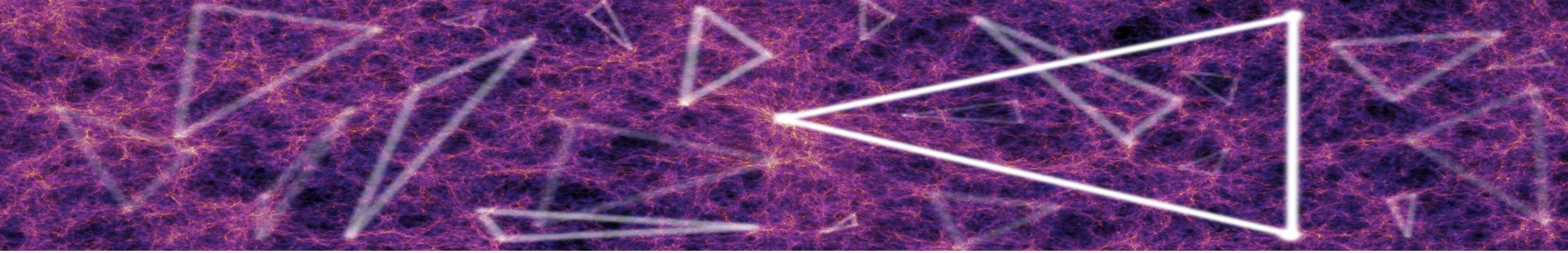


Towards a full modelling of the 3PCF at the BAO scales

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

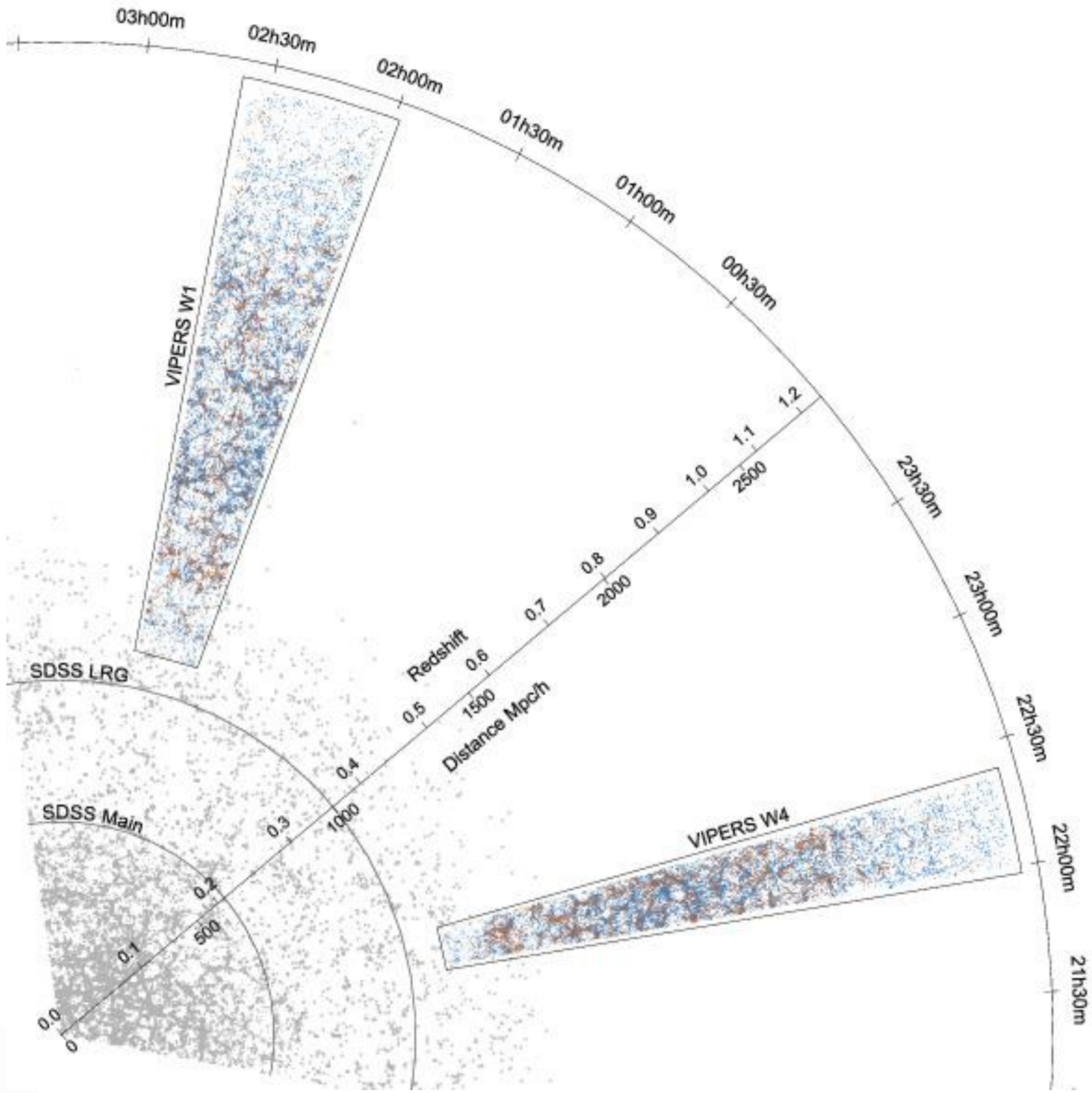
Michele Moresco

Dipartimento di Fisica e Astronomia “Augusto Righi”
Alma Mater Studiorum - Università di Bologna
michele.moresco@unibo.it

in collaboration with

N. Principi, K. Nagainis, M. Guidi (UniBo),
A. Veropalumbo, A. Farina (INAF OABr, INFN-Genova),
B. Granett, I Risso (INAF OABr)

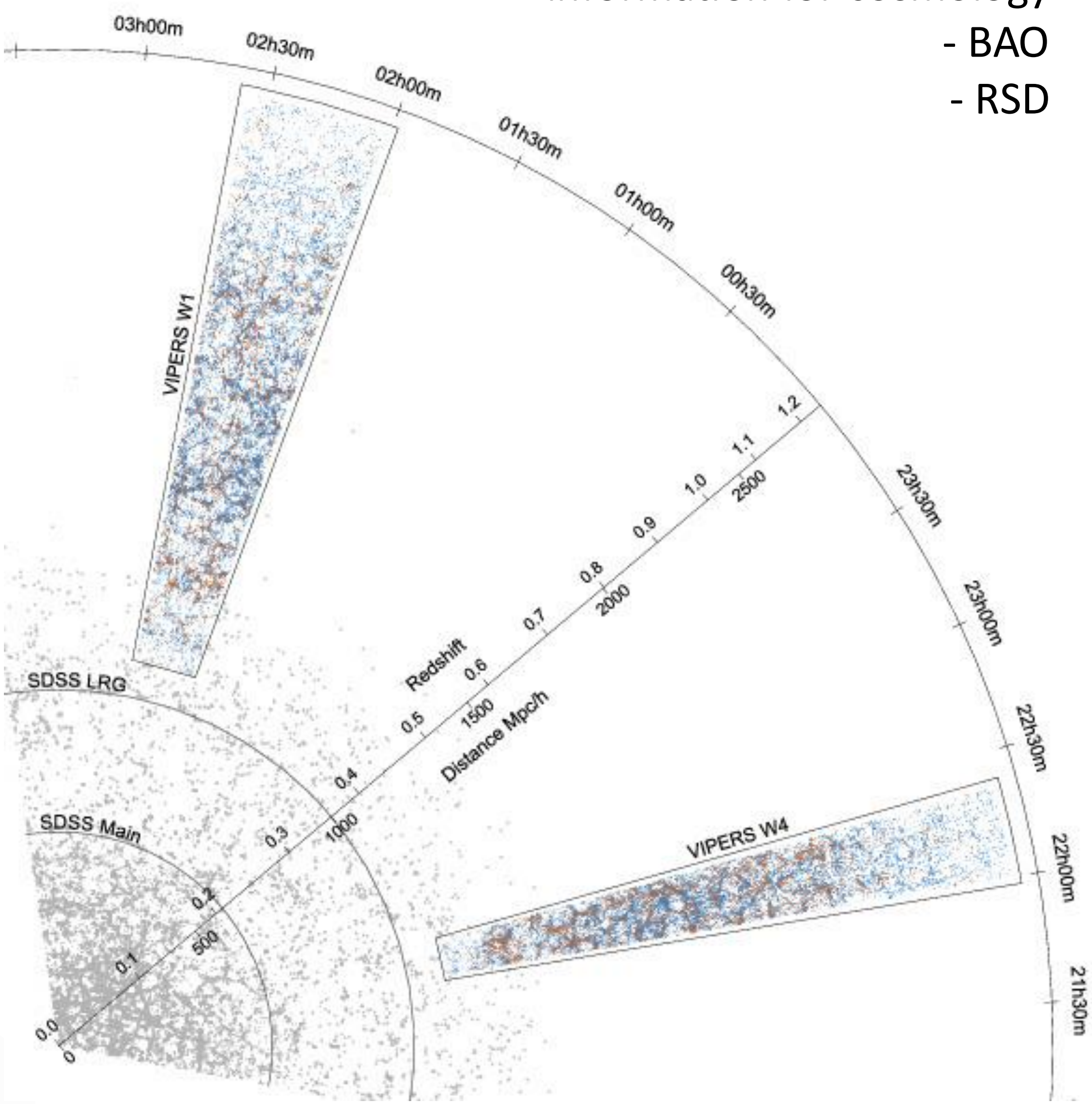
Why going to higher orders



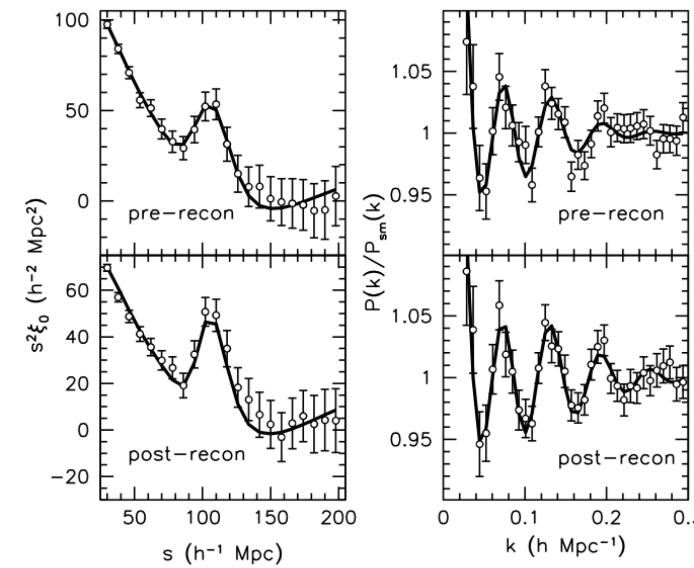
Guzzo et al. (2018)

Why going to higher orders

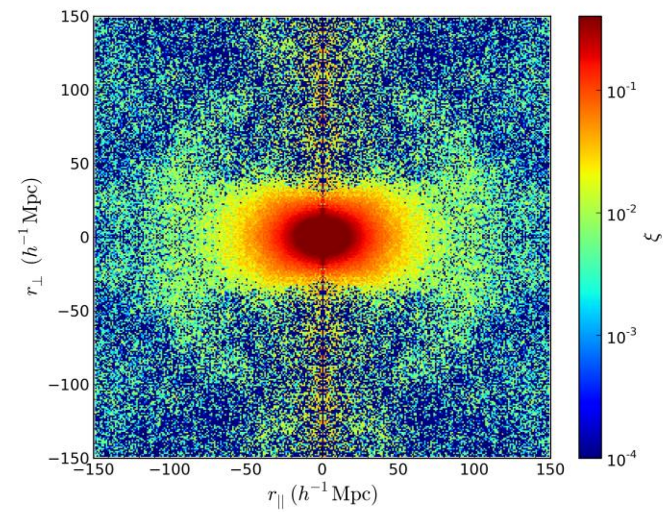
Galaxy correlation functions encode fundamental information for cosmology
- BAO
- RSD



Guzzo et al. (2018)



Anderson et al. (2014)



Samushia et al. (2013)

Why going to higher orders

For a Gaussian Random Field, 2PCF (and/or Pk) would be enough (mean and variance)

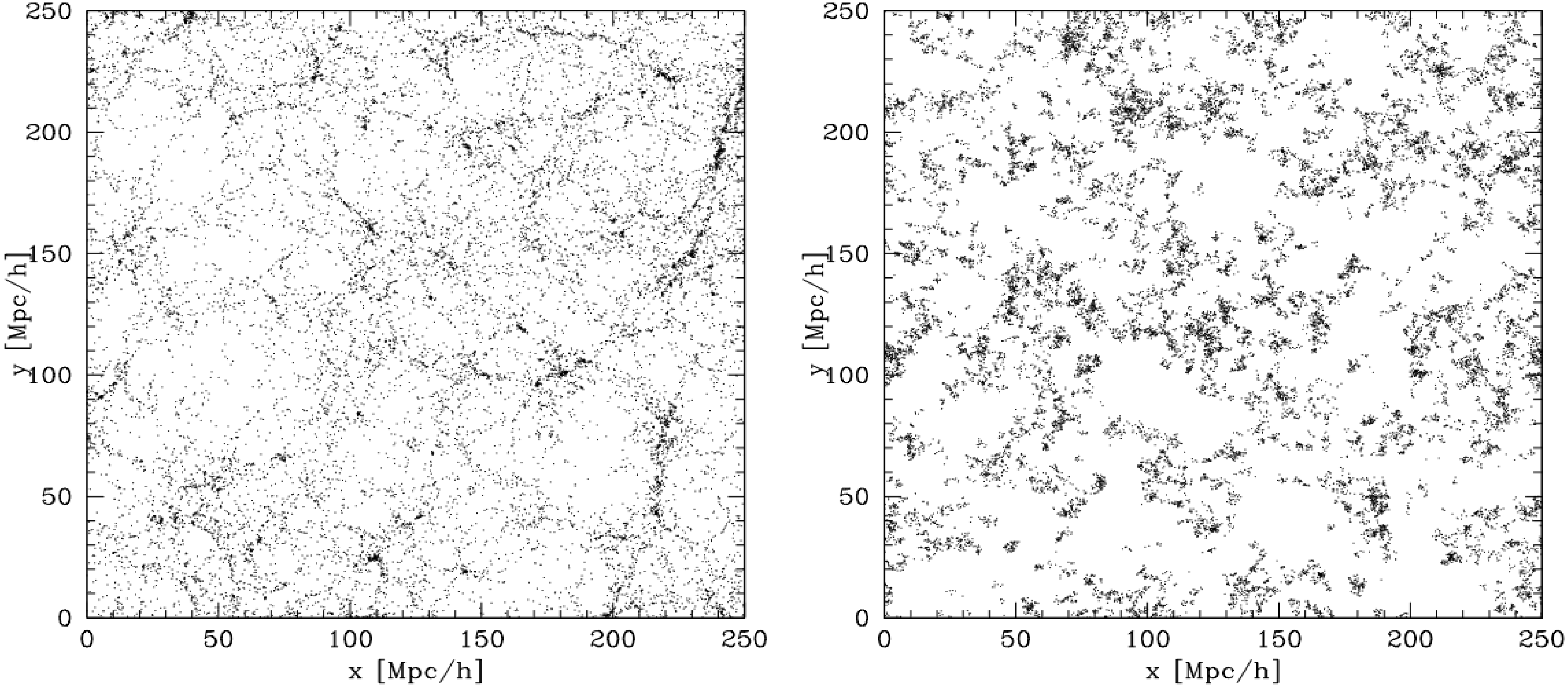
… but the Universe is just not like that!

Why going to higher orders

For a Gaussian Random Field, 2PCF (and/or Pk) would be enough (mean and variance)

… but the Universe is just not like that!

Sefusatti & Scoccimarro (2005)



same 2-pt statistic,
difference in the higher orders

The gain and the price of going to higher orders

Including higher-orders in the analysis is **fundamental for current and next-generation surveys** (e.g. Euclid, DESI), because it improves the cosmological constraints, breaks degenracies, and provides additional information, **but**...

The gain and the price of going to higher orders

Including higher-orders in the analysis is **fundamental for current and next-generation surveys** (e.g. Euclid, DESI), because it improves the cosmological constraints, breaks degenracies, and provides additional information, **but**...

both measuring and modelling it has a **high computational (and theoretical) cost**

The gain and the price of going to higher orders

Including higher-orders in the analysis is **fundamental for current and next-generation surveys** (e.g. Euclid, DESI), because it improves the cosmological constraints, breaks degenracies, and provides additional information, **but**...

both measuring and modelling it has a **high computational (and theoretical) cost**

Currently, several theoretical and analysis efforts are ongoing on this aspect. For the future, it will be crucial to:

- Be able to provide theoretical models more complete and significanly quicker
- Carefully assess the systematics involved and how to treat with those
- Provide new tools to exploit 3PCF and the combination of 2PCF+3PCF for current future surveys

The gain and the price of going to higher orders

Including higher-orders in the analysis is **fundamental for current and next-generation surveys** (e.g. Euclid, DESI), because it improves the cosmological constraints, breaks degenracies, and provides additional information, **but**...

both measuring and modelling it has a **high computational (and theoretical) cost**

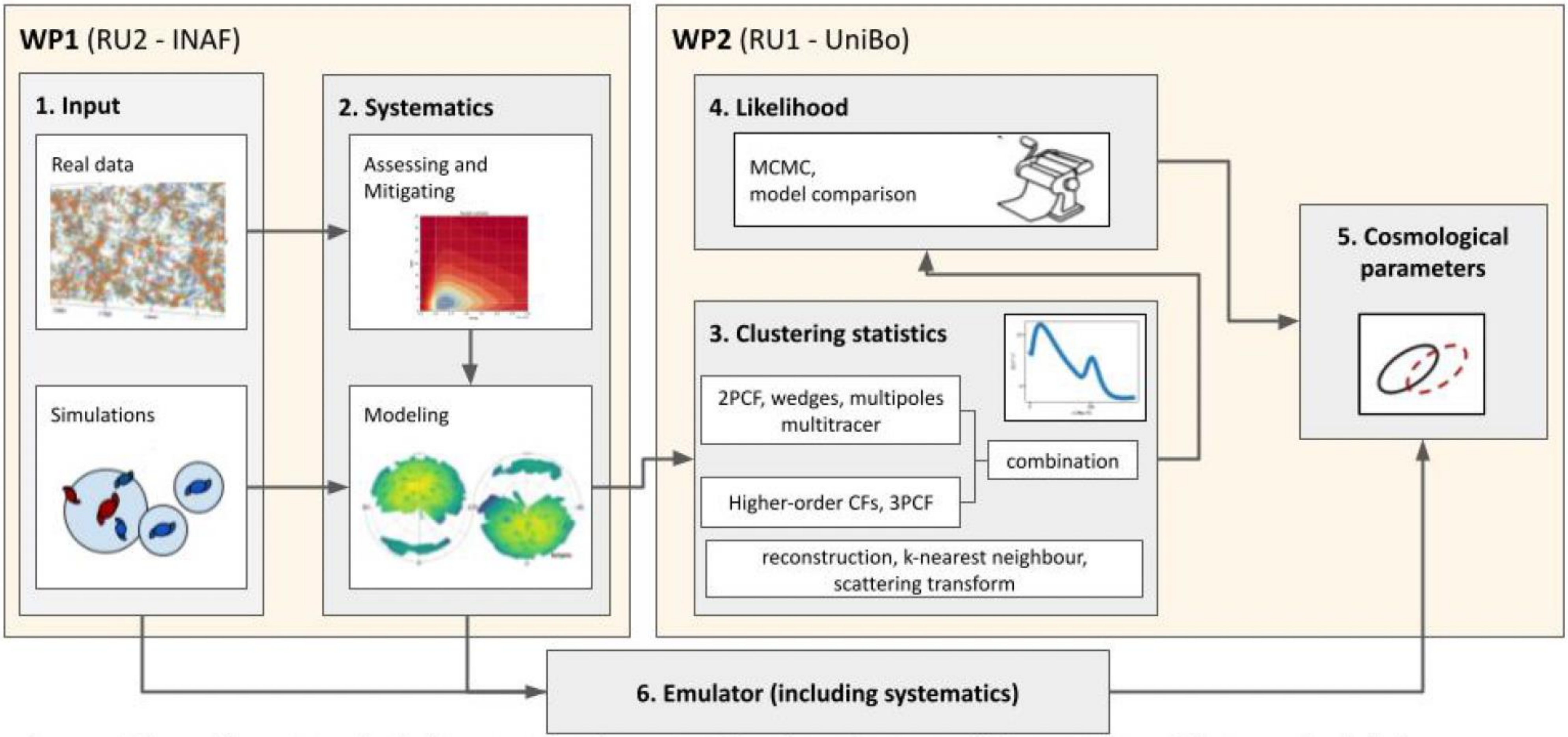
Currently, several theoretical and analysis efforts are ongoing on this aspect. For the future, it will be crucial to:

- Be able to provide theoretical models more complete and significanly quicker
- Carefully assess the systematics involved and how to treat with those
- Provide new tools to exploit 3PCF and the combination of 2PCF+3PCF for current future surveys

Today, we report on:

- Study the **impact of interlopers** on the 3PCF (Master Thesis of **Nicola Principi**)
- Develop **emulator for the anisotropic 3PCF**, and **forecasts on the accuracy** of cosmological parameters **from the 3PCF** (Master Thesis of **Kristers Nagainis**)
- Analysis and cosmological constraints from the 2PCF+3PCF combination (see **M. Guidi**'s talk)

The gain and the price of going to higher orders

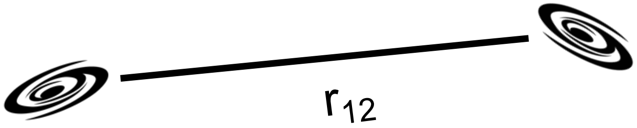


PRIN 2022 "Optimizing the extraction of cosmological information from Large Scale Structure analysis in view of the next large spectroscopic surveys" (2022NY2ZRS 001)

Correlation functions

Probability of finding **pairs** and triplets of objects:

$$dP = n^2[1 + \xi(r)]dV_1 dV_2$$



2PCF estimator
Landy & Szalay (1993)

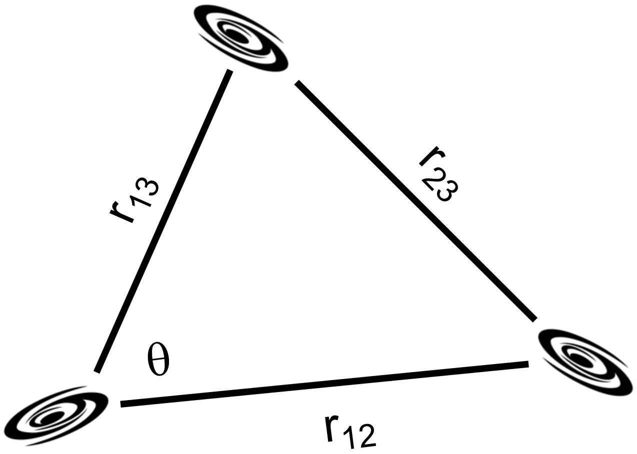
$$\xi(r) = \frac{DD - 2DR - RR}{RR}$$

Correlation functions

Probability of finding **pairs** and **triplets** of objects:

$$dP = n^2[1+\xi(r)]dV_1 dV_2$$

$$dP = n^3[1+\xi(r_{12})+\xi(r_{13})+\xi(r_{23})+\zeta(r_{12},r_{13},r_{23})]\, dV_1 dV_2\, dV_3$$



2PCF estimator
Landy & Szalay (1993)

$$\xi(r) = \frac{DD - 2DR - RR}{RR}$$

3PCF estimator
Szapudi & Szalay (1998)

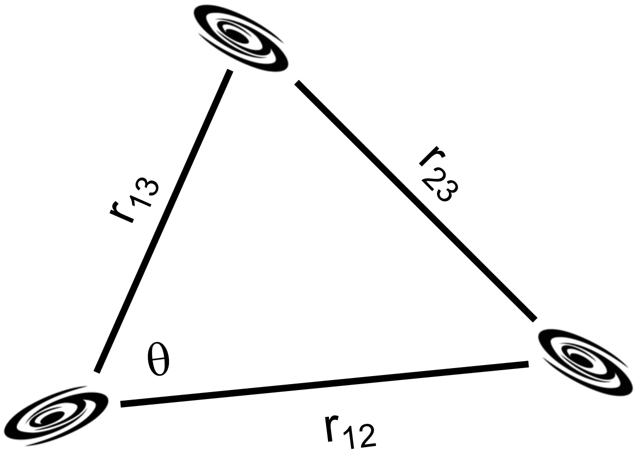
$$\zeta(r_{12},r_{13},r_{23}) = \frac{DDD - 3DDR + 3DRR - RRR}{RRR}$$

Correlation functions

Probability of finding **pairs** and **triplets** of objects:

$$dP = n^2[1 + \xi(r)]dV_1 dV_2$$

$$dP = n^3[1 + \xi(r_{12}) + \xi(r_{13}) + \xi(r_{23}) + \zeta(r_{12}, r_{13}, r_{23})] \\ dV_1 dV_2\, dV_3$$



2PCF estimator
Landy & Szalay (1993)

$$\xi(r) = \frac{DD - 2DR - RR}{RR}$$

3PCF estimator
Szapudi & Szalay (1998)

$$\zeta(r_{12}, r_{13}, r_{23}) = \frac{DDD - 3DDR + 3DRR - RRR}{RRR}$$

connected 3PCF $\quad \zeta(r_{12}, r_{13}, \theta) \quad \propto b^3 \sigma_8^4$

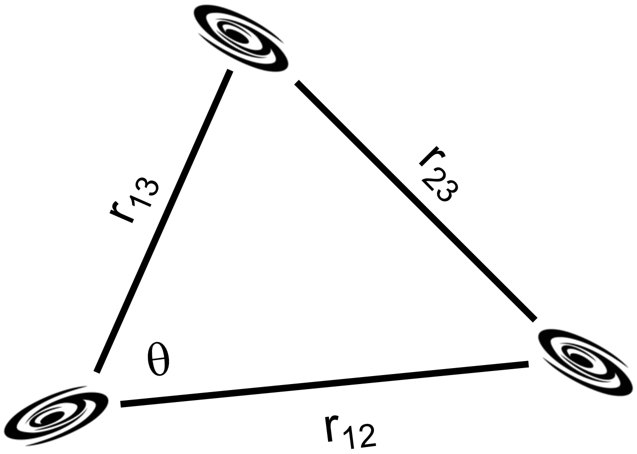
reduced 3PCF $\quad Q(r_{12}, r_{13}, \theta) = \dfrac{\zeta(r_{12}, r_{13}, \theta)}{\xi(r_{12})\xi(r_{23}) + \xi(r_{23})\xi(r_{31}) + \xi(r_{31})\xi(r_{12})} \quad \propto b^{-1}$

Correlation functions

Probability of finding **pairs** and **triplets** of objects:

$$dP = n^2[1 + \xi(r)]dV_1 dV_2$$

$$dP = n^3[1 + \xi(r_{12}) + \xi(r_{13}) + \xi(r_{23}) + \zeta(r_{12}, r_{13}, r_{23})] \\ dV_1 dV_2\, dV_3$$



2PCF estimator
Landy & Szalay (1993)

$$\xi(r) = \frac{DD - 2DR - RR}{RR}$$

3PCF estimator
Szapudi & Szalay (1998)

$$\zeta(r_{12}, r_{13}, r_{23}) = \frac{DDD - 3DDR + 3DRR - RRR}{RRR}$$

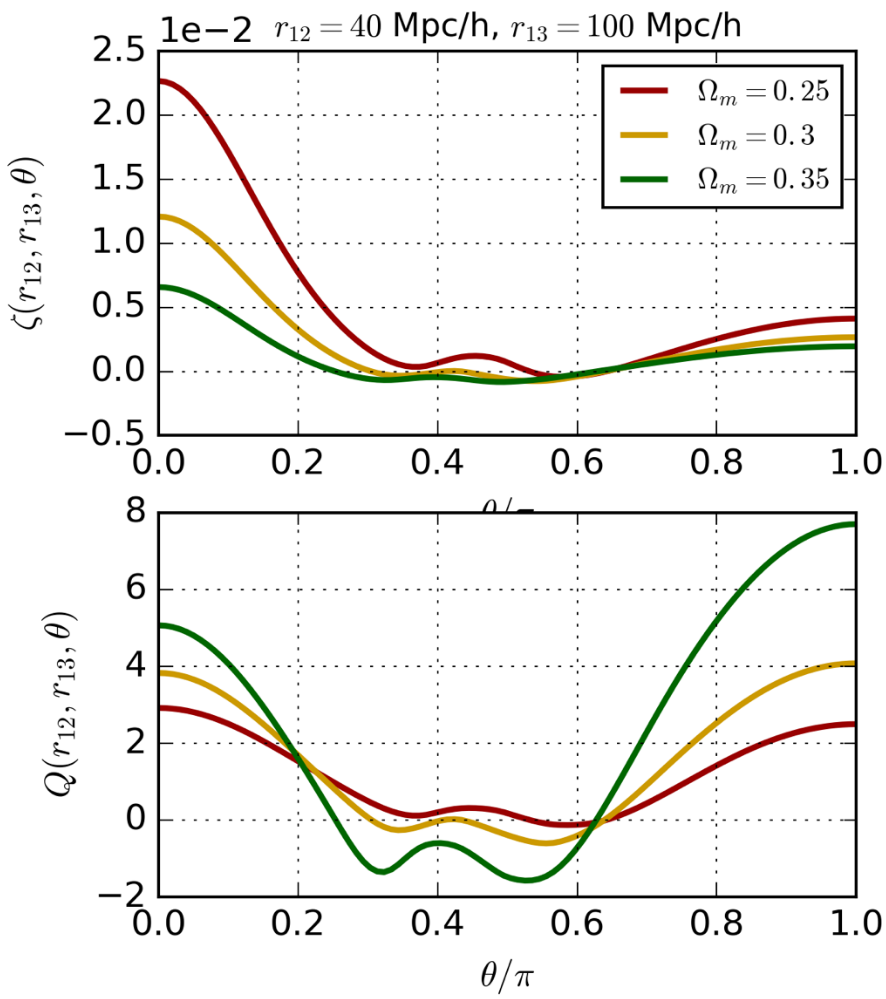
connected 3PCF $\qquad \zeta(r_{12}, r_{13}, \theta) \qquad \propto b^3 \sigma_8^4$

reduced 3PCF $\qquad Q(r_{12}, r_{13}, \theta) = \dfrac{\zeta(r_{12}, r_{13}, \theta)}{\xi(r_{12})\xi(r_{23}) + \xi(r_{23})\xi(r_{31}) + \xi(r_{31})\xi(r_{12})} \qquad \propto b^{-1}$

Theoretical recent advances happening right now! Fast spherical harmonics decomposition estimator (Slepian et al, 2015), anisotropic z-space 3PCF (Slepian et al. 2018), estimate of different 3PCF covariance matrix (Veropalumbo et al. 2022), …

The BAO peak in the 3PCF

Moresco et al. (2021)

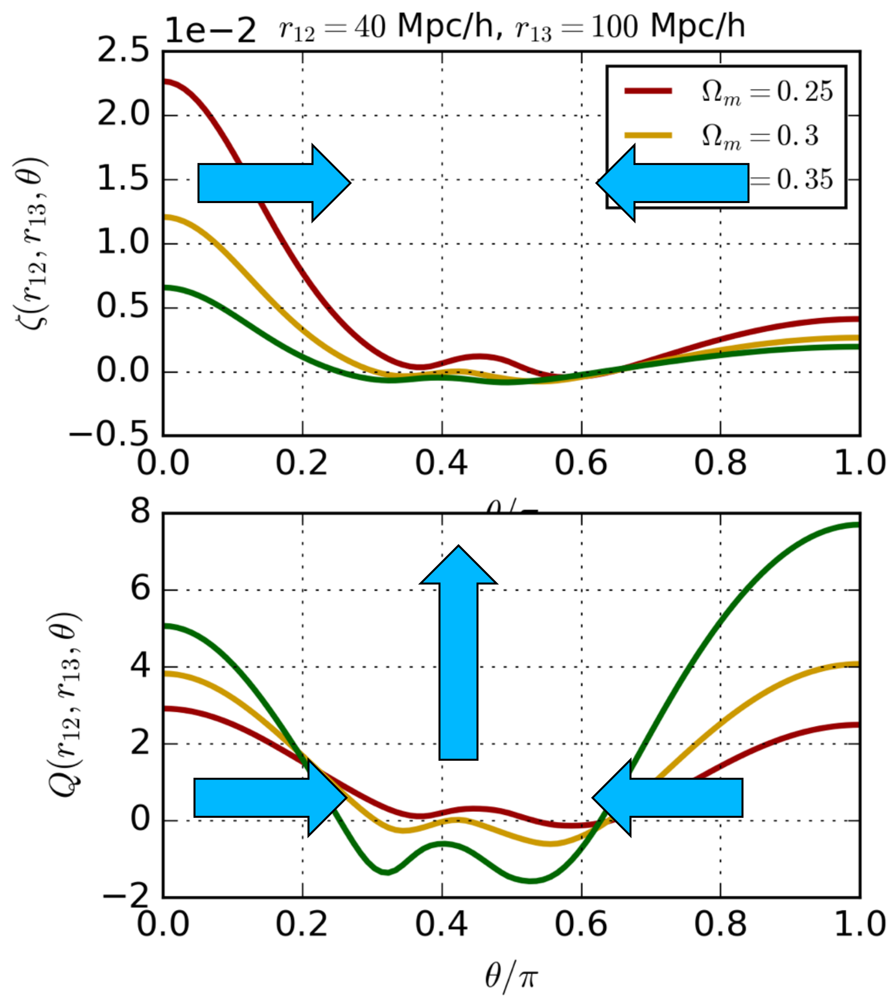


The BAO peak in the 3PCF

Moresco et al. (2021)

BAO signal squeezed

lower number of bins (SNR)
higher contrast



$60 < r_{23}$ [Mpc/h] < 140

The BAO peak in the 3PCF

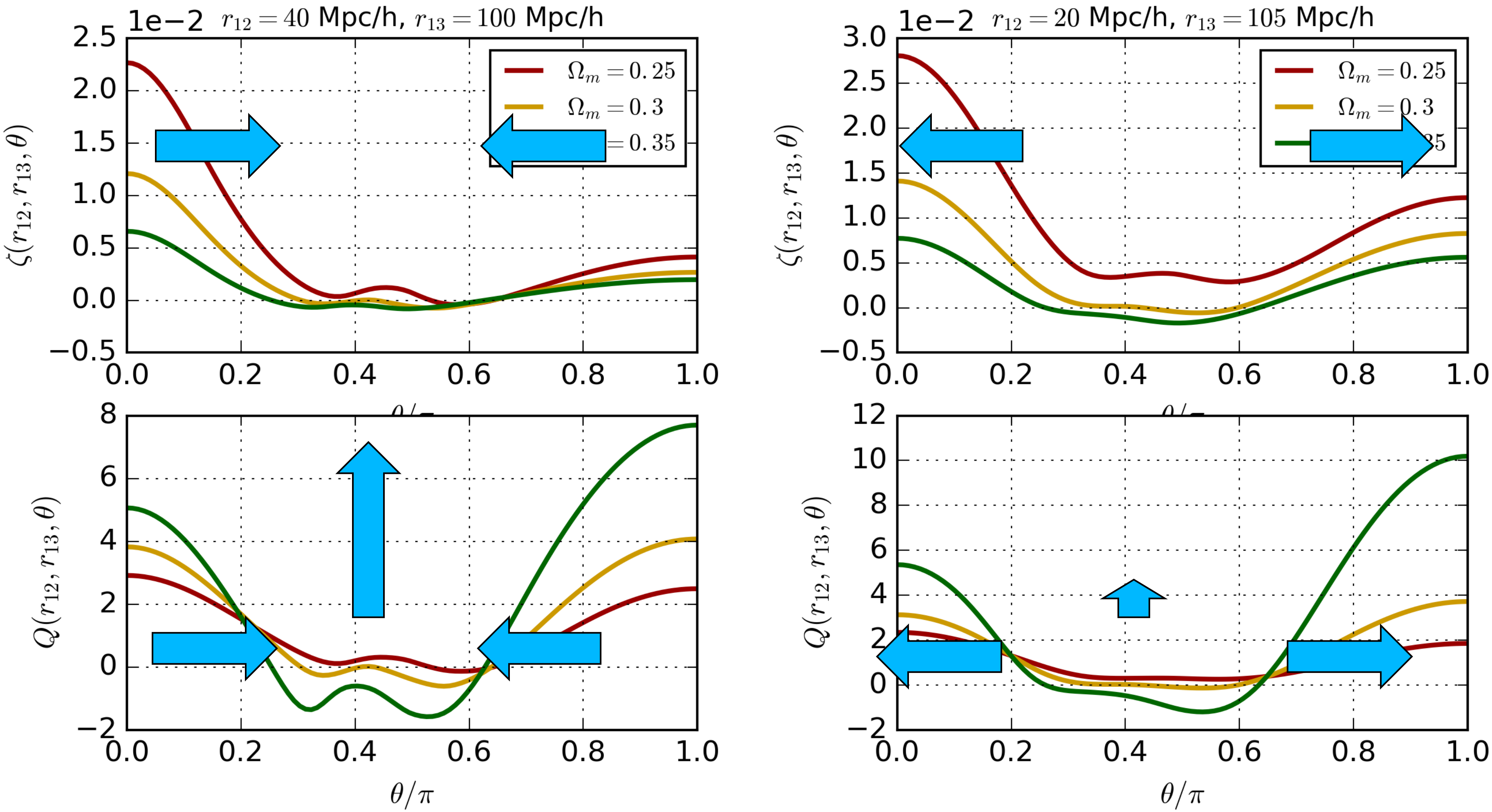
Moresco et al. (2021)

BAO signal squeezed

lower number of bins (SNR)
higher contrast

BAO signal diluted

higher number of bins (SNR)
lower contrast



60 < r_{23} [Mpc/h] < 140

85 < r_{23} [Mpc/h] < 125

Impact on interlopers on the 3PCF

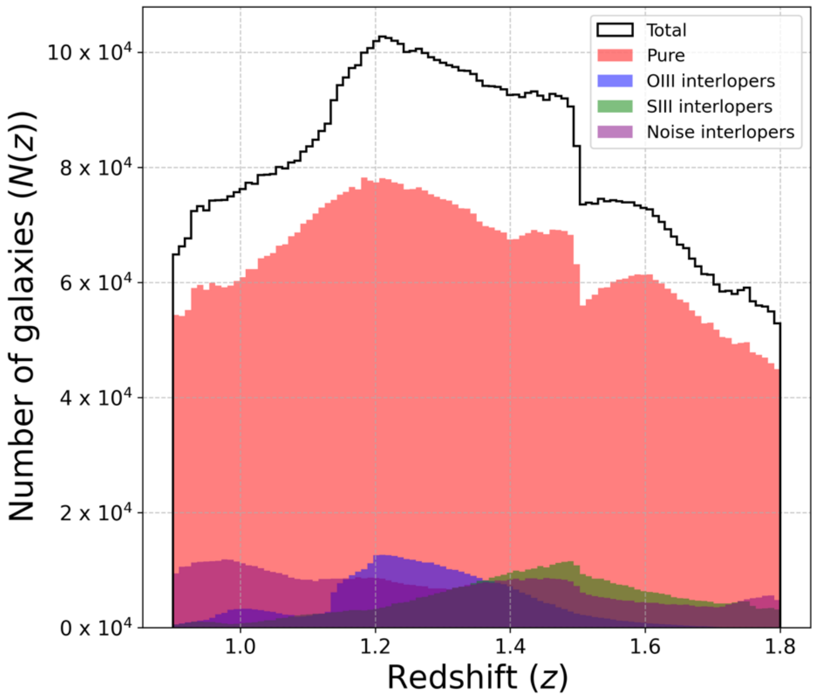
Effect of interlopers in the 3PCF analysis

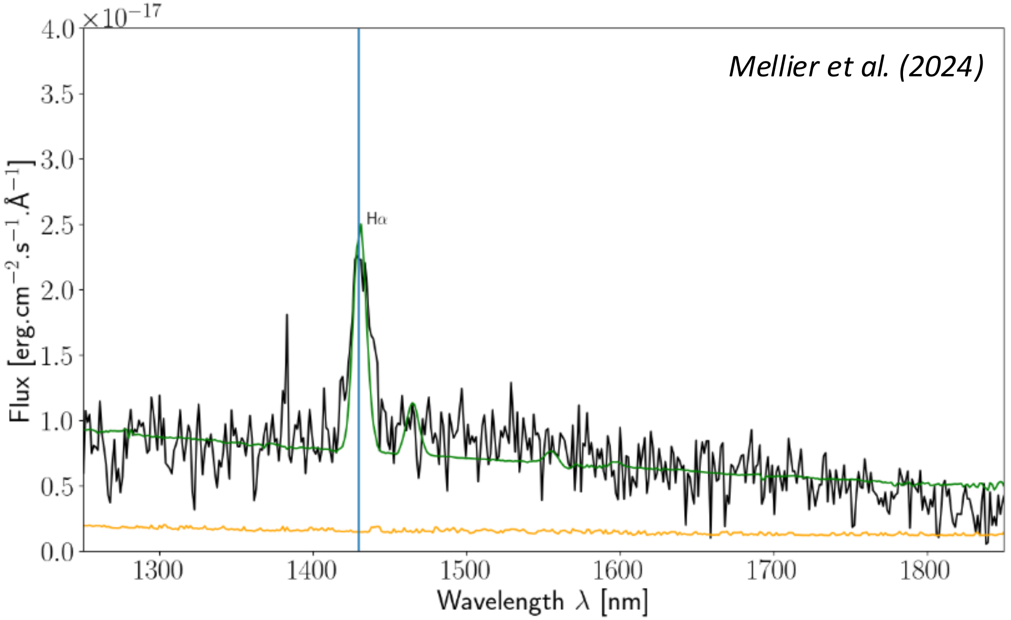
Work by Nicola Principi

Aim: Assess the effect of interlopers on the 3PCF measurements, studying their impact on the derived parameters

Sample: application to the Euclid surveys, considering 2 different catalogs:

- Flagship 2: full octant (+ interlopers) ⇒ realistic catalog (Poissonian errors)
- Euclid Large Mocks: 1000 30deg-field catalogues from PINOCCHIO (+ interlopers) ⇒ covariance





Different redshift intervals, different contaminations

New cross 3PCF estimators

Work by Nicola Principi

$$\hat{\zeta}_{tot} = \sum_i f_i^3 \frac{R_i R_i R_i}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_i + 3 \left[\sum_{i \neq j, i<j} \left(f_i^2 f_j \frac{R_i R_i R_j}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_{iij} + f_i f_j^2 \frac{R_i R_j R_j}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_{ijj} \right) \right]$$
$$+ 6 \sum_{i \neq j \neq k, i<j<k} f_i f_j f_k \frac{R_i R_j R_k}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_{ijk}.$$

New cross 3PCF estimators

Work by Nicola Principi

$$\hat{\zeta}_{tot} = \sum_i f_i^3 \frac{R_i R_i R_i}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_i + 3 \left[\sum_{i \neq j, i<j} \left(f_i^2 f_j \frac{R_i R_i R_j}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_{iij} + f_i f_j^2 \frac{R_i R_j R_j}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_{ijj} \right) \right]$$

$$+ 6 \sum_{i \neq j \neq k, i<j<k} f_i f_j f_k \frac{R_i R_j R_k}{R_{tot} R_{tot} R_{tot}} \hat{\zeta}_{ijk}$$

- f_i^3: fraction of i-th sample
- $\hat{\zeta}_i$: autocorrelation 3PCF
- $\hat{\zeta}_{iij}$: bicross 3PCF
- $\hat{\zeta}_{ijk}$: tricross 3PCF

New cross 3PCF estimators

Work by Nicola Principi

fraction of i-th sample

autocorrelation 3PCF

$$\hat{\zeta}_{tot} = \sum_i f_i^3 \frac{R_iR_iR_i}{R_{tot}R_{tot}R_{tot}} \hat{\zeta}_i + 3\left[\sum_{i\neq j, i<j} \left(f_i^2 f_j \frac{R_iR_iR_j}{R_{tot}R_{tot}R_{tot}} \hat{\zeta}_{iij} + f_i f_j^2 \frac{R_iR_jR_j}{R_{tot}R_{tot}R_{tot}} \hat{\zeta}_{ijj} \right)\right]$$

$$+ 6 \sum_{i\neq j\neq k, i<j<k} f_i f_j f_k \frac{R_iR_jR_k}{R_{tot}R_{tot}R_{tot}} \hat{\zeta}_{ijk}$$

bicross 3PCF

tricross 3PCF

$$\hat{\zeta}_i = \frac{D_iD_iD_i - 3D_iD_iR_i + 3D_iR_iR_i - R_iR_iR_i}{R_iR_iR_i},$$

$$\hat{\zeta}_{iij} = \frac{D_iD_iD_j - D_iD_iR_j - 2D_iR_iD_j + 2D_iR_iR_j + R_iR_iD_j - R_iR_iR_j}{R_iR_iR_j},$$

$$\hat{\zeta}_{ijj} = \frac{D_iD_jD_j + D_iR_jR_j - 2D_iD_jR_j + 2R_iD_jR_j - R_iD_jD_j - R_iR_jR_j}{R_iR_jR_j},$$

$$\hat{\zeta}_{ijk} = \frac{D_iD_jD_k - D_iR_jD_k - D_iD_jR_k - R_iD_jD_k + R_iR_jD_k + R_iD_jR_k + D_iR_jR_k - R_iR_jR_k}{R_iR_jR_k}$$

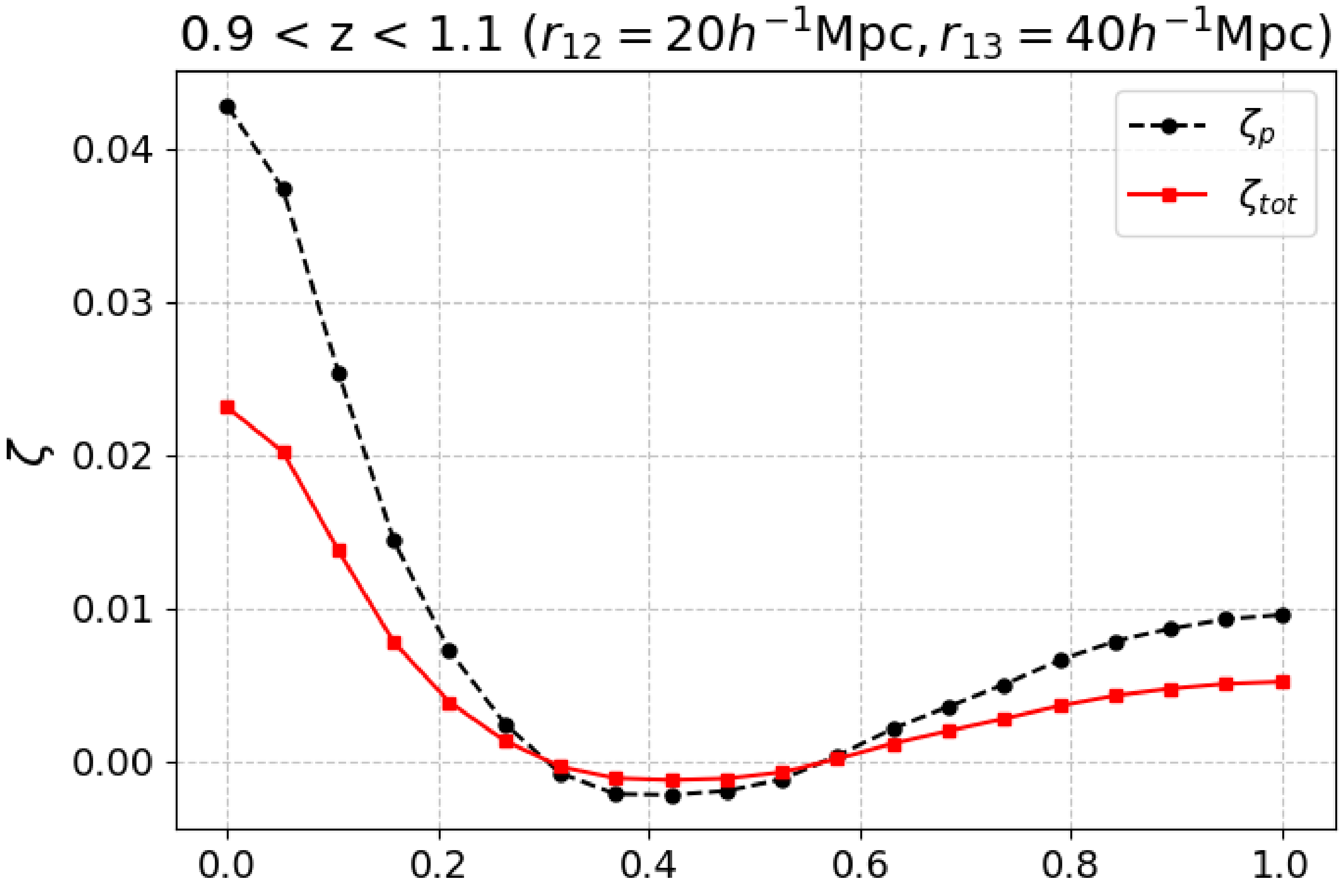
4 new classes and 19 new functions implemented in the CosmoBolognaLib suite

Autocorrelations and crosscorrelations

Work by Nicola Principi

Analysis on Flagship 2 catalog

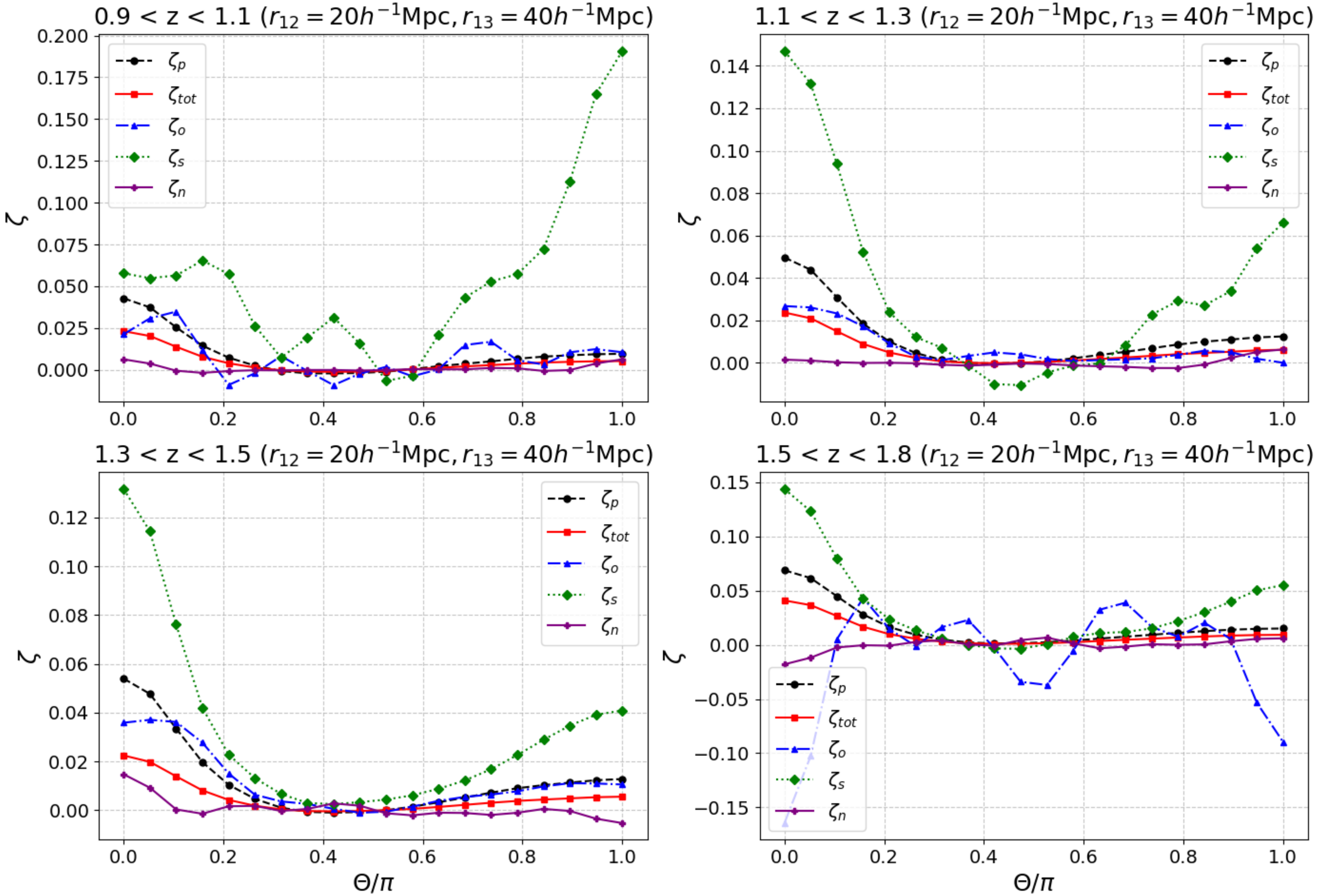
Global effect: **damping of the signal**



Autocorrelations and crosscorrelations

Work by Nicola Principi

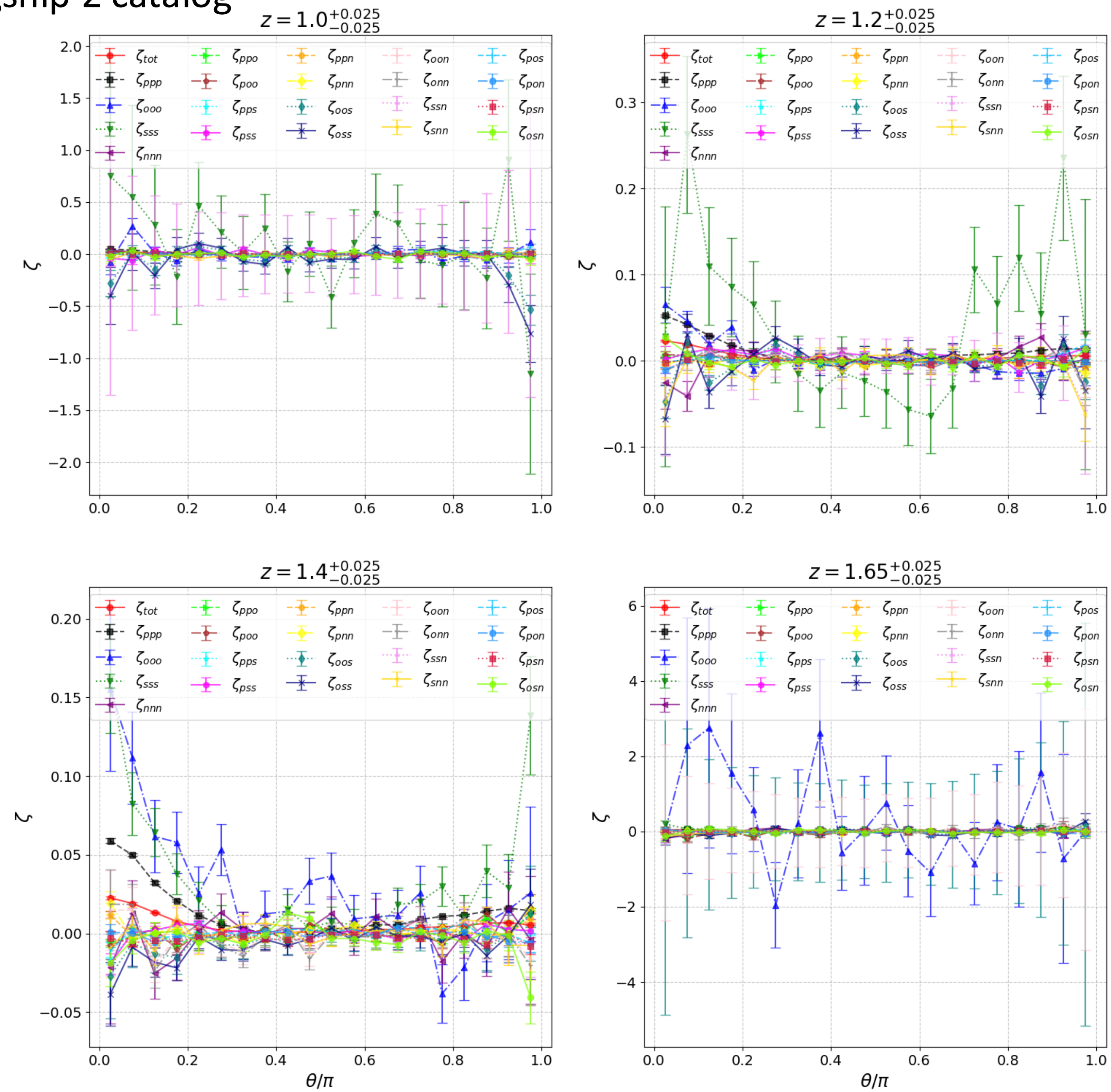
Analysis on Flagship 2 catalog



Autocorrelations and crosscorrelations

Work by Nicola Principi

Analysis on Flagship 2 catalog

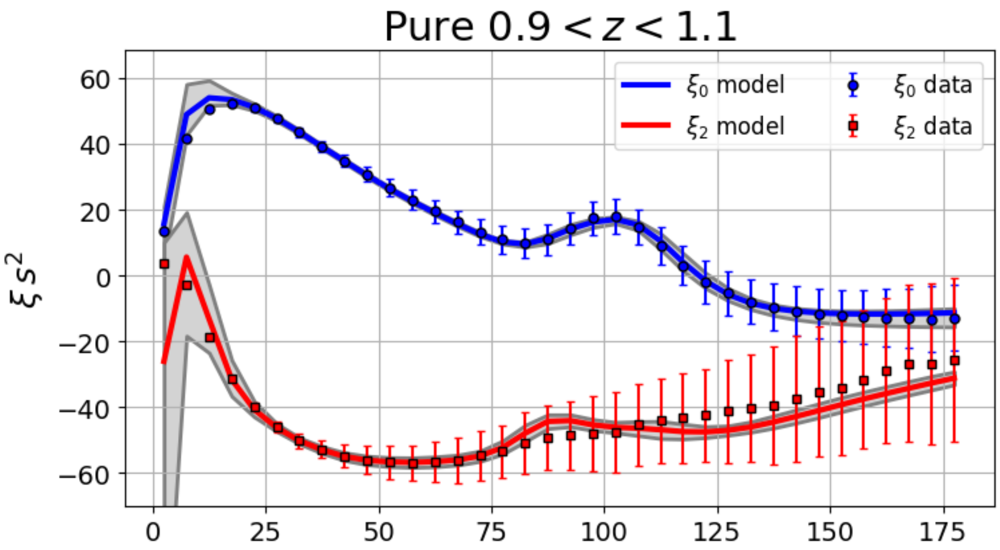


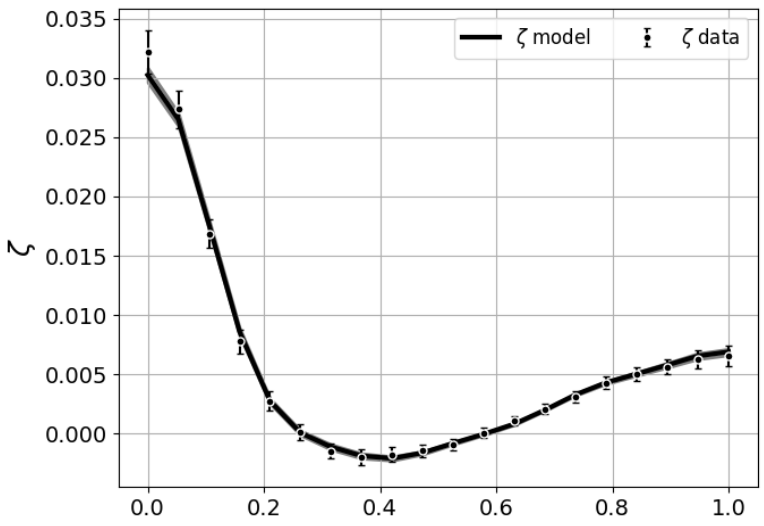
Measurements of the 2PCF and 3PCF

Work by Nicola Principi

Analysis on Euclid Large Mocks

Constraints on both 2PCF (+ multipoles) and 3PCF (isotropic) + covariance



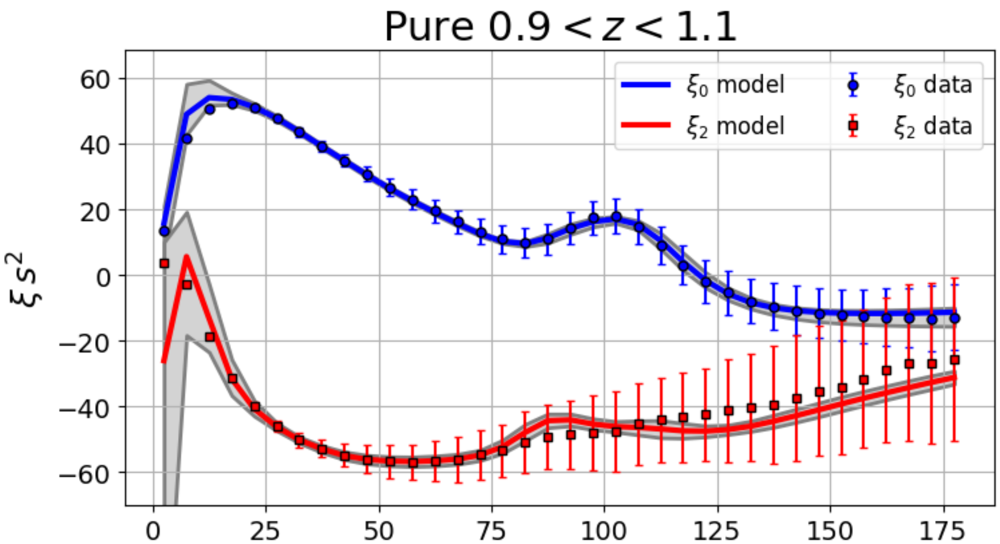


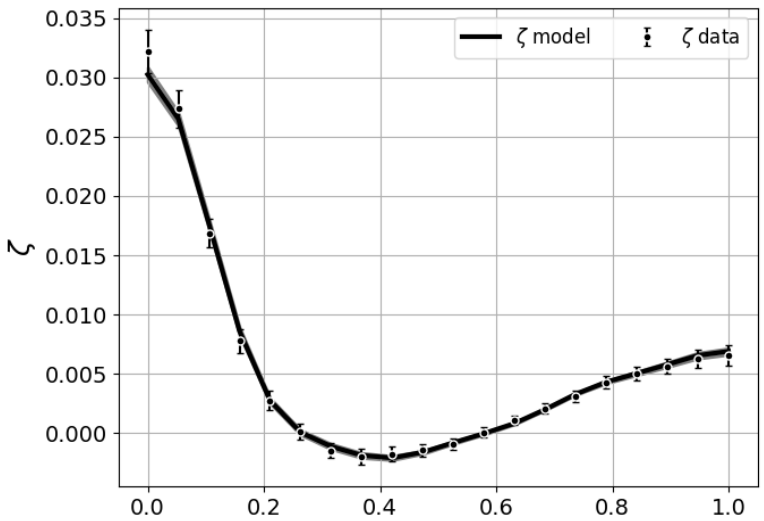
Measurements of the 2PCF and 3PCF

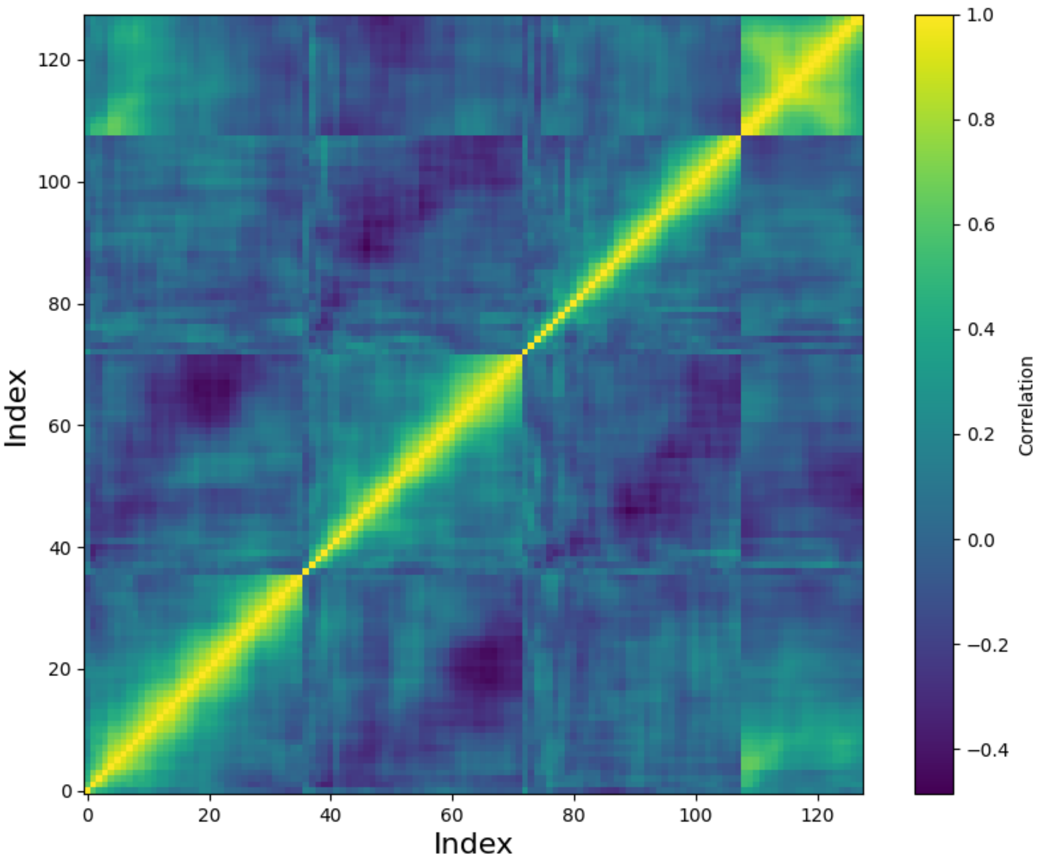
Work by Nicola Principi

Analysis on Euclid Large Mocks

Constraints on both 2PCF (+ multipoles) and 3PCF (isotropic) + covariance





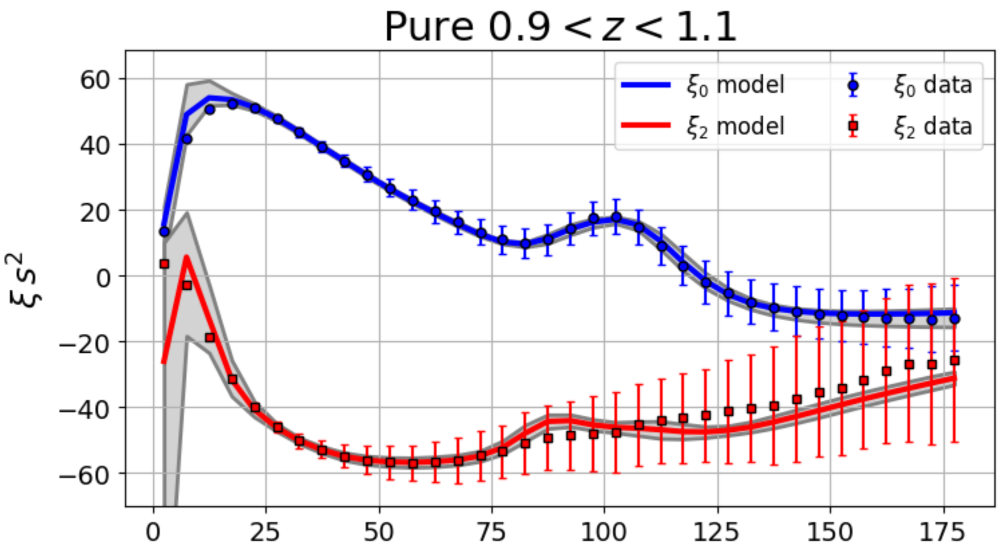


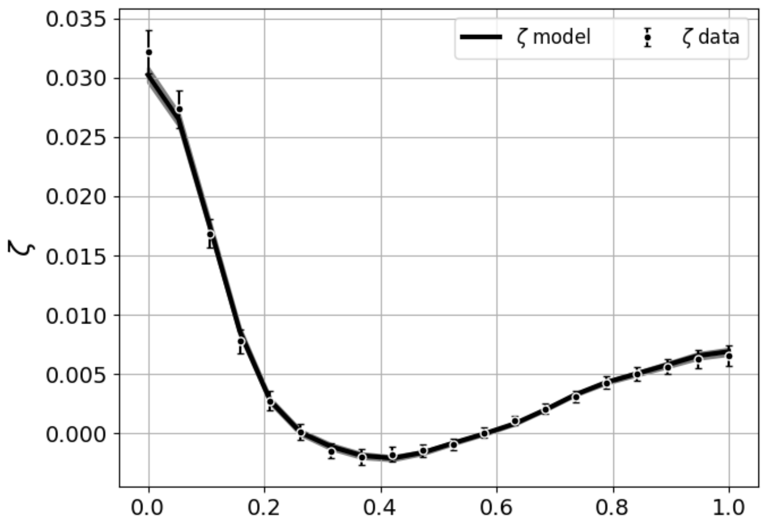
Measurements of the 2PCF and 3PCF

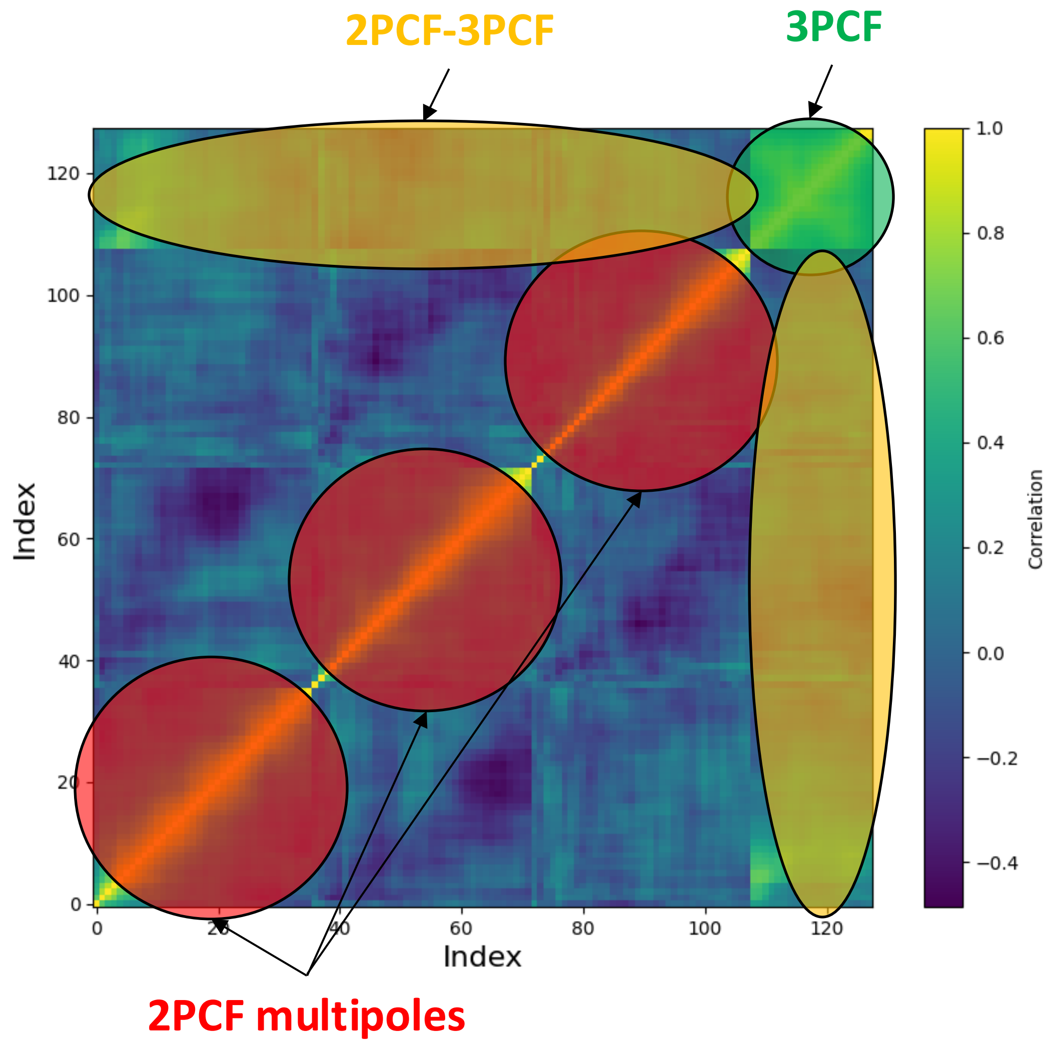
Work by Nicola Principi

Analysis on Euclid Large Mocks

Constraints on both 2PCF (+ multipoles) and 3PCF (isotropic) + covariance



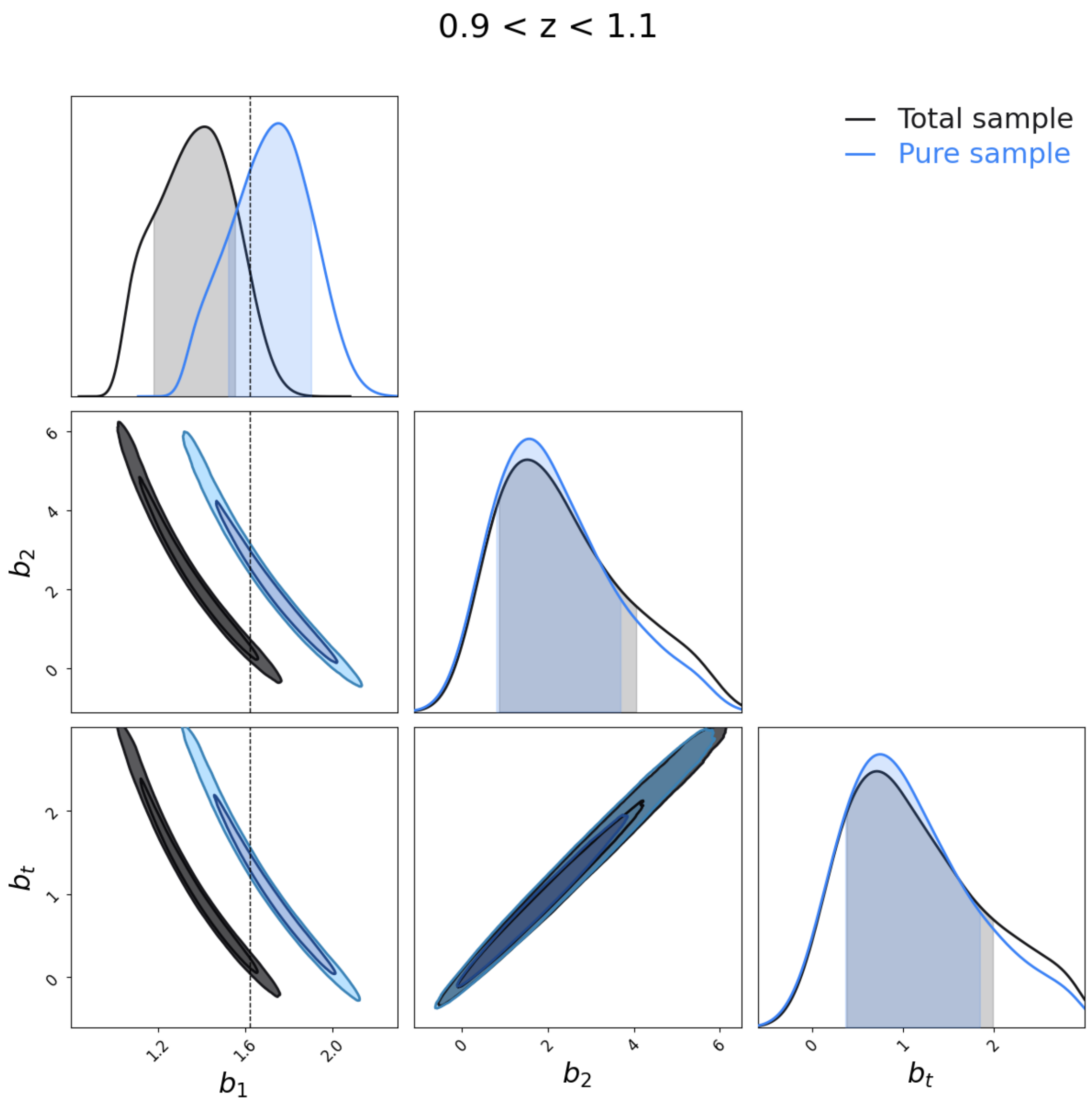




Constraints from the 3PCF

Work by Nicola Principi

Analysis on Euclid Large Mocks



Offset in b_1, but not in b_2 and b_t

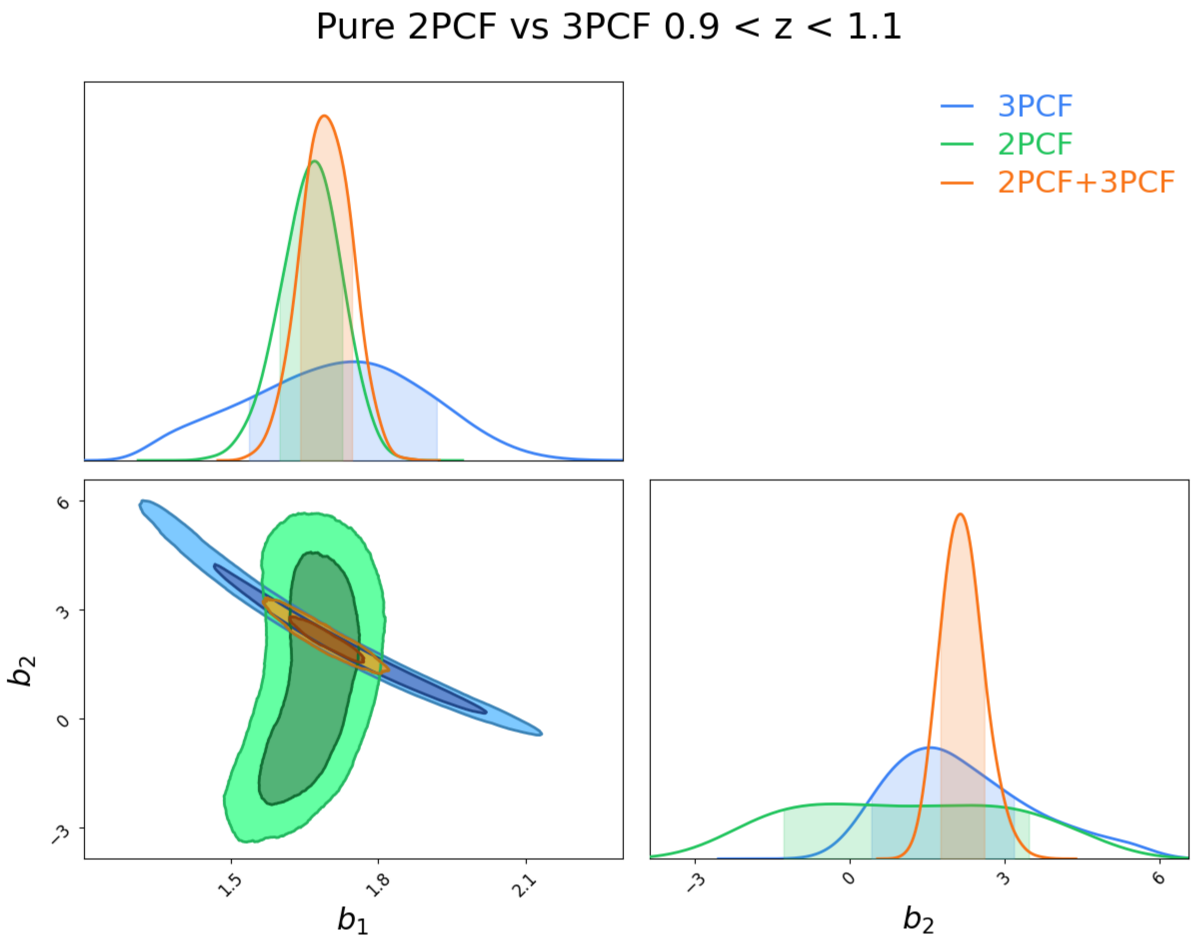
Could be due to a simple offset between contaminated and pure 3PCF

(under investigation...)

Combining 2PCF and 3PCF

Work by Nicola Principi

Analysis on Euclid Large Mocks



Combining 2PCF and 3PCF significantly improves the constraints (see also M. Guidi's talk)

Emulating the anispotropic matter 3PCF

(and what about BAO?)

Modelling the anisotropic 3PCF at BAO scales

Work by Kristers Nagainis

Aim: Develop an emulator for the anisotropic 3PCF of matter, to significantly speed up its computation and be able to provide forecasts on the accuracy of the constraints on cosmological parameters

ISOTROPIC

$$\zeta_{\rm s}(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2) = \sum_\ell \zeta_\ell(r_1, r_2) P_\ell(\hat{r}_1 \cdot \hat{r}_2)$$

$$\zeta_\ell(r_1, r_2) =$$

$$(-1)^\ell \int \frac{k_1^2 k_2^2 dk_1 dk_2}{(2\pi^2)^2} B_{{\rm s},\ell}(k_1, k_2) j_\ell(k_1 r_1) j_\ell(k_2 r_2),$$

Slepian et al. (2017)

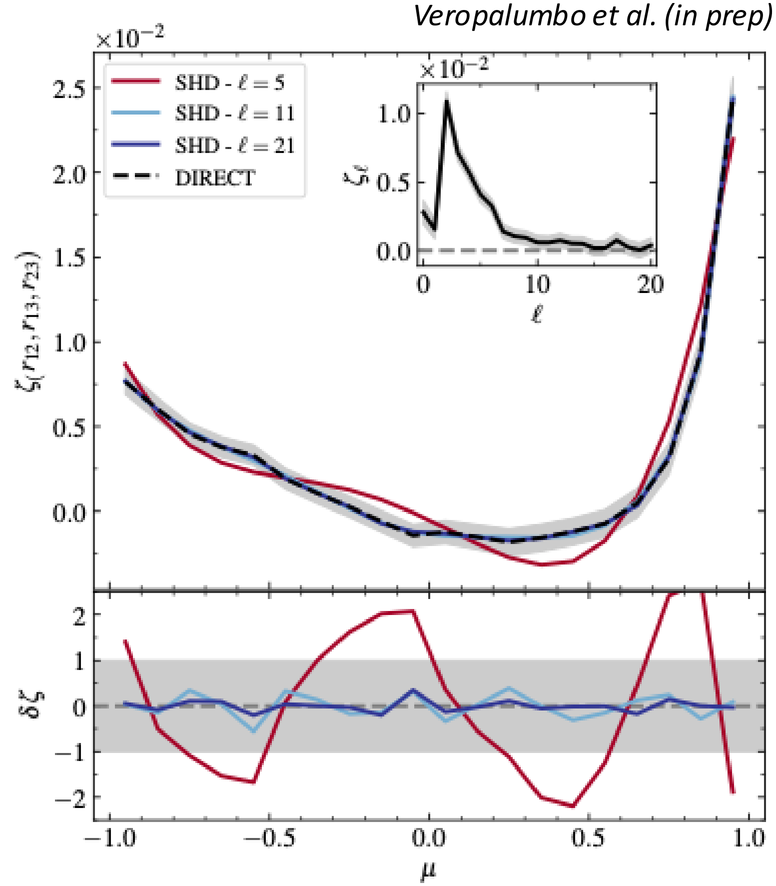
From 3PCF to 3PCF multipoles

Modelling the anisotropic 3PCF at BAO scales

Work by Kristers Nagainis

Aim: Develop an emulator for the anisotropic 3PCF of matter, to significantly speed up its computation and be able to provide forecasts on the accuracy of the constraints on cosmological parameters

Veropalumbo et al. (in prep)



ISOTROPIC

$$\zeta_{\rm s}(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2) = \sum_\ell \zeta_\ell(r_1, r_2) P_\ell(\hat{r}_1 \cdot \hat{r}_2)$$

$$\zeta_\ell(r_1, r_2) =$$

$$(-1)^\ell \int \frac{k_1^2 k_2^2 dk_1 dk_2}{(2\pi^2)^2} B_{{\rm s},\ell}(k_1, k_2) j_\ell(k_1 r_1) j_\ell(k_2 r_2),$$

Slepian et al. (2017)

From 3PCF to 3PCF multipoles

Modelling the anisotropic 3PCF at BAO scales

Work by Kristers Nagainis

Aim: Develop an emulator for the anisotropic 3PCF of matter, to significantly speed up its computation and be able to provide forecasts on the accuracy of the constraints on cosmological parameters

ISOTROPIC

$$\zeta_s(r_1, r_2; \hat{r}_1 \cdot \hat{r}_2) = \sum_\ell \zeta_\ell(r_1, r_2) P_\ell(\hat{r}_1 \cdot \hat{r}_2)$$

$$\zeta_\ell(r_1, r_2) = (-1)^\ell \int \frac{k_1^2 k_2^2 dk_1 dk_2}{(2\pi^2)^2} B_{s,\ell}(k_1, k_2) j_\ell(k_1 r_1) j_\ell(k_2 r_2),$$

Slepian et al. (2017)

ANISOTROPIC

$$B_{\ell_1 \ell_2 L}(k_1, k_2) = (-i)^{\ell_1+\ell_2}(4\pi)^2 \int dr_1 r_1^2 \int dr_2 r_2^2 \times j_{\ell_1}(k_1 r_1) j_{\ell_2}(k_2 r_2) \zeta_{\ell_1 \ell_2 L}(r_1, r_2)$$

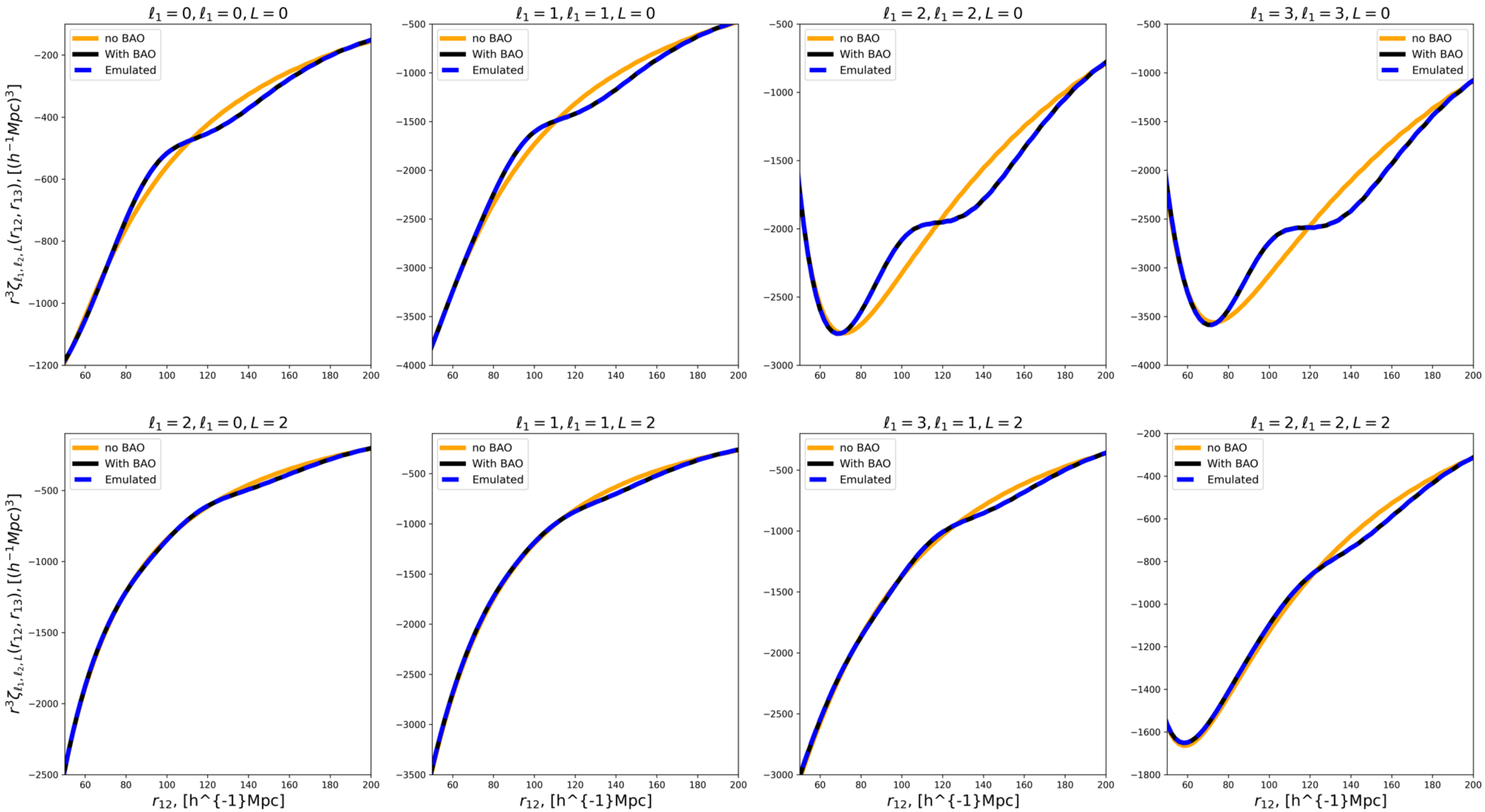
$$\zeta_{\ell_1 \ell_2 L}(r_1, r_2) = i^{\ell_1+\ell_2} \int \frac{dk_1 k_1^2}{2\pi^2} \int \frac{dk_2 k_2^2}{2\pi^2} \times j_{\ell_1}(r_1 k_1) j_{\ell_2}(r_2 k_2) B_{\ell_1 \ell_2 L}(k_1, k_2),$$

Sugiyama et al. (2019)

Modelling the anisotropic 3PCF at BAO scales

Work by Kristers Nagainis

Aim: Develop an emulator for the anisotropic 3PCF of matter, to significantly speed up its computation and be able to provide forecasts on the accuracy of the constraints on cosmological parameters

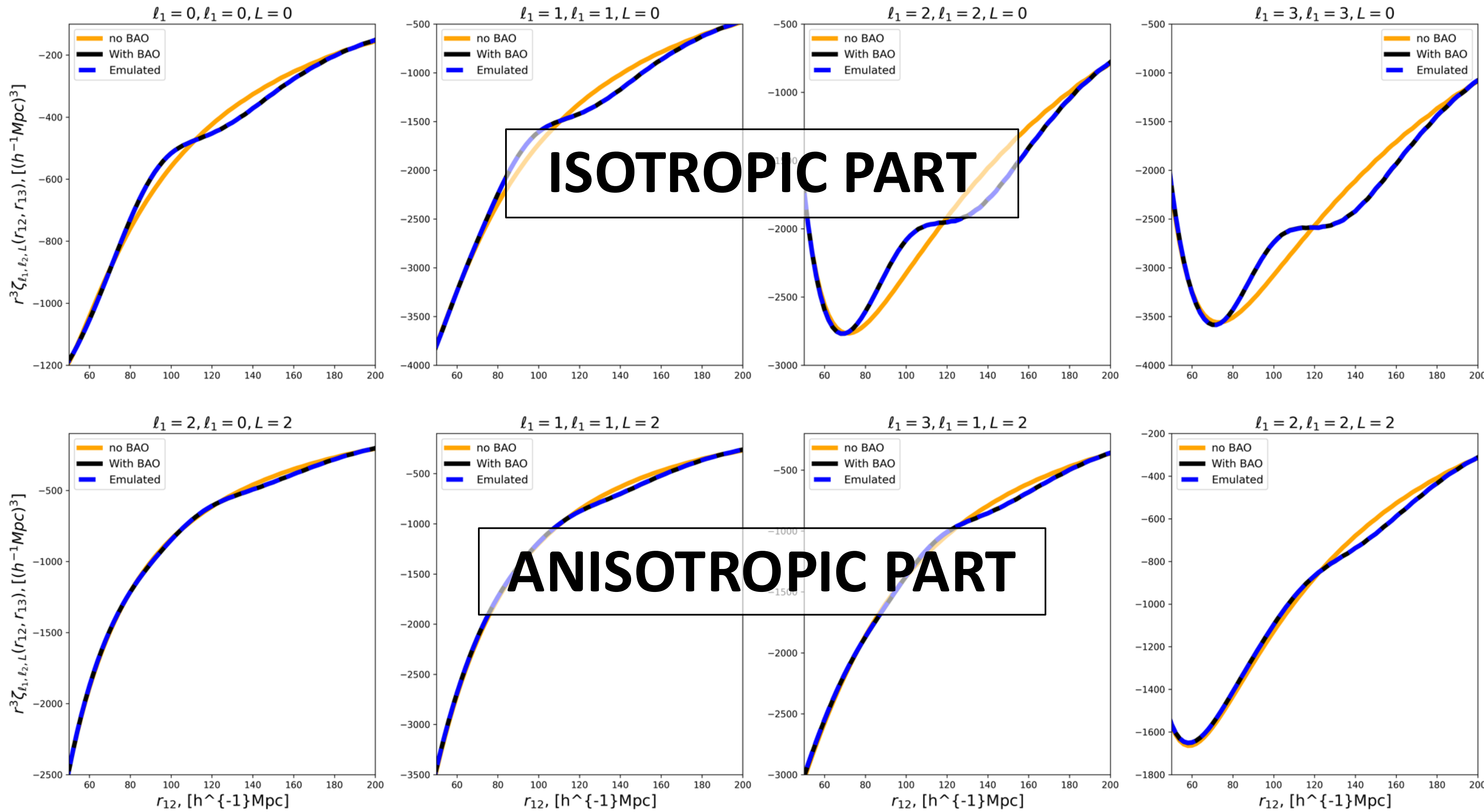


Models created with Mod3l *(Farina et al. 2024)*

Modelling the anisotropic 3PCF at BAO scales

Work by Kristers Nagainis

Aim: Develop an emulator for the anisotropic 3PCF of matter, to significantly speed up its computation and be able to provide forecasts on the accuracy of the constraints on cosmological parameters



Models created with Mod3l (Farina et al. 2024)

Step 0: setup an emulator

Work by Kristers Nagainis

Current typical time for estimating 1 model: **30-40 minutes!**

Idea: Compute a large library of models for the anisotropic 3PCF for matter with Mod3l, and use CosmoPower (Spurio Mancini 2021) to train an emulator

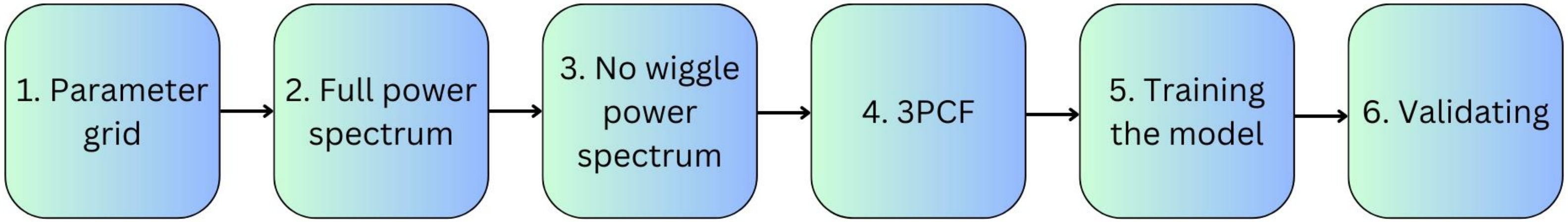
Step 0: setup an emulator

Work by Kristers Nagainis

Current typical time for estimating 1 model: **30-40 minutes!**

Idea: Compute a large library of models for the anisotropic 3PCF for matter with Mod3l, and use CosmoPower (Spurio Mancini 2021) to train an emulator

Starting point: emulating P(k) (more control, easier, allows us to test configurations



Step 0: setup an emulator

Work by Kristers Nagainis

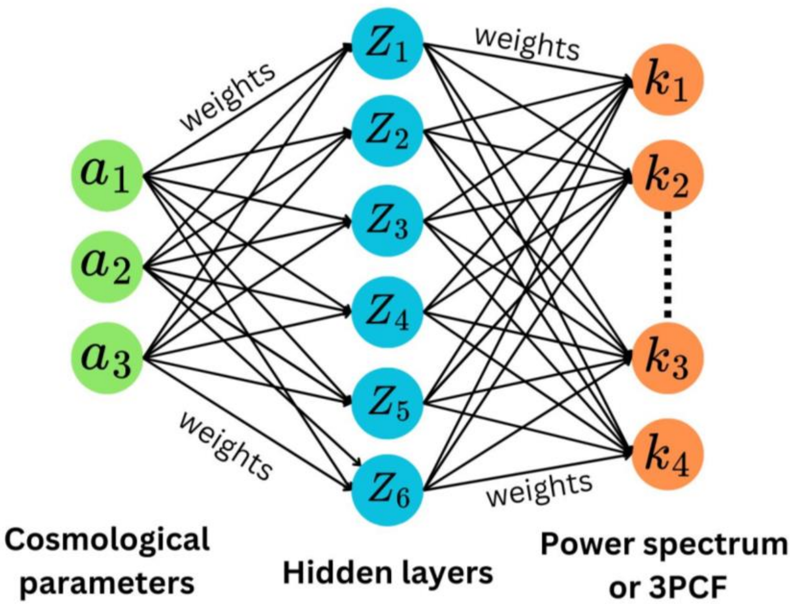
Current typical time for estimating 1 model: **30-40 minutes!**

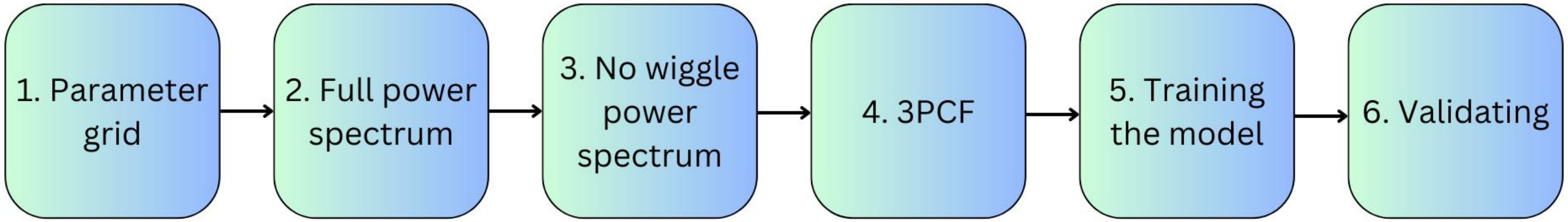
Idea: Compute a large library of models for the anisotropic 3PCF for matter with Mod3l, and use CosmoPower (Spurio Mancini 2021) to train an emulator

Starting point: emulating P(k) (more control, easier, allows us to test configurations

Large grid of configurations tested: batch size, numer of layers, number of neurons, binning of P(k), range of cosmological parameters

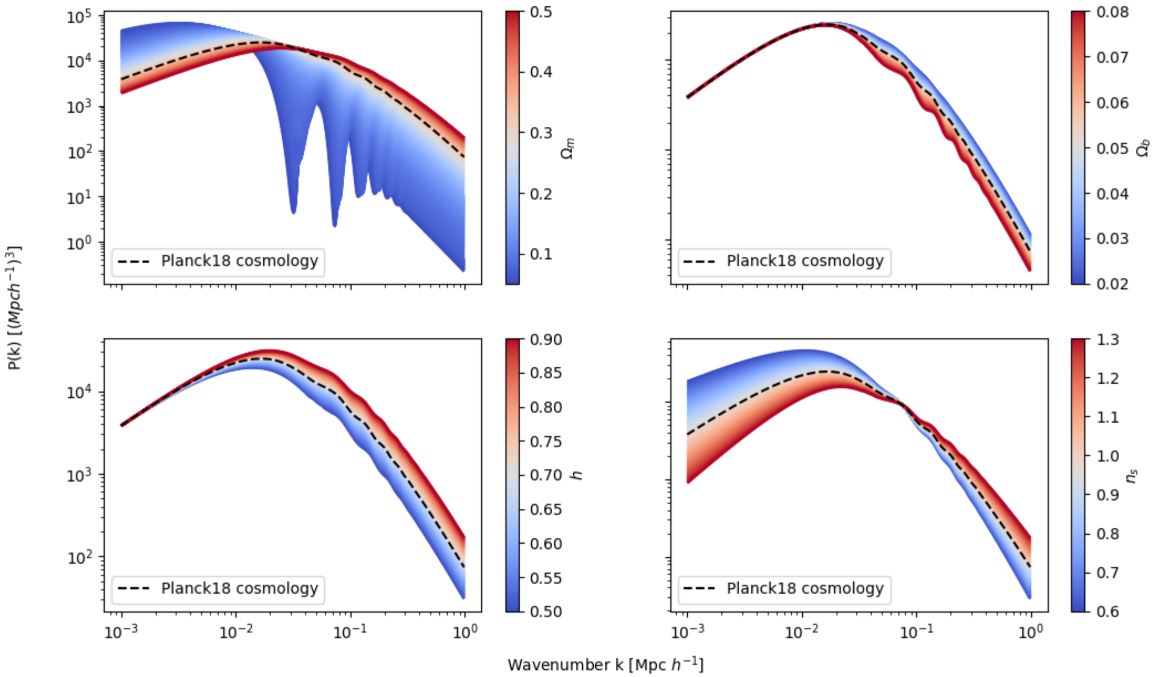
Aim: reach sub-% accuracy on the reconstructed P(k)





Step 1: emulating the power spectrum

Work by Kristers Nagainis



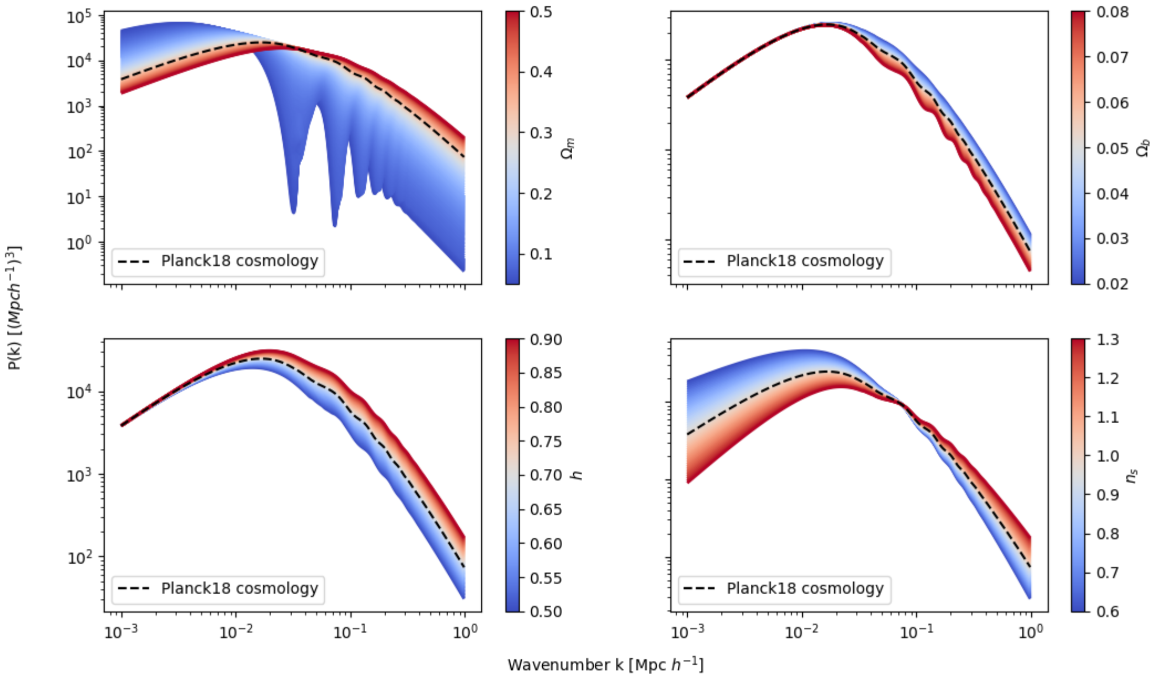
4 cosmological parameters considered: Ω_m, h, Ω_b, n_s

Emulated:

- $P_{\text{no wiggle}}(k)$
- $P(k)$ $\longrightarrow$ ingredients for B(k)
- $P_{\text{only wiggle}}(k)$

Step 1: emulating the power spectrum

Work by Kristers Nagainis

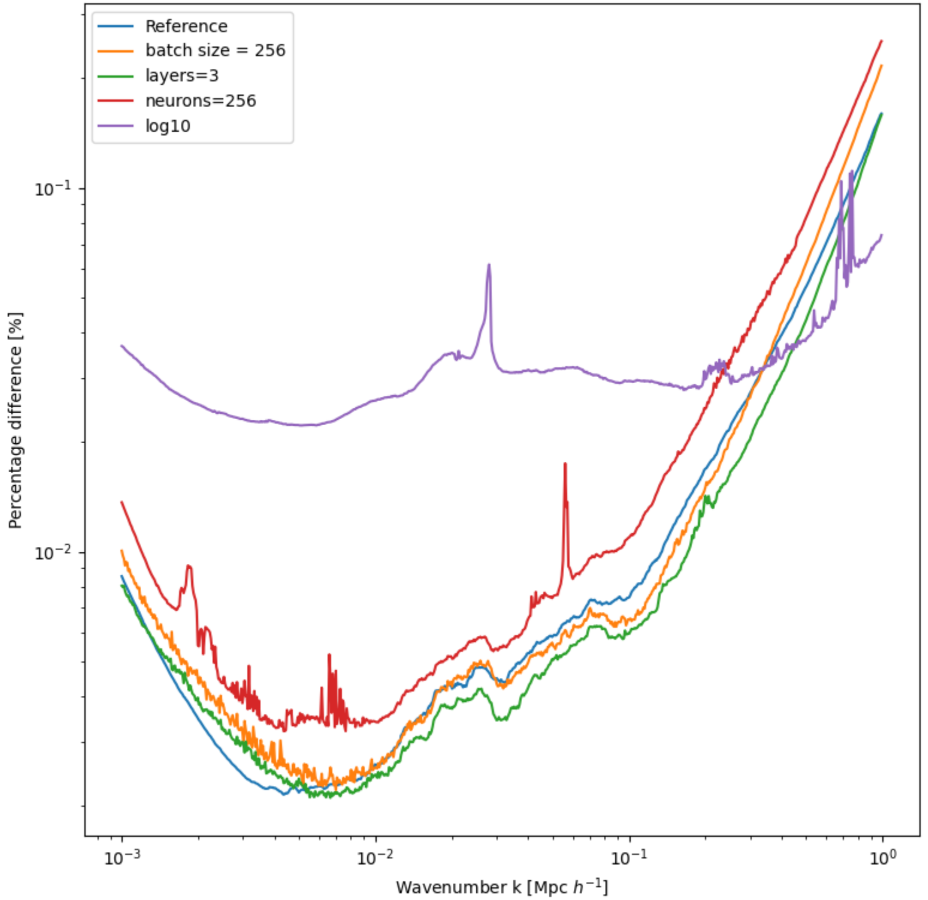


4 cosmological parameters considered: Ω_m, h, Ω_b, n_s

Emulated:

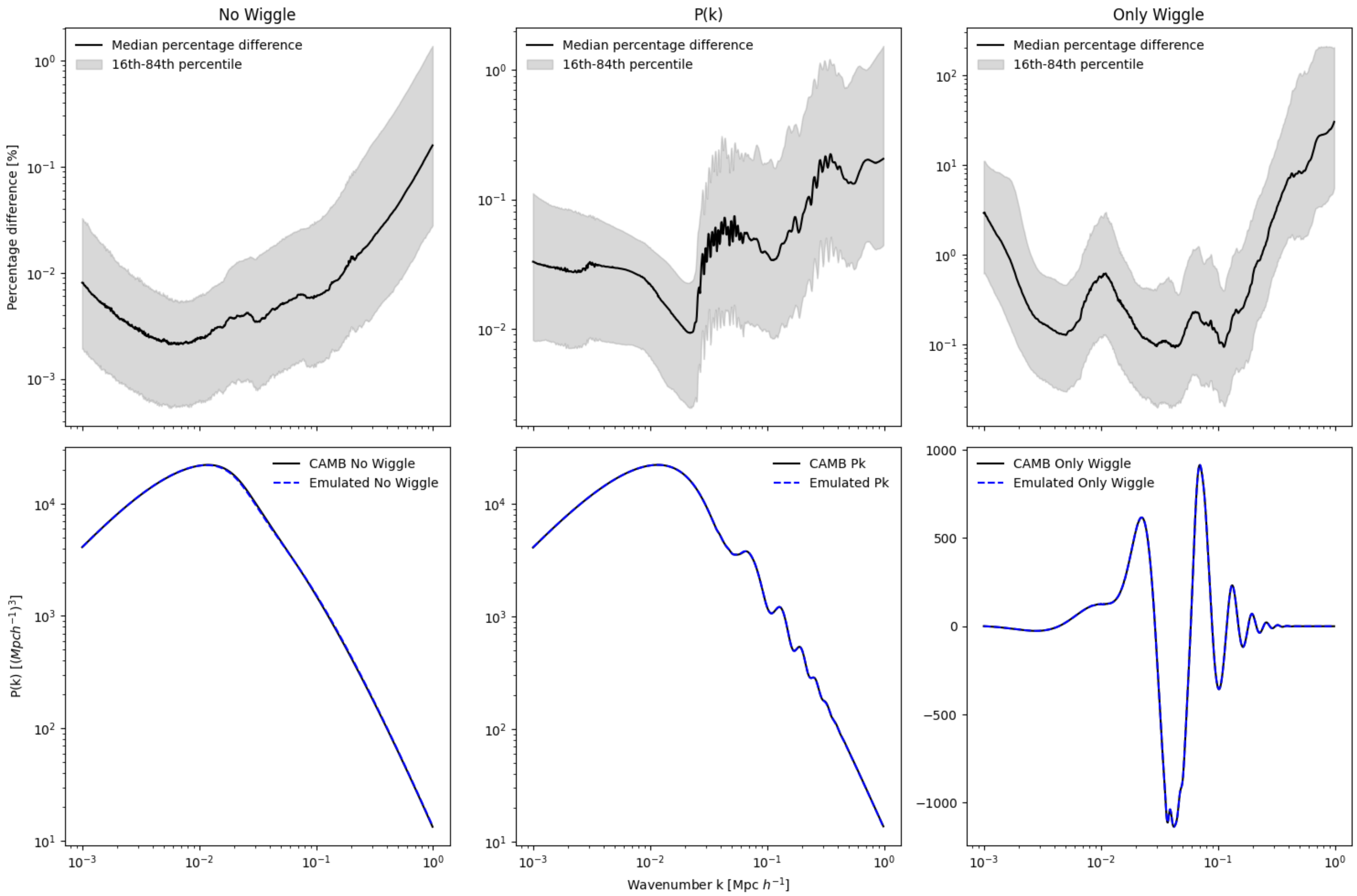
- $P_{\text{no wiggle}}(k)$
- P(k) ⟶ ingredients for B(k)
- $P_{\text{only wiggle}}(k)$

Thorough tests and otimization to get the best performances



Step 1: emulating the power spectrum

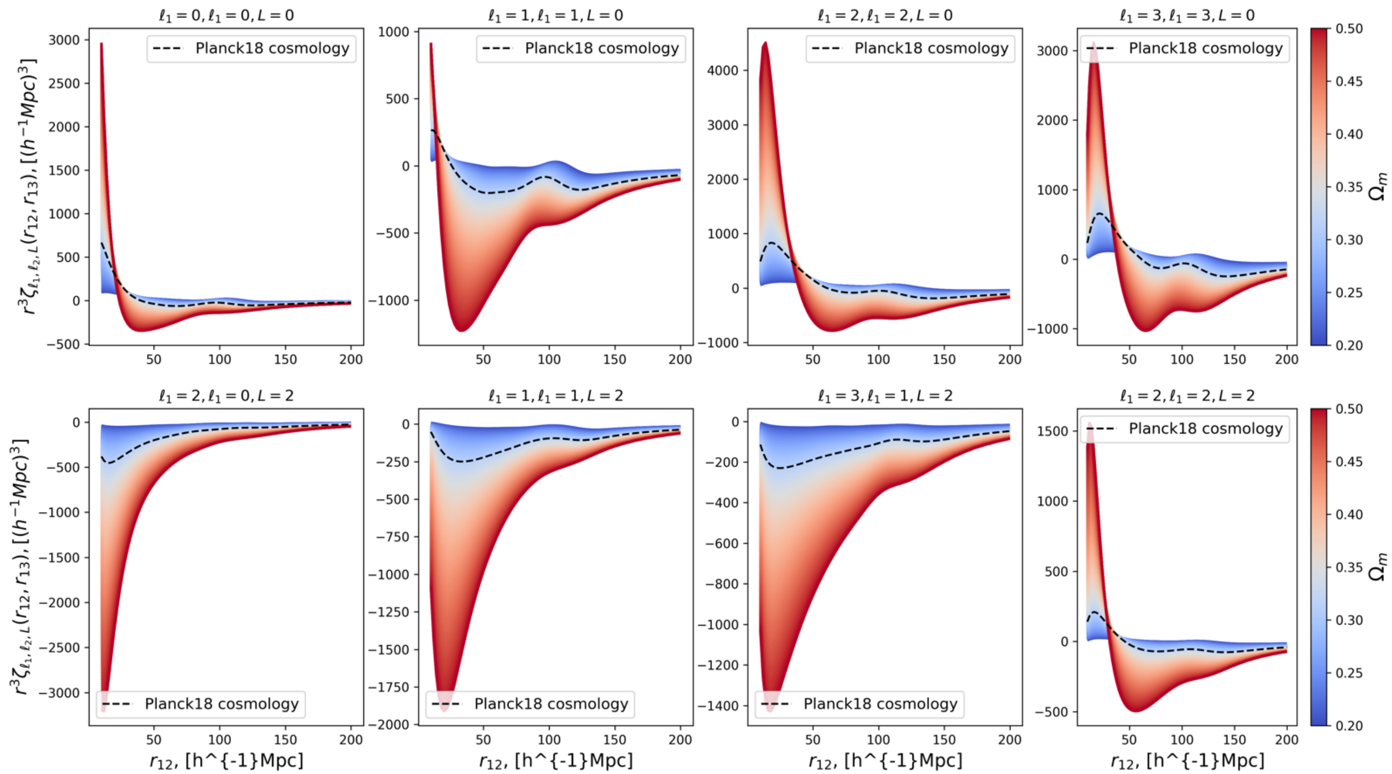
Work by Kristers Nagainis



- Sub-% accuracy reached in all configurations almost at all scales
- Emulated models implemented inside CosmoBolognaLib
- Significant gain in computational time

Step 2: emulating the 3PCF

Work by Kristers Nagainis

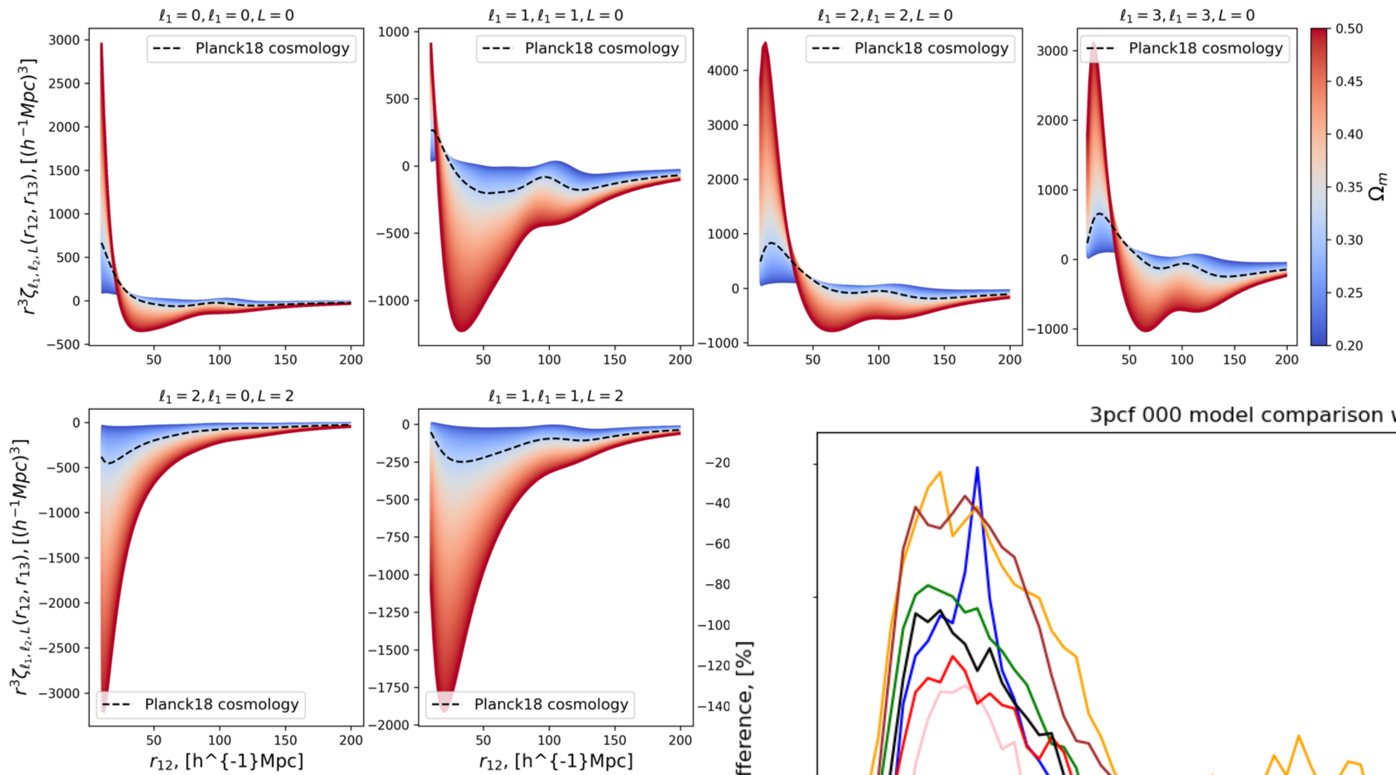


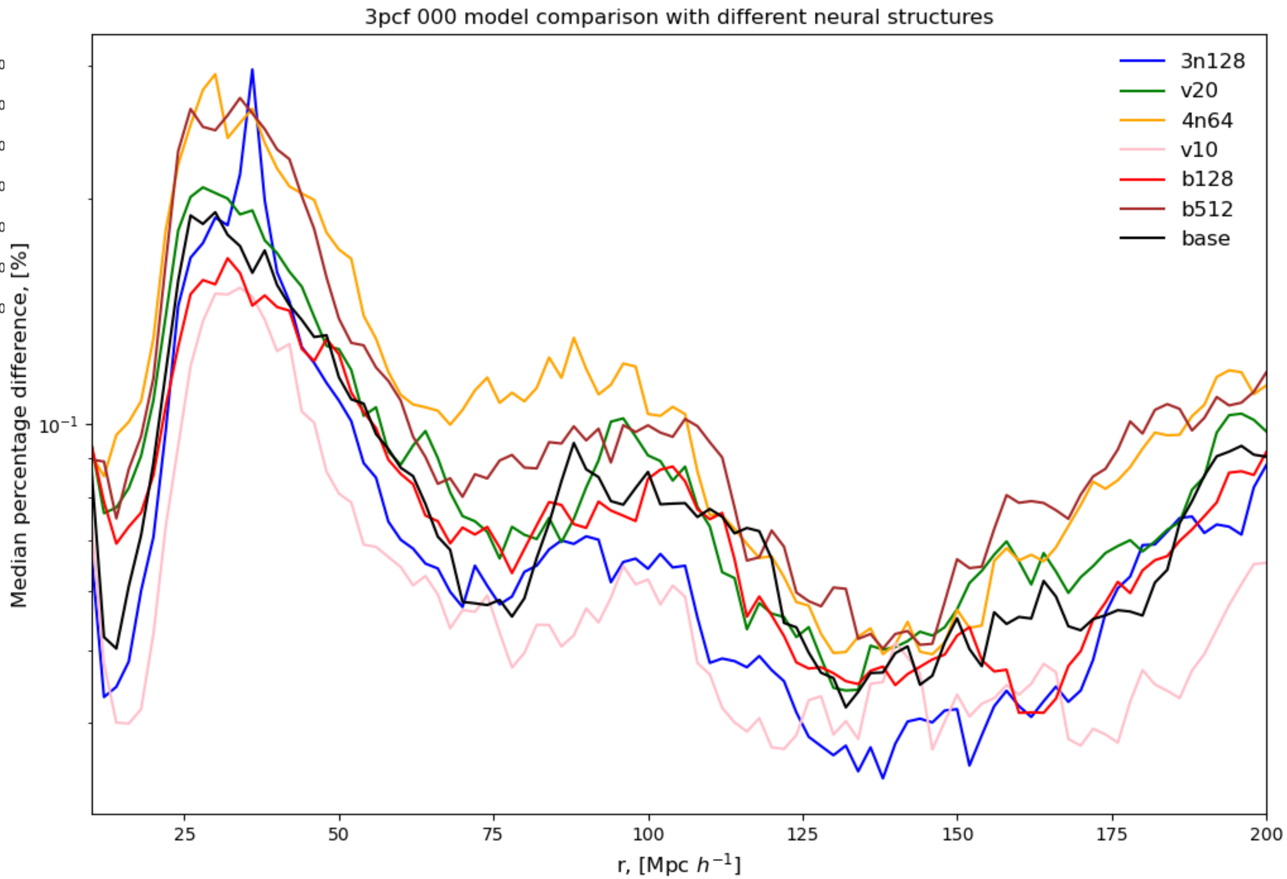
Similar approach to P(k), but required further implementations on CosmoPower

8 different emulators

Step 2: emulating the 3PCF

Work by Kristers Nagainis





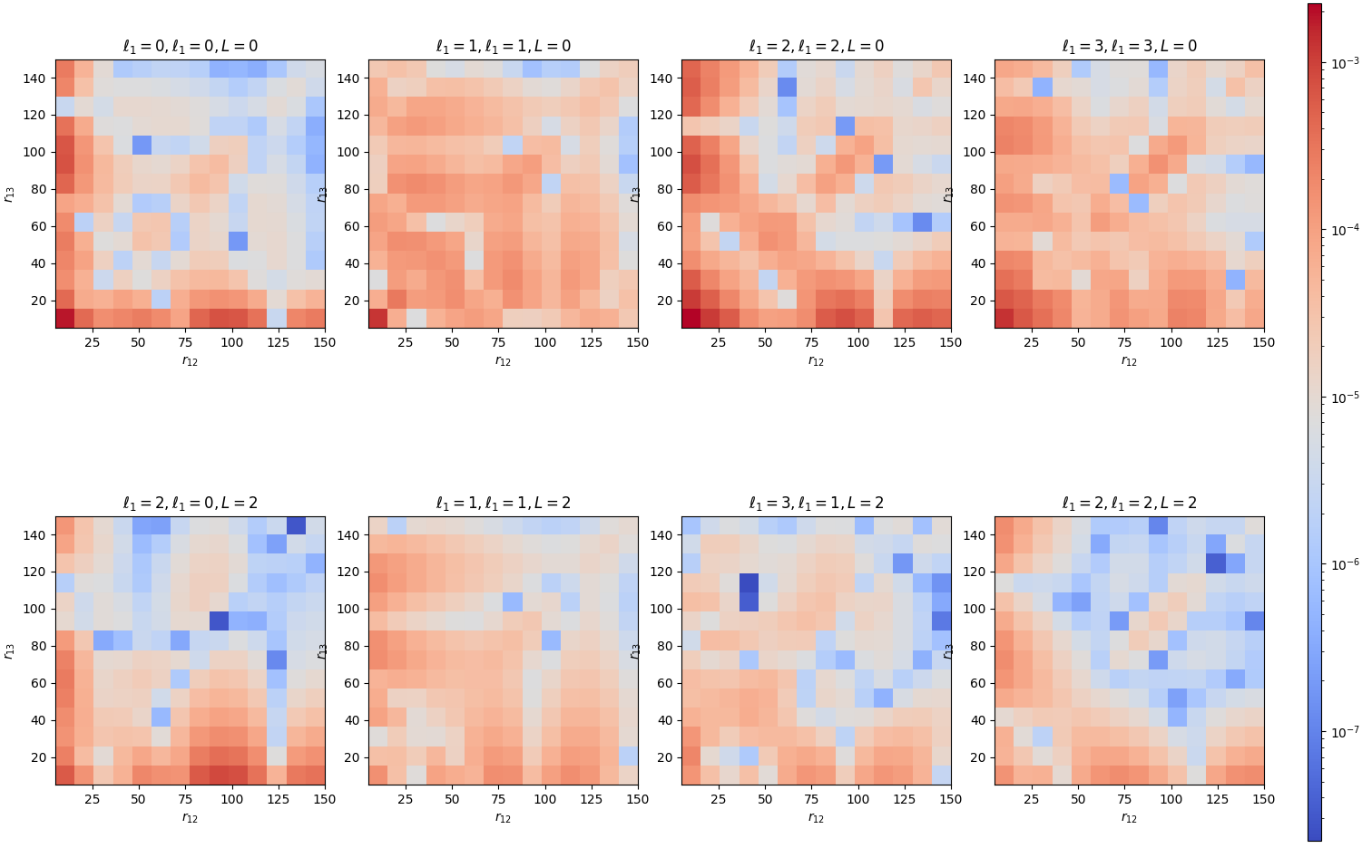
Similar approach to P(k), but required further implementations on CosmoPower

8 different emulators

Also **sub-% accuracy on the emulated 3PCF for all multipoles**

Step 3: BAO detectability

Work by Kristers Nagainis

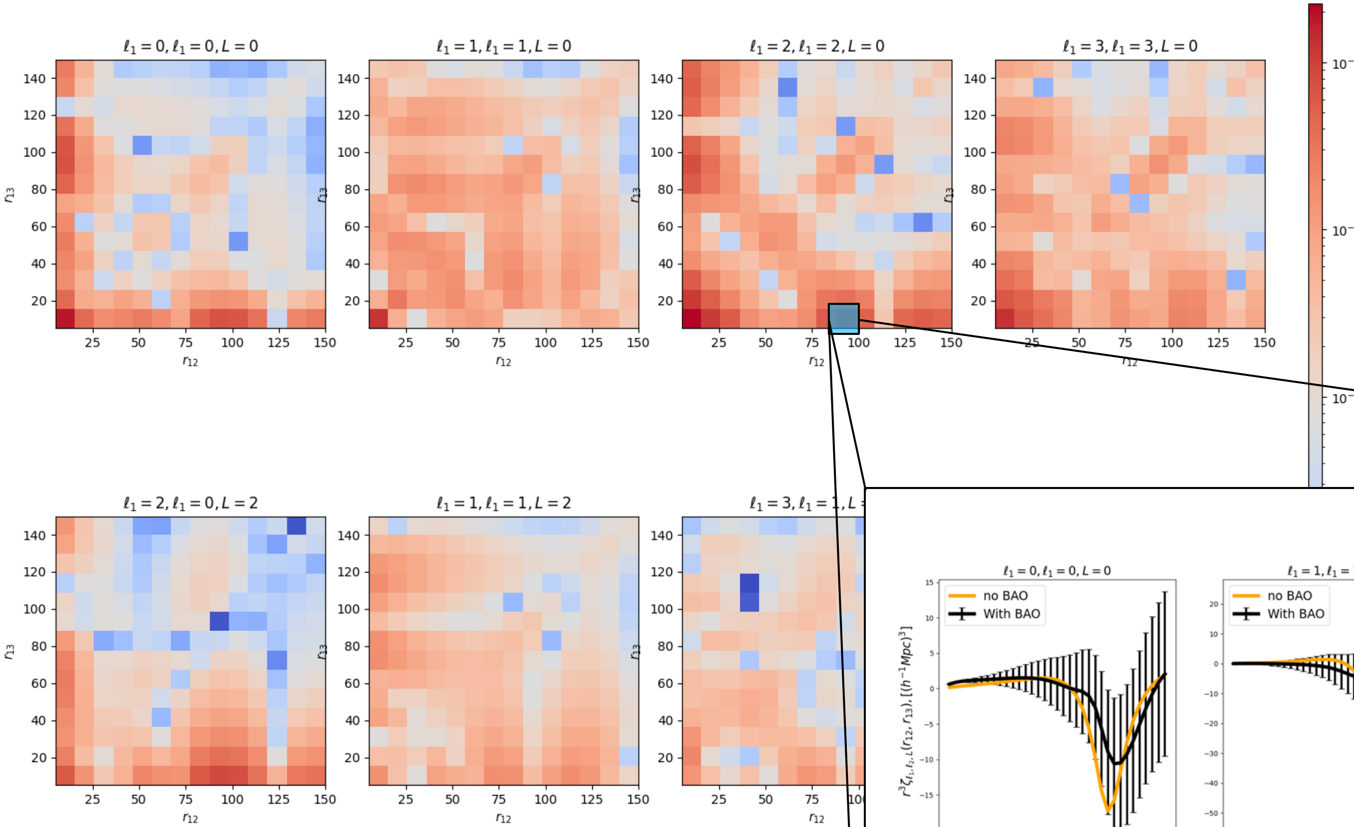


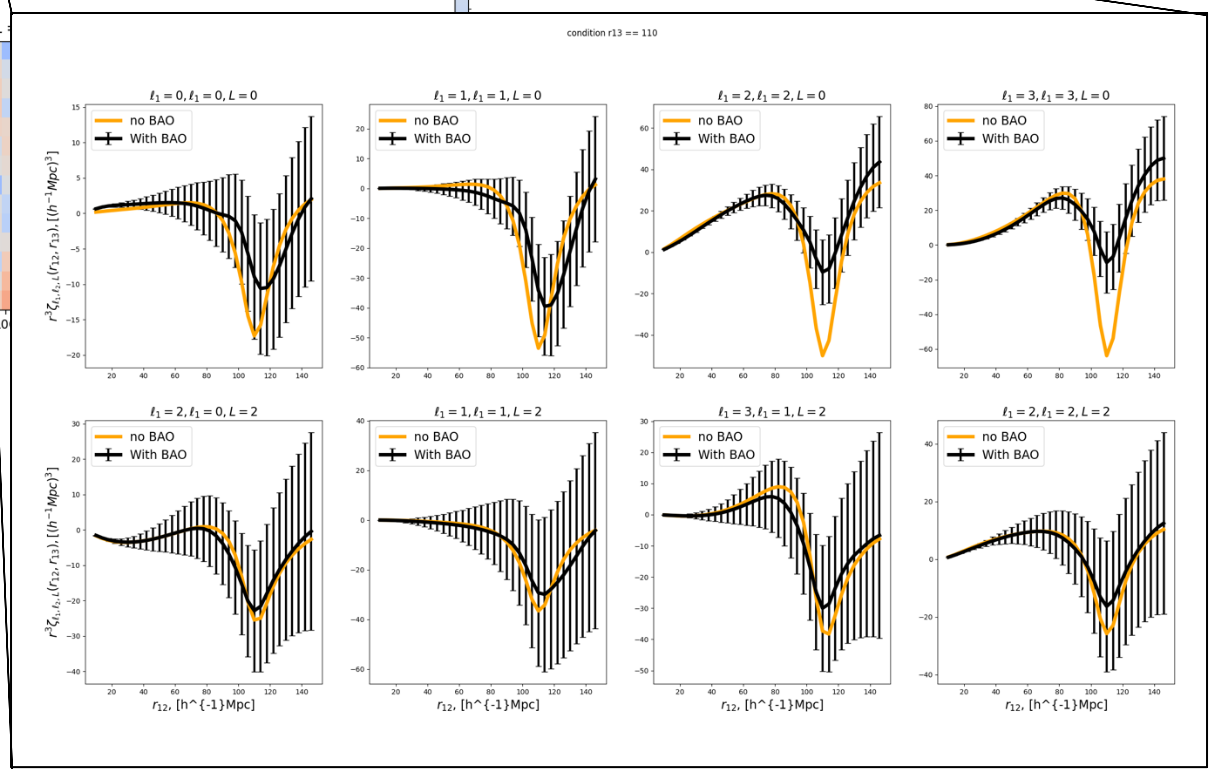
Analysis to detect the configurations providing the maximum signal to identify the BAO features

Theoretical covariance (thanks to A. Veropalumbo)

Step 3: BAO detectability

Work by Kristers Nagainis



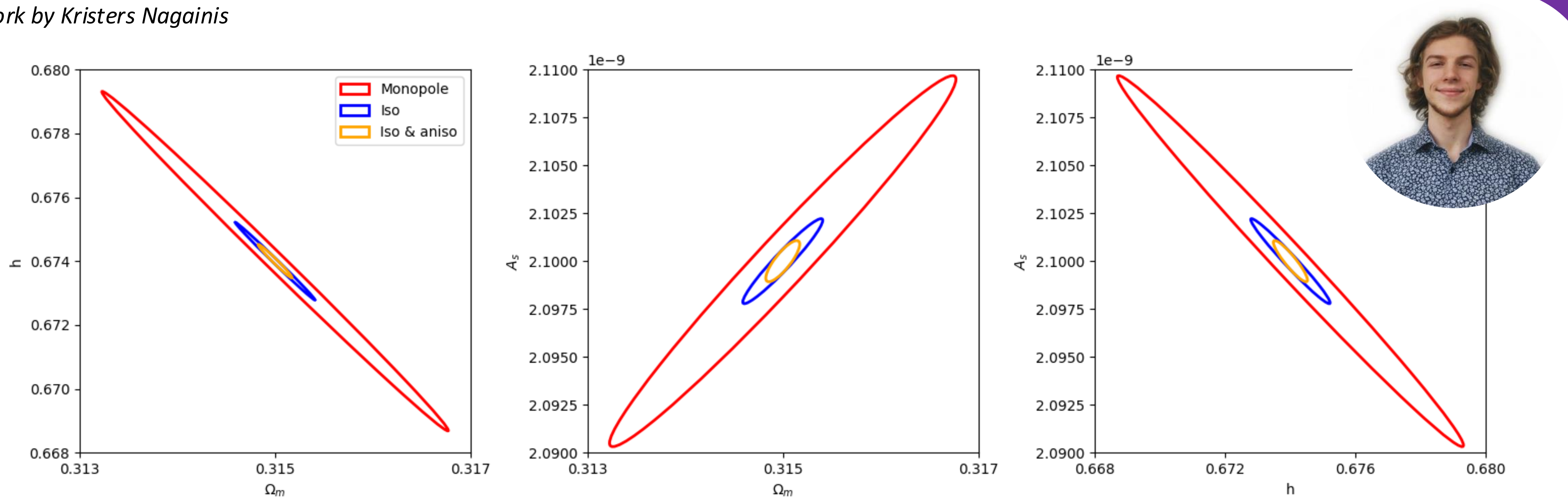


Analysis to detect the configurations providing the maximum signal to identify the BAO features

Theoretical covariance (thanks to A. Veropalumbo)

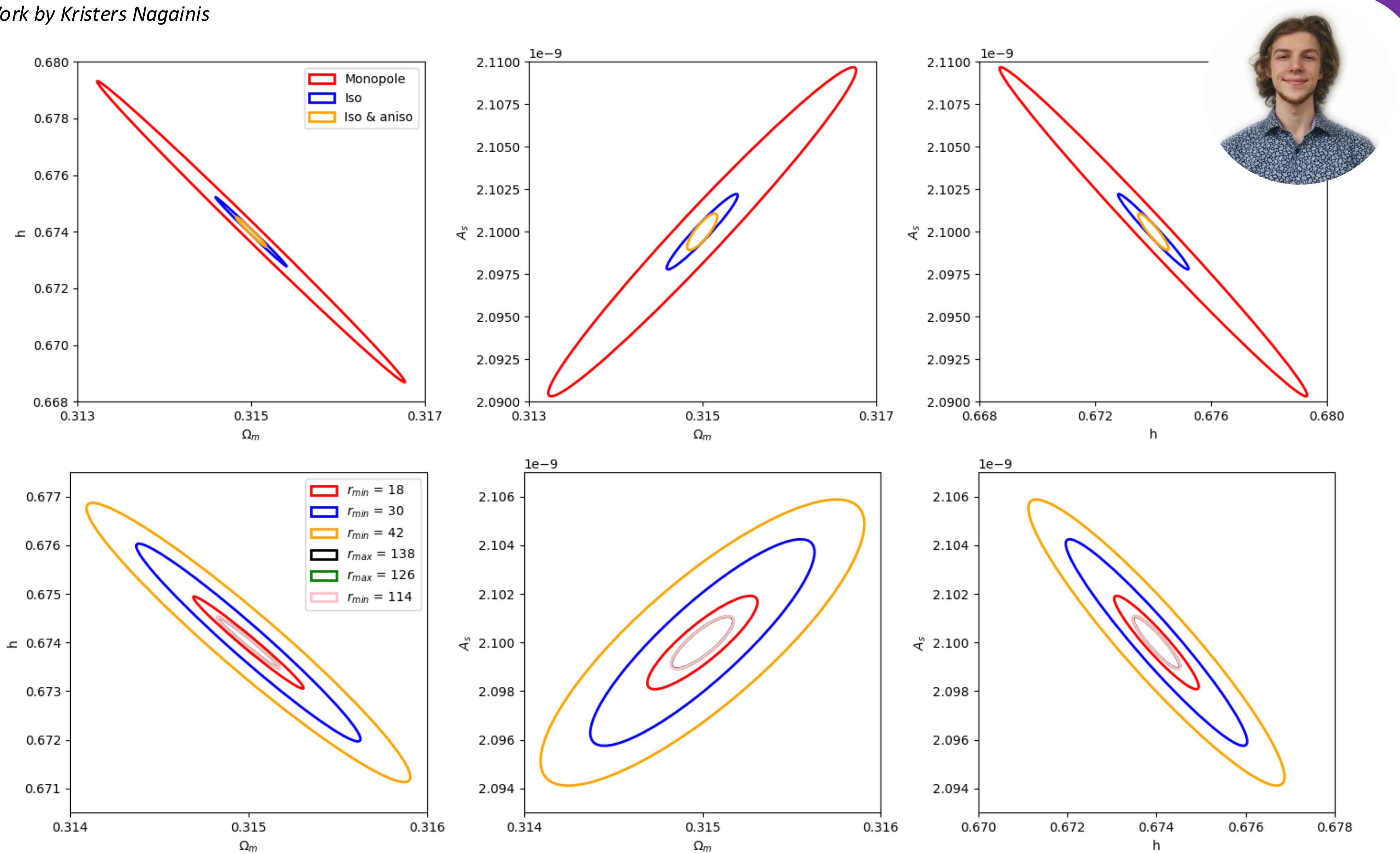
Step 4: Fisher forecasts

Work by Kristers Nagainis



Step 4: Fisher forecasts

Work by Kristers Nagainis



Promising results: **including the anisotropic component appears to improve the constraints**

Conclusions

- 3PCF as powerful statistical tool complementary to 2PCF
  - shape of structures depends on cosmological models
  - provide independent constraints on the bias and cosmological parameters

- **first assessment of the impact of interlopers on 3PCF**: bias which is crucial to be quantified and corrected
  - models to compensate for it under evaluation
  - combination of 2PCF and 3PCF crucial to break degeneracies and improve the constraints

- **For 3PCF cosmological applications and MCMC constraints it is fundamental to have ways to significantly speed up 3PCF models computation**:
  - created a framework for emulating P(k) and 3PCF, inside CosmoBolognaLib environment
  - emulated P(k) and anisotropic 3PCF models with sub-% accuracy
  - explored the configurations that maximize the BAO signal in the anisotropic 3PCF
  - Fisher forecasts on the accuracy of cosmological constraints from isotropic+anisotropic 3PCF

- **higher-order correlation function will be fundamental for future cosmological surveys** to maximize the extraction of cosmological information