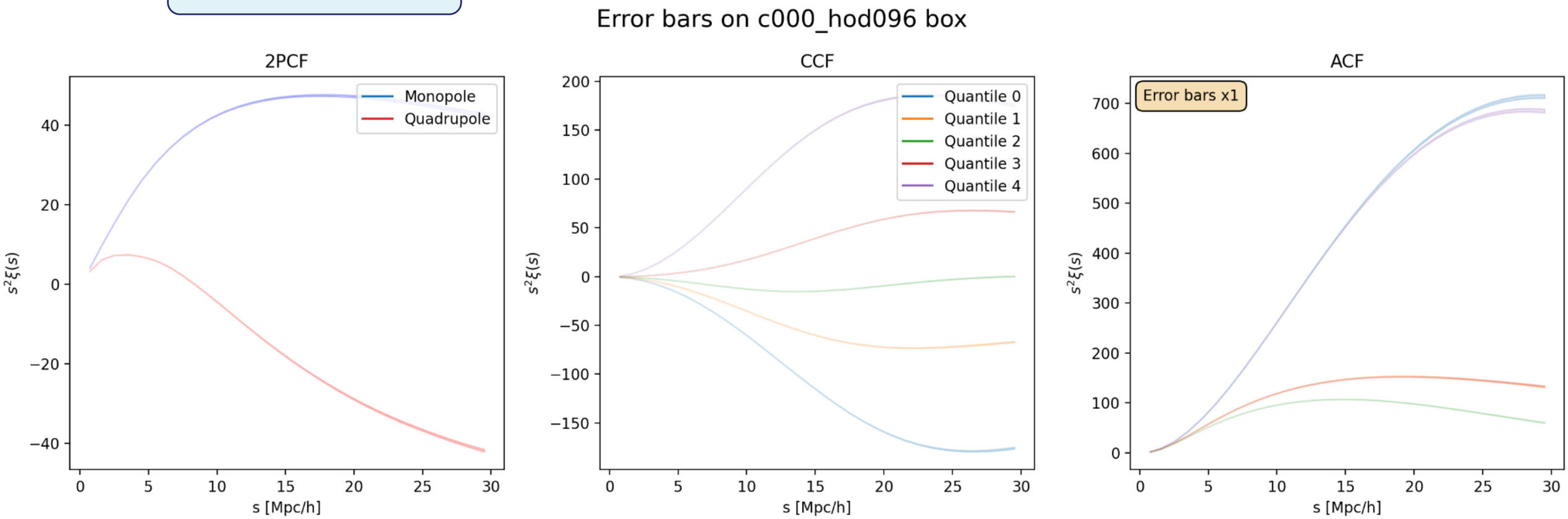


DS monopole of 93 HOD catalogs (Uchuu in gray, Abacus in black)



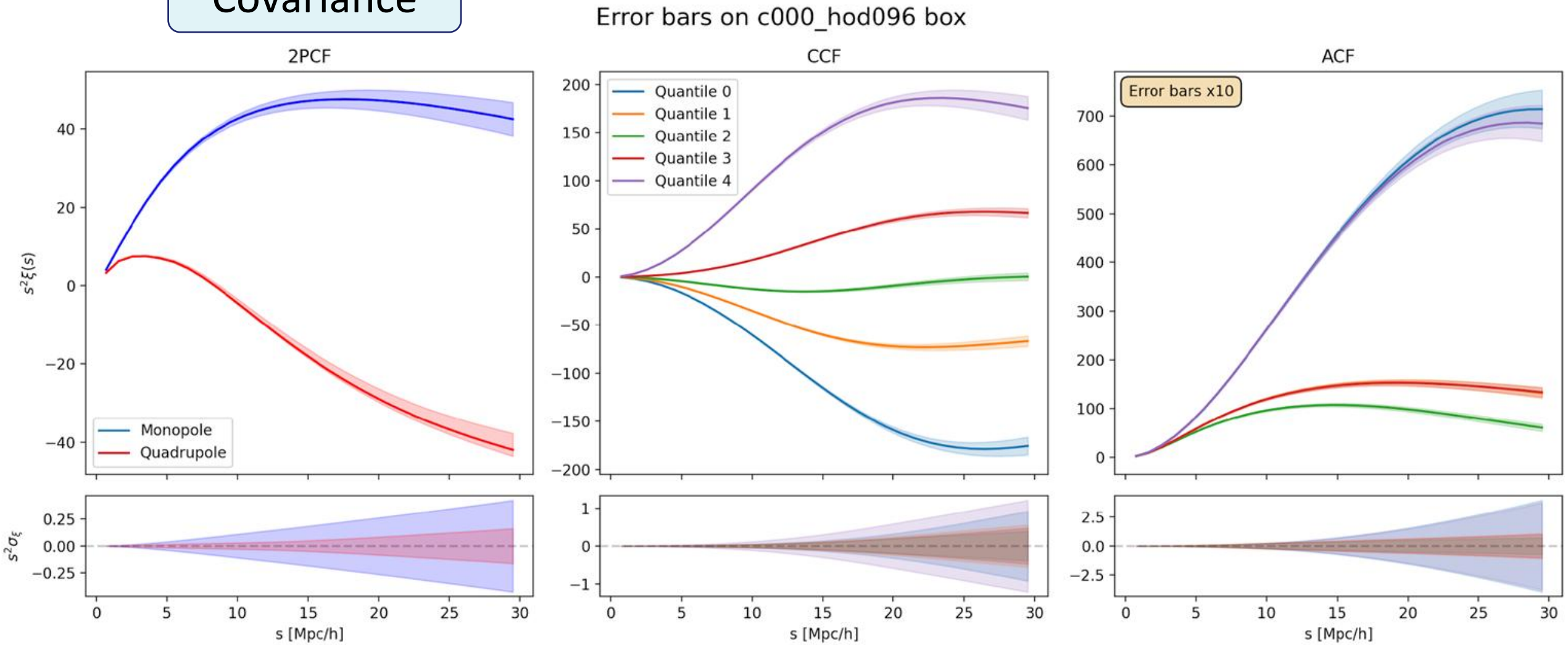
Simulations

Covariance



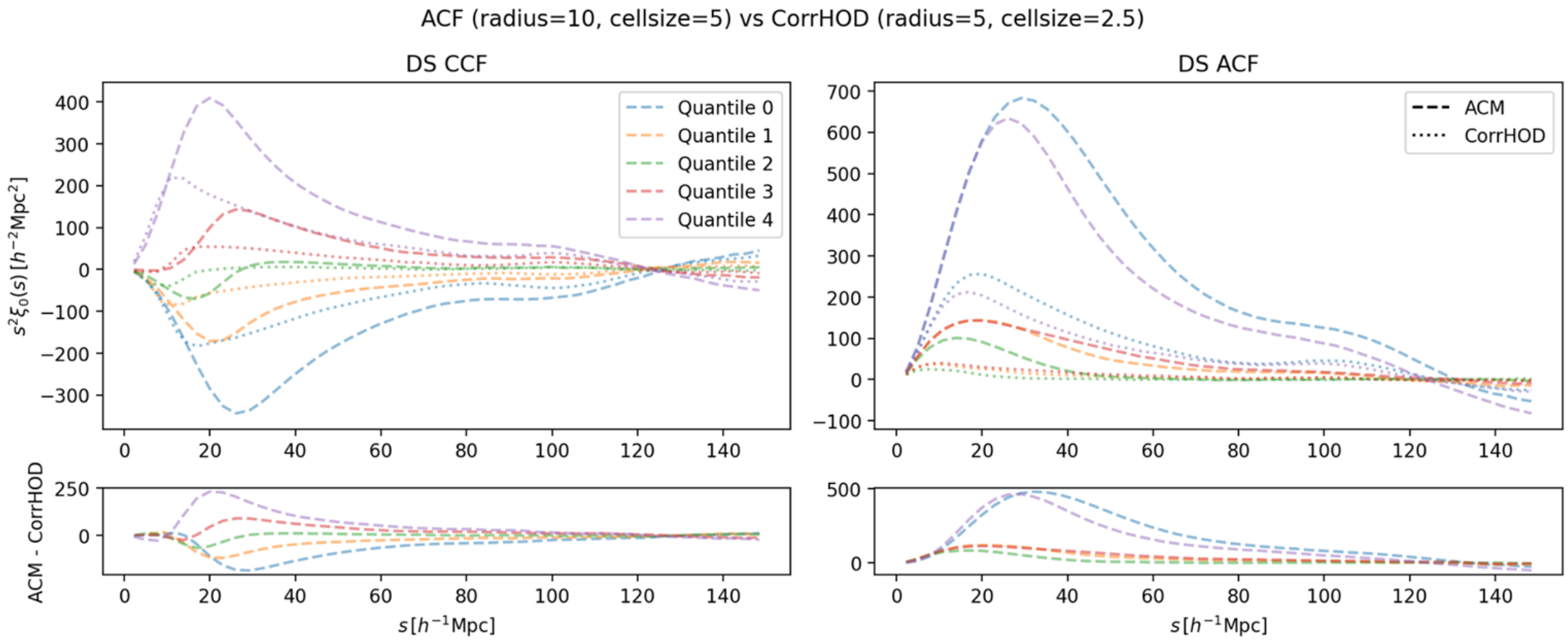
Simulations

Covariance

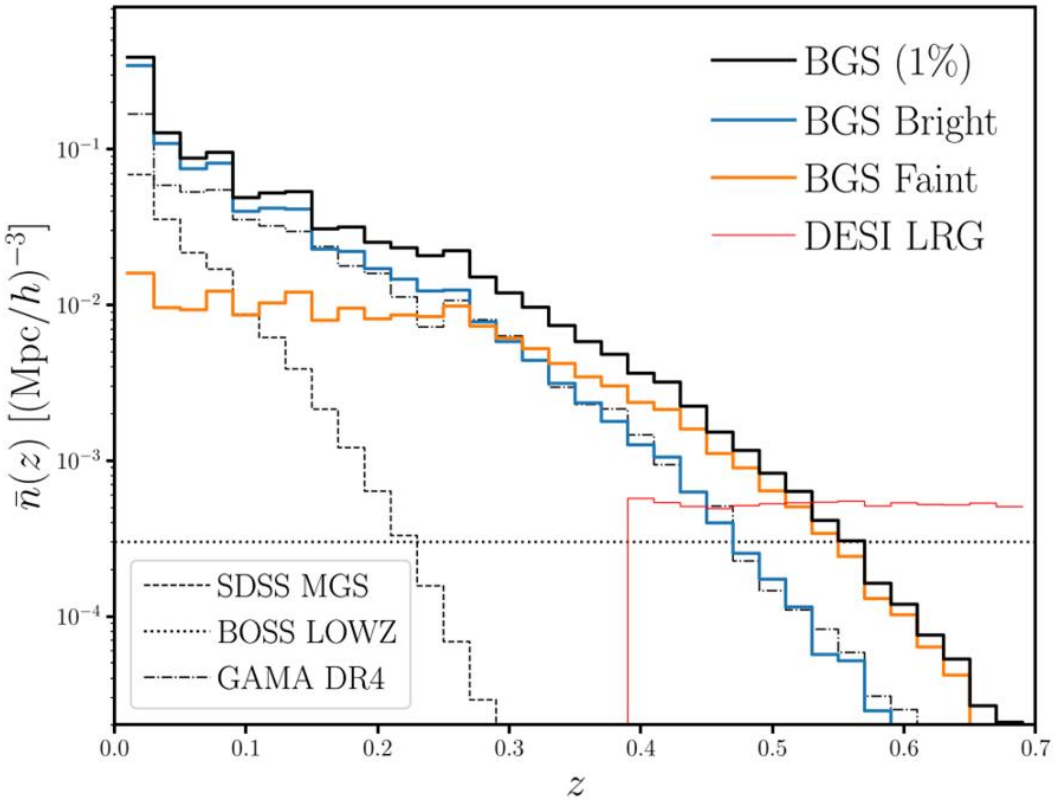


Other parameters

R_s impact



BGS Density



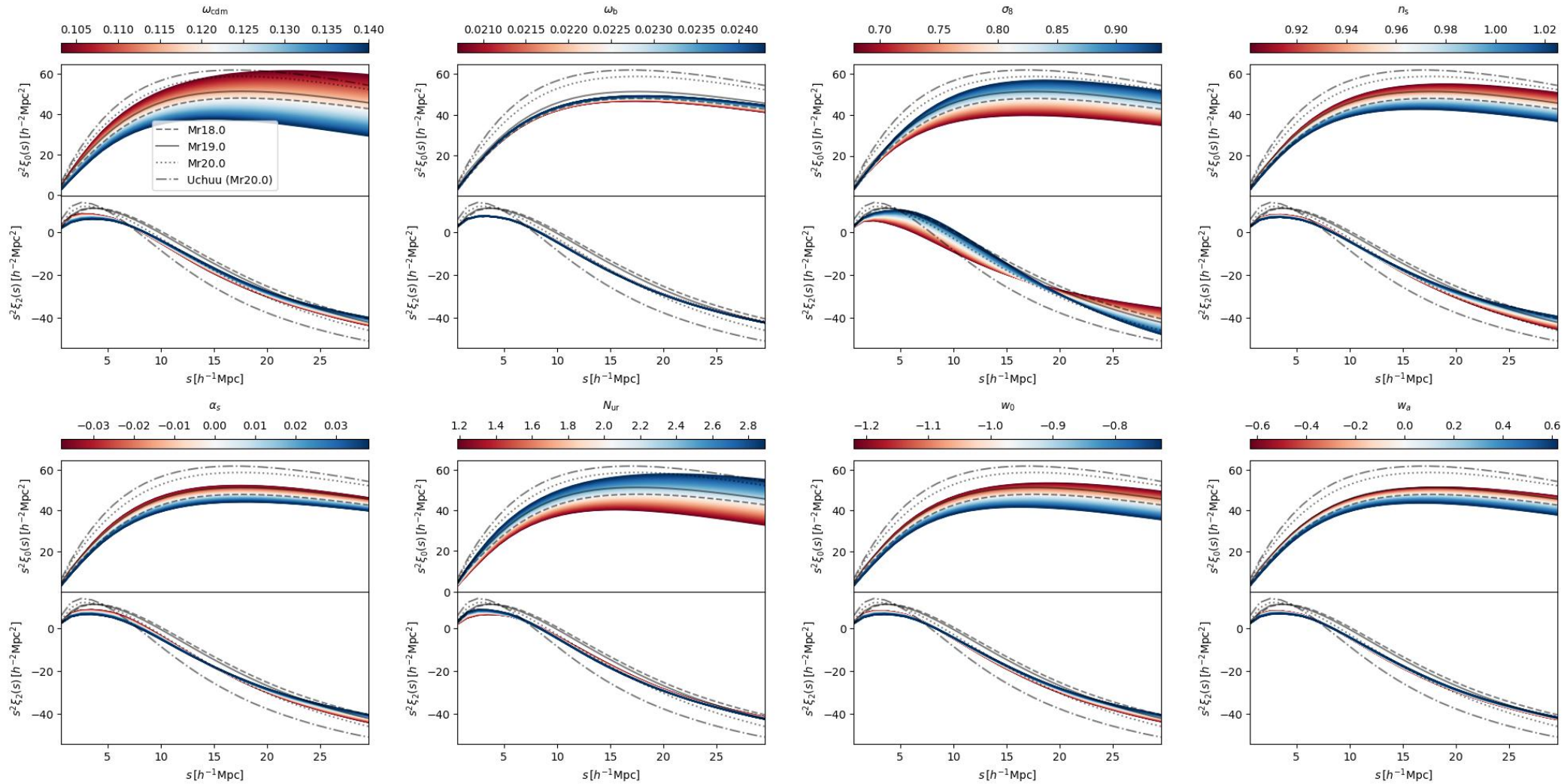
Source : Hahn et al. (2023)

Emulator

Sensitivity (cosmo)

→ 2PCF only here

→ Variation of each parameter around bestfit mock

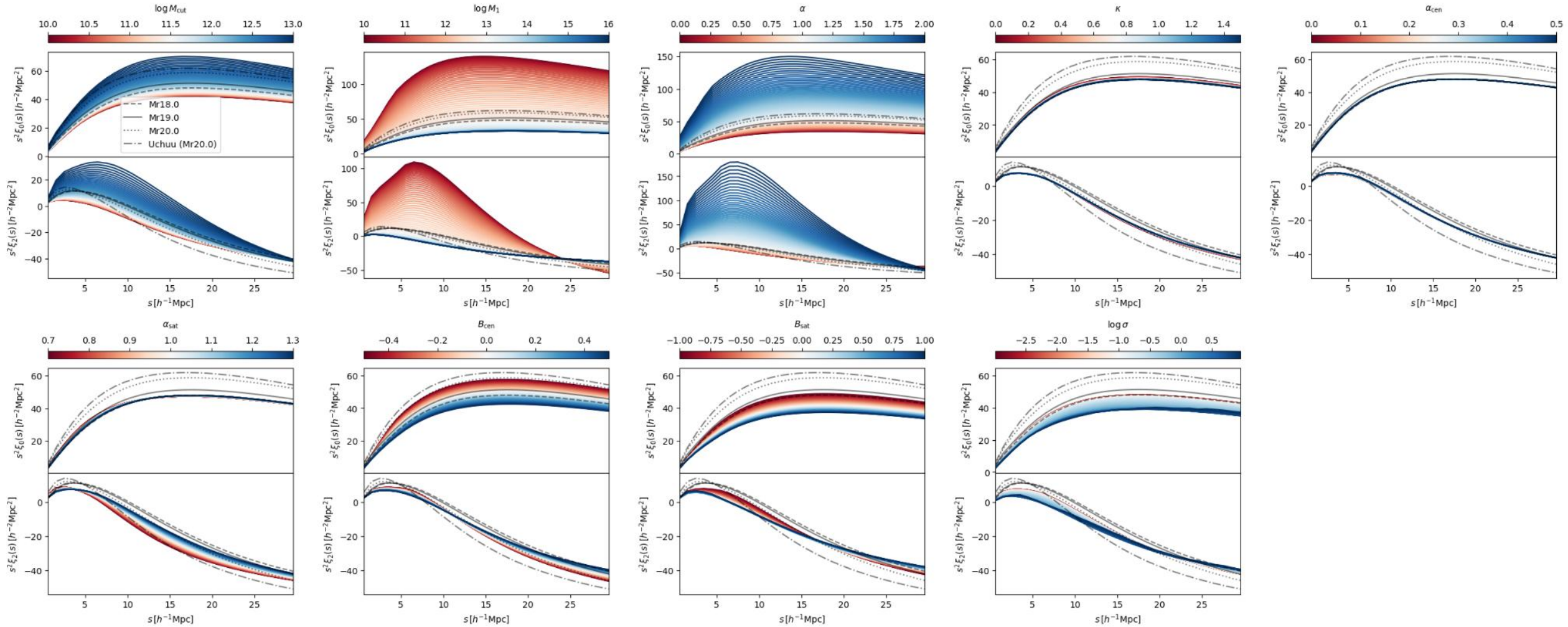


Emulator

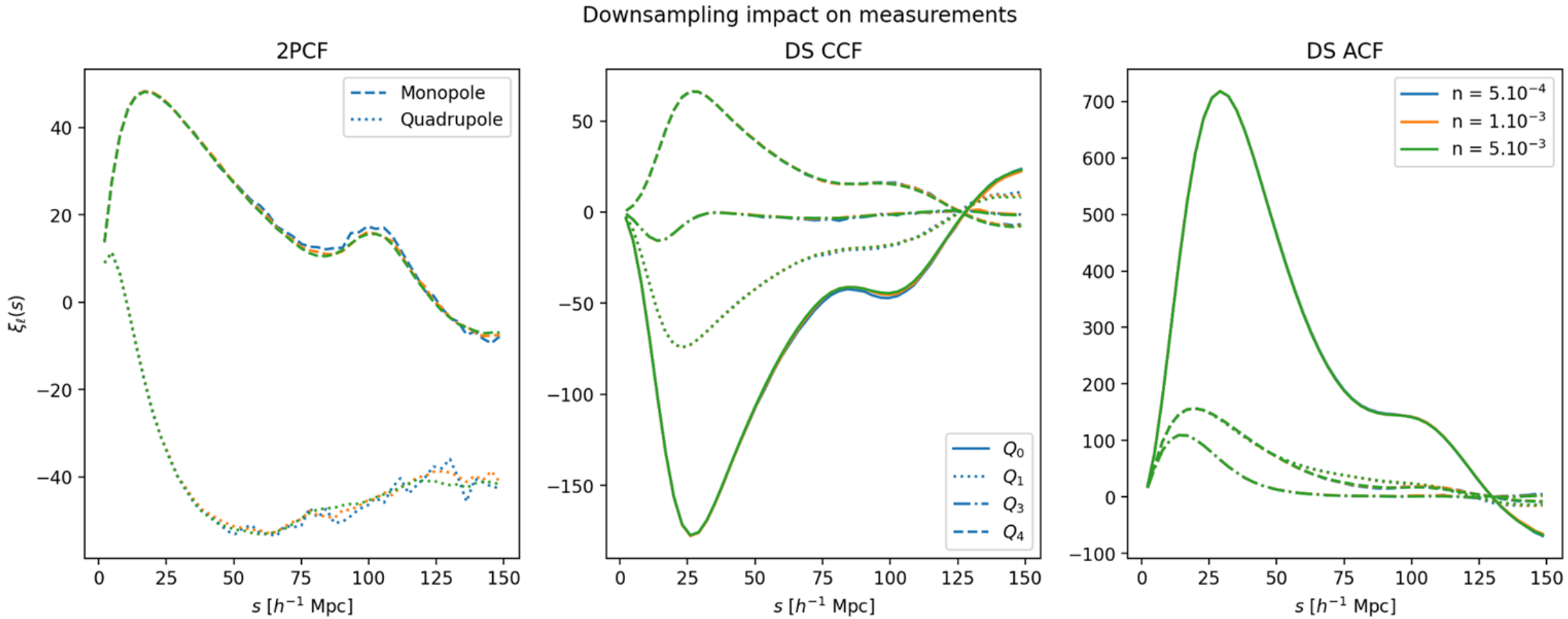
Sensitivity (HOD)

→ 2PCF only here

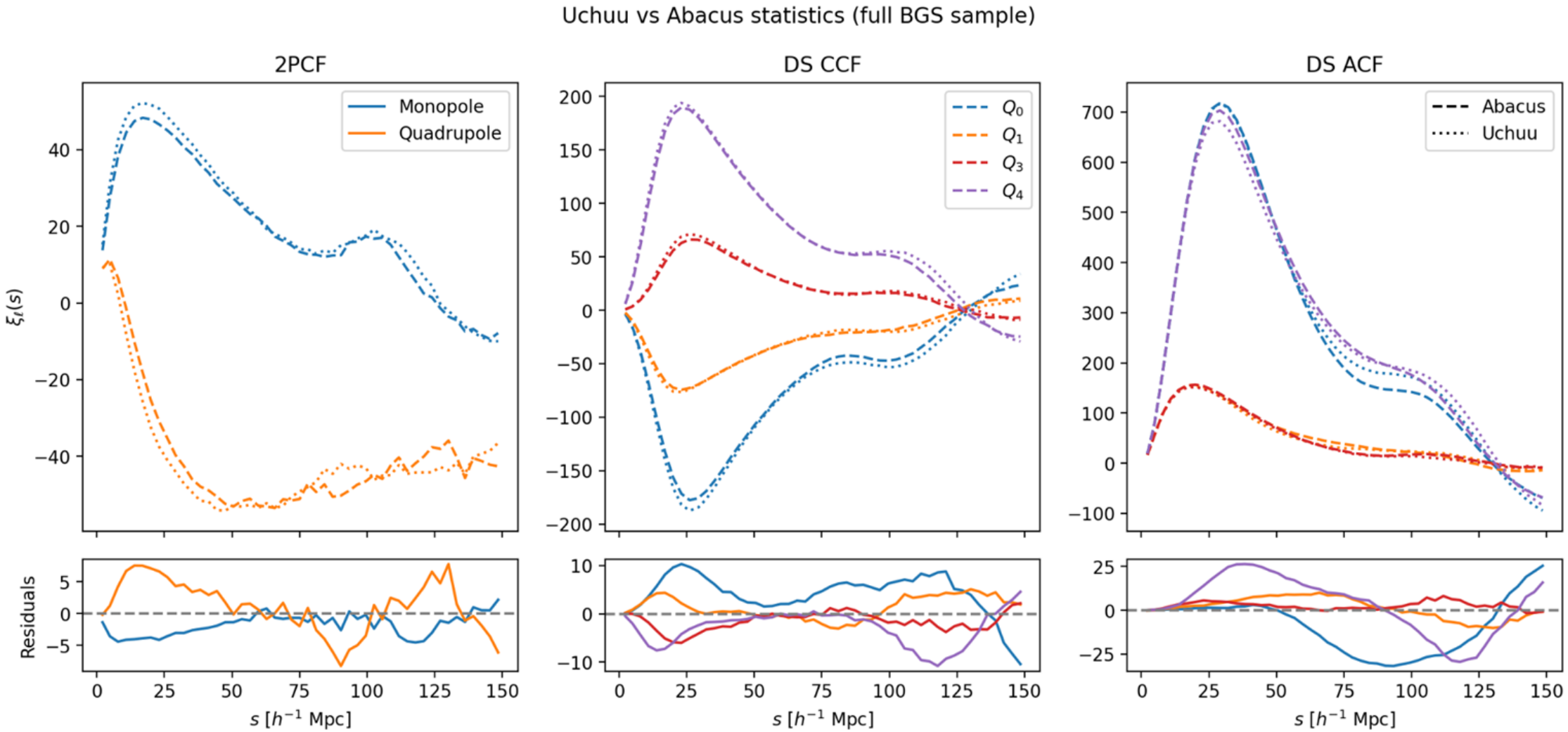
→ Variation of each parameter around bestfit mock



Data downsampling



Other Mocks


Uchuu vs Abacus statistics (full BGS sample)
2PCF
DS CCF
DS ACF
Monopole
Quadrupole
Q_0
Q_1
Q_3
Q_4
Abacus
Uchuu
$\xi_\ell(s)$
Residuals
s [h^{-1} Mpc]