Model-independent constraints on clustering and growth of cosmic structures from BOSS DR12 galaxies in harmonic space [arXiv: 2107.00026]

FZU
Institute of Physics of the Czech Academy of Sciences

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research, Development and Education

MŠMT
MINISTRY OF EDUCATION, YOUTH AND SPORTS

Central European Institute for Cosmology and Fundamental Physics, Prague

Konstantinos Tanidis

Unbearable Lightness of the Universe Workshop,
September 16th 2021

What is galaxy number counts or galaxy clustering?

Distribution of galaxies usually studied with the 3D galaxy Fourier PS: Fourier mode of the 3D separation between pairs of galaxies in the sky at a given redshift:

$$P_{gg}(k, \bar{z}) = b(\bar{z})^2 \, D^2(\bar{z}) \, P_{lin}(k) \text{ (approximation; leading order)}$$

What are the redshift-space distortions (RSD)?

- The background galaxies recede: expanding universe

- Galaxies also have their own peculiar velocities whose contributions are added to the main component of cosmological recession

- Result: Distribution of the galaxies in the redshift space is squashed and deformed : Kaiser effect (linear) & FoG (non-linear)

$$P_{gg}(k, \bar{z}) = [b(\bar{z}) + f(\bar{z})\mu^2]^2 \, D^2(\bar{z}) \, P_{lin}(k) \text{ [Kaiser effect]}$$

with $f(\bar{z}) = -d \, lnD(\bar{z})/d \, ln(1+\bar{z})$ and $\mu = \hat{\boldsymbol{r}} \cdot \boldsymbol{k}/k, \;\; k = |\boldsymbol{k}|$

Why to use the harmonic space PS instead of the galaxy Fourier PS?

- Natural tool and a direct observable (just measuring redshifts and angles); No Alcock-Paczynski correction
- Accounts for cosmic evolution
- Wide angle effects included
- Lensing is naturally included
- Tomography between different redshift shells is possible

$$C^{gg}_{ij,\ell} := \langle \Delta^i_{\ell m} \Delta^{j*}_{\ell m} \rangle = \frac{2}{\pi} \int dk\, k^2\, P_{lin}(k) \Delta^i_\ell(k) \Delta^j_\ell(k)\,, \quad \text{[FULL]}$$

$$\Delta^i_\ell(k) = \int dr\, n^i(r) D(r) [b(r) j_\ell(kr) - f(r) j''_\ell(kr)]$$

We can introduce the following derived quantities

$$b\sigma_8(z) := b(z)D(z)\sigma_8$$

$$\mathrm{f}\sigma_8(z) := f(z)D(z)\sigma_8$$

where σ_8 the r.m.s variance of clustering in spheres of radius $8h^{-1}$Mpc

Terms $b\sigma_8(z)$ and f$\sigma_8(z)$ can be factorised out of the integral for almost arbitrarily narrow redshift bins (possible in galaxy surveys with spectroscopic redshift accuracy). f$\sigma_8(z)$ is a good probe to discriminate between different Dark Energy models.

$$C^{gg}_{ij,\ell} = b\sigma_8^i b\sigma_8^j C^{\delta\delta}_{ij,\ell} + f\sigma_8^i f\sigma_8^j C^{VV}_{ij,\ell} - 2b\sigma_8^{(i} f\sigma_8^{j)} C^{\delta V}_{ij,\ell}\,,$$ [APPROX.]

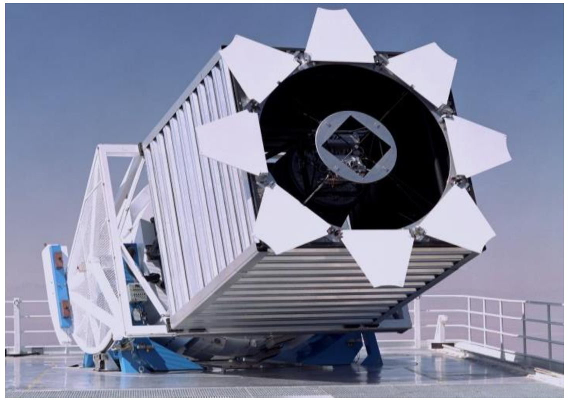
$$C^{\delta\delta}_{ij,\ell} := \frac{2}{\pi}\int dk\, dr\, d\tilde{r}\, k^2 \frac{P_{lin}(k)}{\sigma_8^2} n^i(r) n^j(\tilde{r}) j_\ell(kr) j_\ell(k\tilde{r})\,,$$

$$C^{VV}_{ij,\ell} := \frac{2}{\pi}\int dk\, dr\, d\tilde{r}\, k^2 \frac{P_{lin}(k)}{\sigma_8^2} n^i(r) n^j(\tilde{r}) j''_\ell(kr) j''_\ell(k\tilde{r})\,,$$

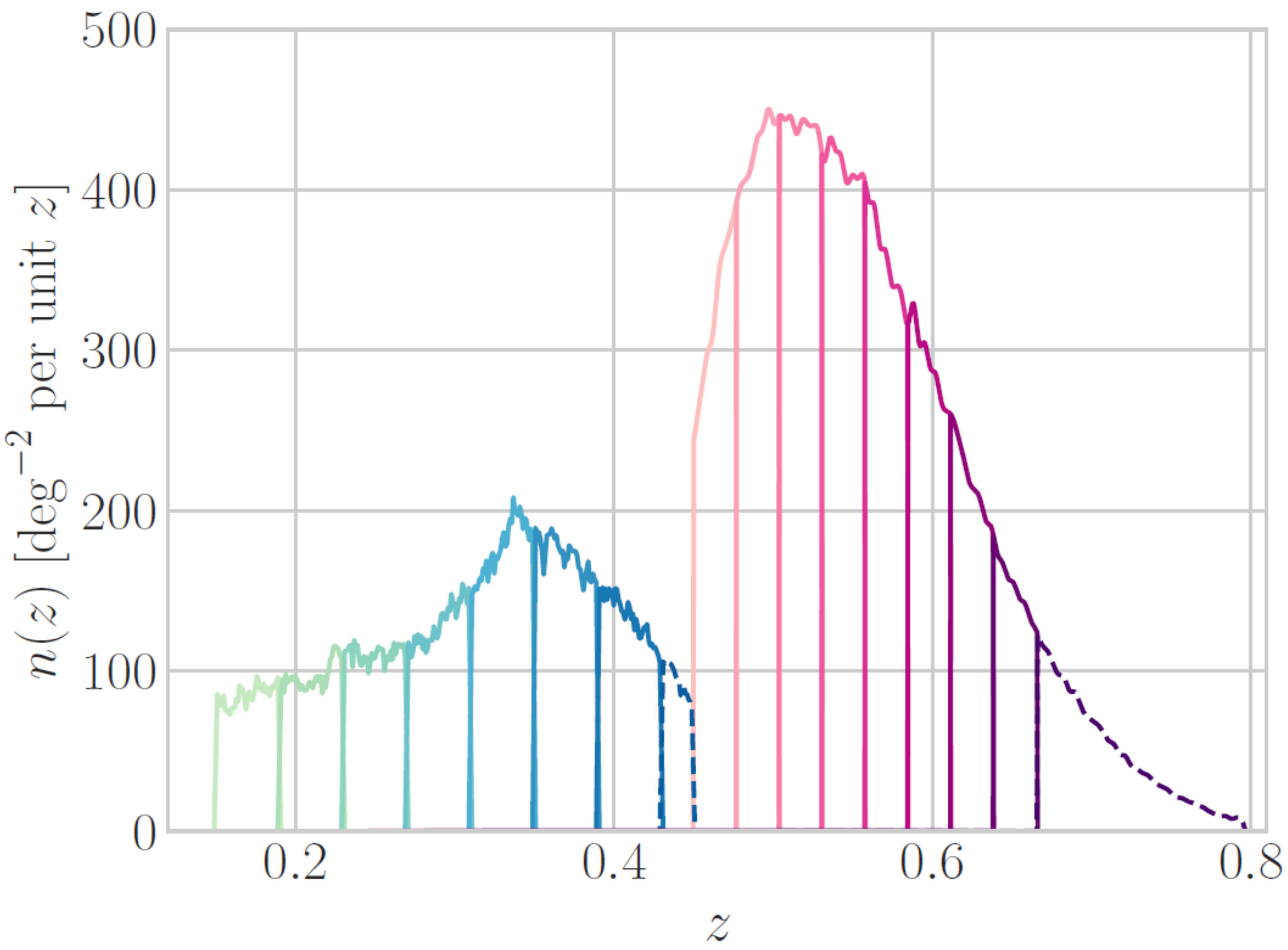
$$C^{\delta V}_{ij,\ell} := \frac{2}{\pi}\int dk\, dr\, d\tilde{r}\, k^2 \frac{P_{lin}(k)}{\sigma_8^2} n^{(i}(r) n^{j)}(\tilde{r}) j_\ell(kr) j''_\ell(k\tilde{r})\,,$$

Let’s measure it with BOSS data

BOSS DR12 spectroscopic survey

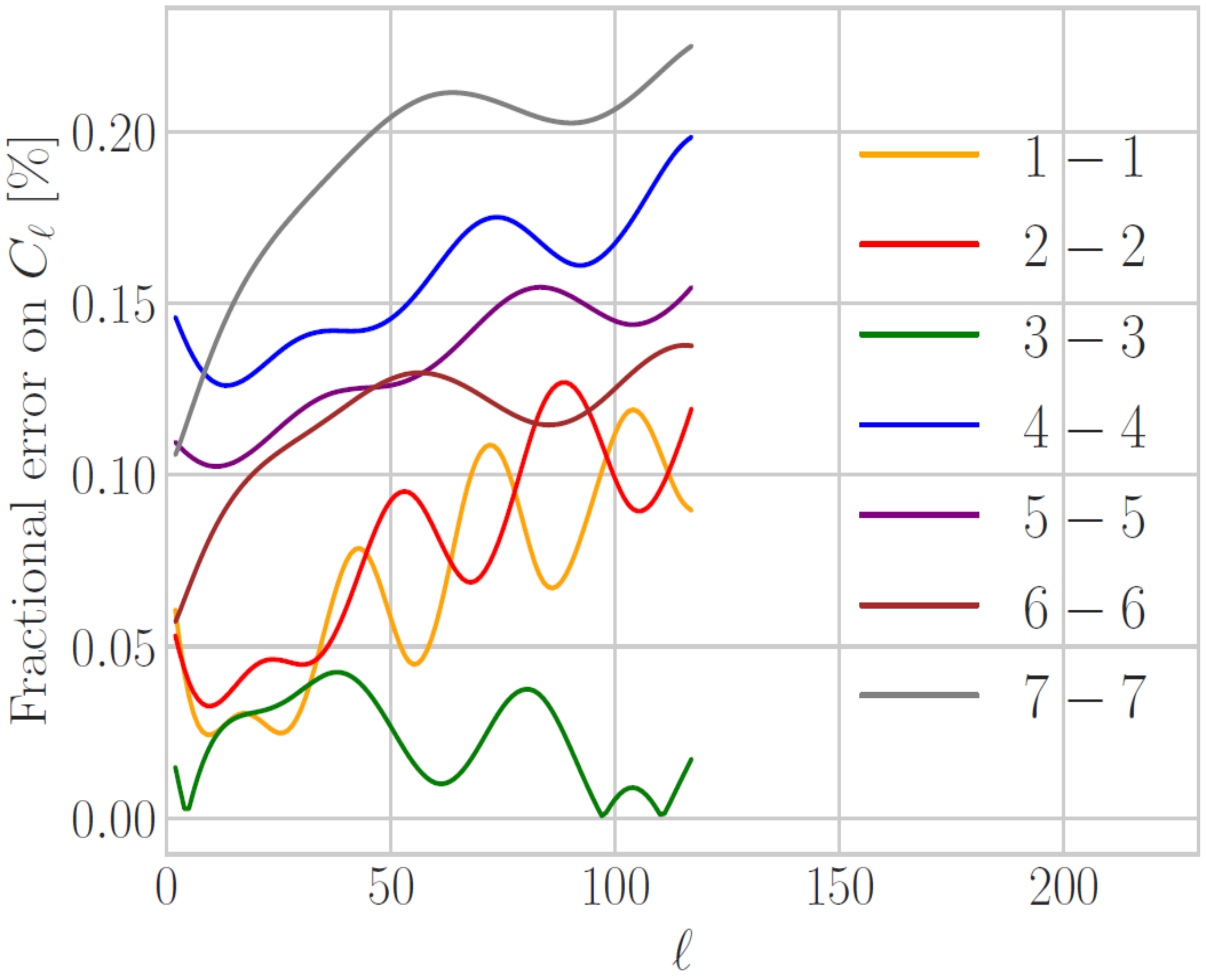


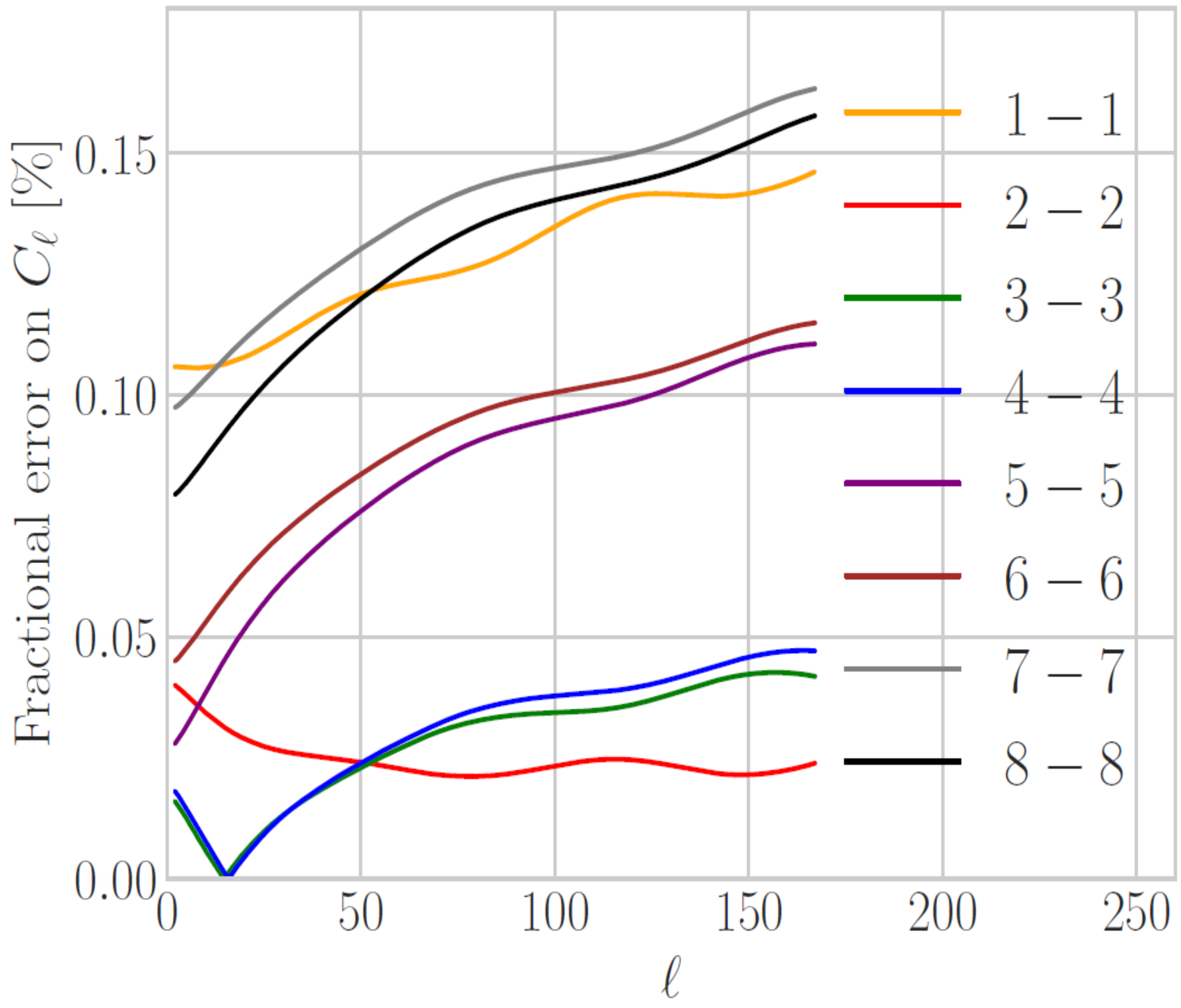
- Apache Point Observatory (APO): 2.5m telescope for 5 yr (2009-2014)
- Part of SDSS-III project: Baryon Oscillation Spectroscopic Survey
- About 10,000 sq.deg
- Two LRG subsamples: LOWZ and CMASS
- About 10^6galaxies in total



LOWZ CMASS

|FULL/APPROX.-1| X 100%





- Restrict analysis to strictly linear scales, $k_{max} = 0.1\ h\ Mpc^{-1}$,

$$\ell^i_{max} = k_{max} r(\bar{z}_i)$$

- Minimum multipole resolved, $\ell_{min} = \frac{\pi}{2 f_{sky}}$

Table 1. Redshift range, number of sources, shot noise, and maximum multipole for the redshift bins considered in the analysis.

Bin ID	$[z_{min}, z_{max}]$	# of gals	shot noise [sr]	ℓ^i_{max}
LOWZ-1	[0.150, 0.190]	33 906	8.08×10^{-5}	49
LOWZ-2	[0.190, 0.230]	38 728	7.07×10^{-5}	60
LOWZ-3	[0.230, 0.270]	44 291	6.18×10^{-5}	71
LOWZ-4	[0.270, 0.310]	51 781	5.27×10^{-5}	81
LOWZ-5	[0.310, 0.350]	70 879	3.86×10^{-5}	91
LOWZ-6	[0.350, 0.390]	68 701	3.98×10^{-5}	101
LOWZ-7	[0.390, 0.430]	53 191	5.15×10^{-5}	111
CMASS-1	[0.450, 0.477]	81 836	3.71×10^{-5}	123
CMASS-2	[0.477, 0.504]	112 589	2.67×10^{-5}	129
CMASS-3	[0.504, 0.531]	118 915	2.55×10^{-5}	135
CMASS-4	[0.531, 0.558]	113 139	2.68×10^{-5}	141
CMASS-5	[0.558, 0.585]	99 672	3.04×10^{-5}	147
CMASS-6	[0.585, 0.612]	80 914	3.75×10^{-5}	153
CMASS-7	[0.612, 0.639]	61 551	4.93×10^{-5}	159
CMASS-8	[0.639, 0.670]	44 057	6.89×10^{-5}	164

Construct the galaxy overdensity maps

- Galaxy catalogues public from BOSS collaboration

- Random catalogues are also available; 50 times denser than galaxy catalogues, taking into account:

1. Completeness masks (extent of complete observations)
2. Veto masks (various observational effects)

- Can be used to build binary angular masks
  - $N_{side} = 1024$ using HEALPix
  - Construct galaxy overdensity maps as

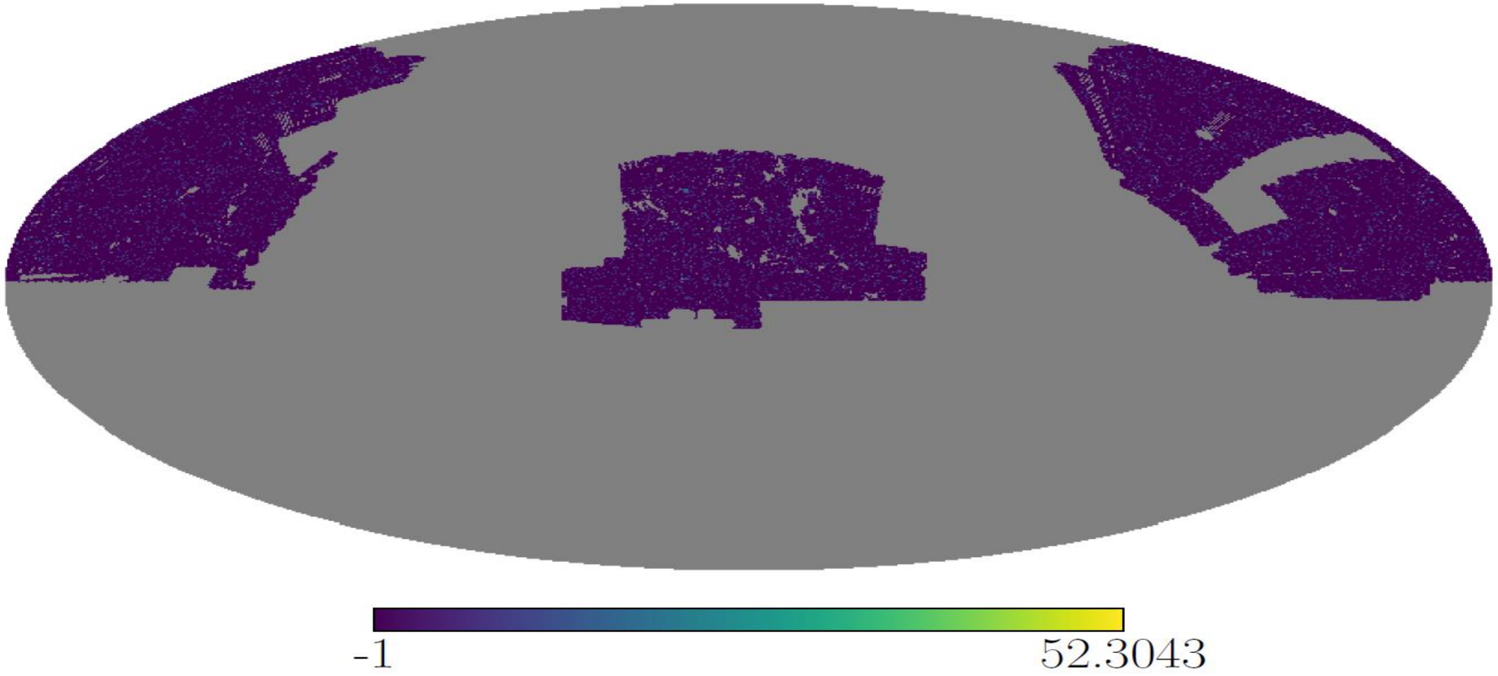
$$\delta_g^p = (n_g^p - \bar{n}_g) / \bar{n}_g$$

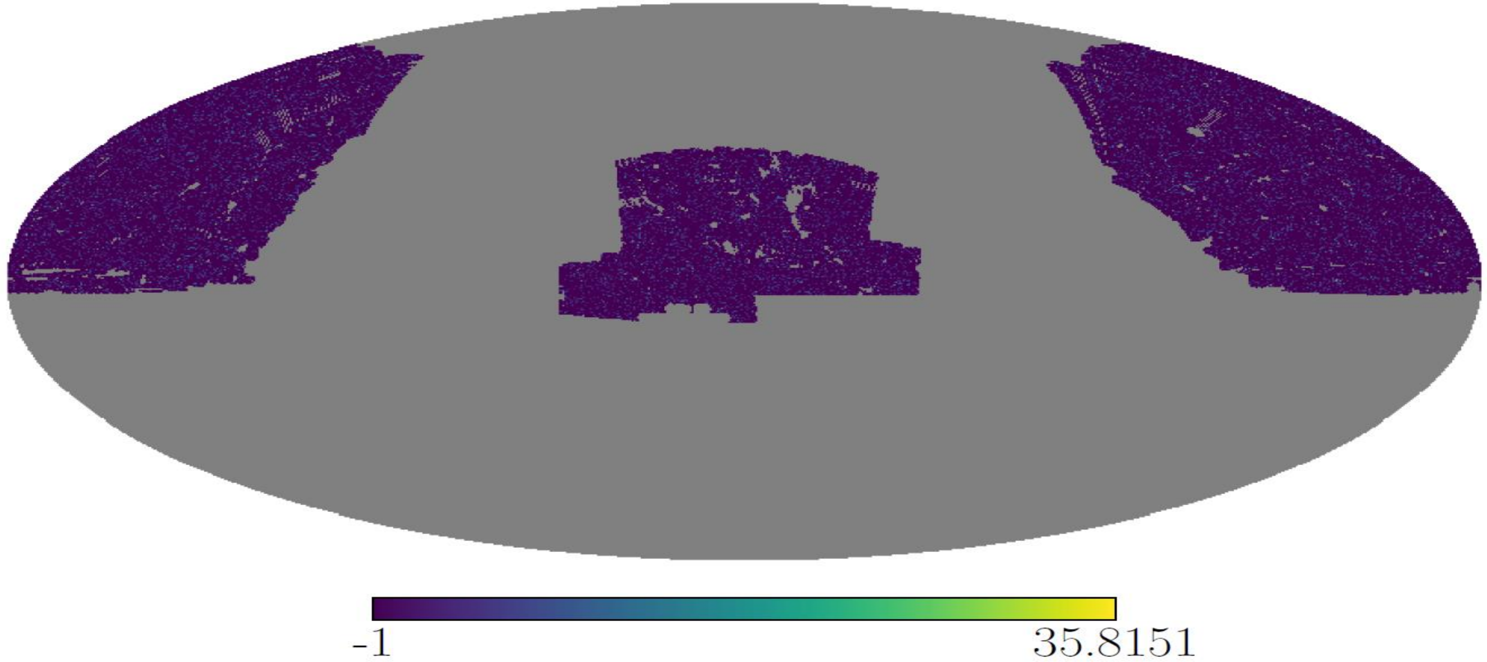
with n_g^p the number of weighted galaxies in a given pixel p
and $\bar{n}_g$ the mean number of weighted galaxies per pixel

$f_{sky} = 0.2081$ for LOWZ

$f_{sky} = 0.2416$ for CMASS

A common $\ell_{min} \approx 7$





Pseudo-C_ℓ approach & mask effects

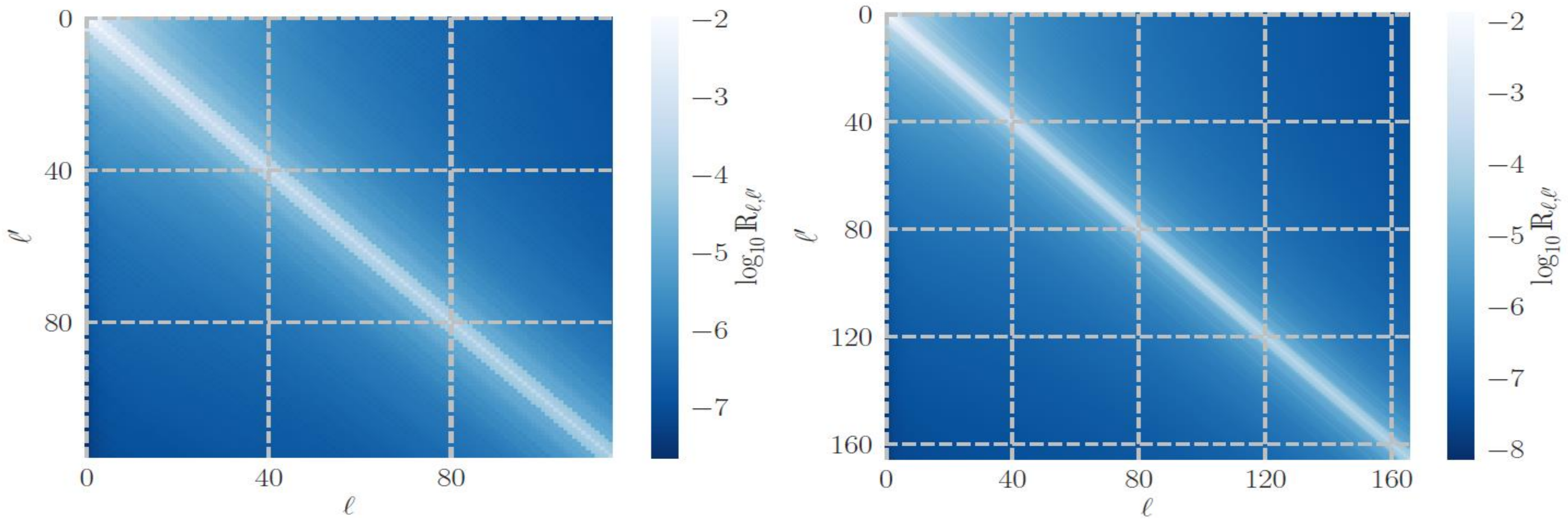
- Project the observed field onto the celestial sphere
- Decompose it to spherical harmonics
- Analyse statistically the coefficients of the decomposition

$$\widetilde{C_{ij,\ell}^{gg}} = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}\left|\delta_g^{\ell m}\right|^2 - \delta_{ij}^K/\bar{n}_g^i$$

- The mask introduces coupling between different modes: Measured field is related to the unmasked field

$$\left\langle\widetilde{C_{ij,\ell}^{gg}}\right\rangle = \sum_{\ell'}\mathcal{R}_{\ell\ell'}\,\widetilde{C_{ij,\ell}^{gg}}$$

With the coupling matrix $\mathcal{R}_{\ell\ell'} = \frac{2\ell'+1}{4\pi}\sum_{\ell''}(2\ell''+1)W_{\ell''}\begin{pmatrix}\ell & \ell' & \ell''\\ 0 & 0 & 0\end{pmatrix}^2$



- A way to overcome this is with the forward modelling; convolve the coupling matrix with the theory spectra (usually time consuming)
- Another is to invert the coupling matrix and deconvolve it from the data; Direct inversion impossible since matrix becomes singular; A solution is to introduce bandpowers (pymaster package)

Let s a set of 8 multipoles with weights w_s^ℓ, and $\sum_{\ell \in s} w_s^\ell = 1$

The coupled pseudo-C_ℓ in s bandpower:

$$\widetilde{B_{ij,\ell}^{gg}} = \sum_{\ell \in S} w_s^\ell \langle \widetilde{C_{ij,\ell}^{gg}} \rangle,$$

With expectation value:

$$\left\langle \widetilde{B_{ij,\ell}^{gg}} \right\rangle = \sum_{\ell \in S} w_s^\ell \sum_{\ell'} \boldsymbol{\mathcal{R}}_{\ell\ell'} \widetilde{C_{ij,\ell'}^{gg}}$$

Thus assuming the true power spectrum is a stepwise function related to bandpowers:

$$\widetilde{C_{ij,\ell}^{gg}} = \sum_{s} \widetilde{B_{ij,s}^{gg}} \, \Theta(\ell \in s)$$

Then an unbiased estimator can be derived:

$$B_{ij,s}^{gg} = \sum_{s'} \boldsymbol{\mathcal{M}}_{ss'}^{-1} \widetilde{B_{ij,s'}^{gg}}$$

With the binned coupling matrix now $\boldsymbol{\mathcal{M}}_{ss'} = \sum_{\ell \in s} \sum_{\ell' \in s'} w_s^\ell \boldsymbol{\mathcal{R}}_{\ell\ell'}$

The theory spectra should be binned accordingly

$$B_{ij,s}^{gg,th} = \sum_{\ell} \left(\sum_{s'} \boldsymbol{\mathcal{M}}_{ss'}^{-1} \sum_{\ell' \in S'} w_{s'}^{\ell'} \boldsymbol{\mathcal{R}}_{\ell'\ell} \right) \cdot C_{ij,\ell}^{gg}$$

Construct the Gaussian likelihood

$$\chi^2(\boldsymbol{\theta}) = [\boldsymbol{d} - \boldsymbol{t}(\boldsymbol{\theta})]^T \, \boldsymbol{C}^{-1}[\boldsymbol{d} - \boldsymbol{t}(\boldsymbol{\theta})]$$

with $\boldsymbol{d} = \{B_{ij,s}^{gg}\}$, $\boldsymbol{t}(\boldsymbol{\theta}) = \{B_{ij,s}^{gg,th}(\boldsymbol{\theta})\}$ and $\boldsymbol{\theta} = \{b\sigma_8^i, f\sigma_8^i\}$

For the theory spectra we assume a ΛCDM fiducial cosmology as from Planck 2016

We will put constraints on the parameter set $\boldsymbol{\theta}$ using three different approaches to estimate the data covariance matrix $\boldsymbol{C}$

Gaussian Covariance: diagonal in multipole space

$$\boldsymbol{C}_{ss'}^{ij,kl} = \frac{\delta_{ss'}^{K}}{(2s+1)\Delta s\, \mathrm{f_{sky}}} \left(\widehat{B_{ik,s}^{gg}}\ \widehat{B_{jl,s}^{gg}} + \widehat{B_{il,s}^{gg}}\ \widehat{B_{jk,s}^{gg}}\right)$$

with Δs the multipole range in the bandpower binning

$$\widehat{B_{ij,s}^{gg}} = B_{ij,s}^{gg} + \frac{\delta_K^{ij}}{\bar{n}_g^i}$$

PolSpice Covariance: accounting for coupling due to the mask

$$\mathcal{C}_{ss'} = \mathcal{M}^{-1}_{ss'} \mathcal{V}_{ss'} [(\mathcal{M}^{-1})_{ss'}]^T$$

with $\mathcal{V}_{ss'} = \frac{2 \boldsymbol{d}_s \boldsymbol{d}_{s'} \mathcal{M}_{ss'}}{2s'+1}$

and $\boldsymbol{d}_s$ the data vector in bandpower s

Mocks Covariance

- Realistic way to estimate covariances is with data simulations

- BOSS offers simulated galaxy catalogues with PATCHY and QPM mocks

  1. Constructed with a fixed cosmological model
  2. Include various observational and systematics effects

- Due to their different redshift ranges, for LOWZ we use QPM mocks ($z>1.5$) while for CMASS the PATCHY ones ($0.2<z<0.75$)

- A further test to validate our analysis by using different mocks for each sample

$$\mathcal{C}_{ss'} = \frac{1}{\mathrm{N_m} - 1} \sum_{m=1}^{N_m} (d_s^m - \bar{d}_s^m)\,(d_{s'}^m - \bar{d}_{s'}^m)^T$$

With $m = 1 \ldots N_m$, $N_m = 1000(2048)$ mocks for QPM(PATCHY), $d_s^m = \{B_{ik,s}^{gg,m}\}$ and

$$\bar{d}_s^m = \frac{1}{\mathrm{N_m}} \sum_{m=1}^{N_m} \bar{d}_s^m$$

Modify the likelihood:

- The inverse of the simulated covariance matrix is not an unbiased estimator of the true covariance

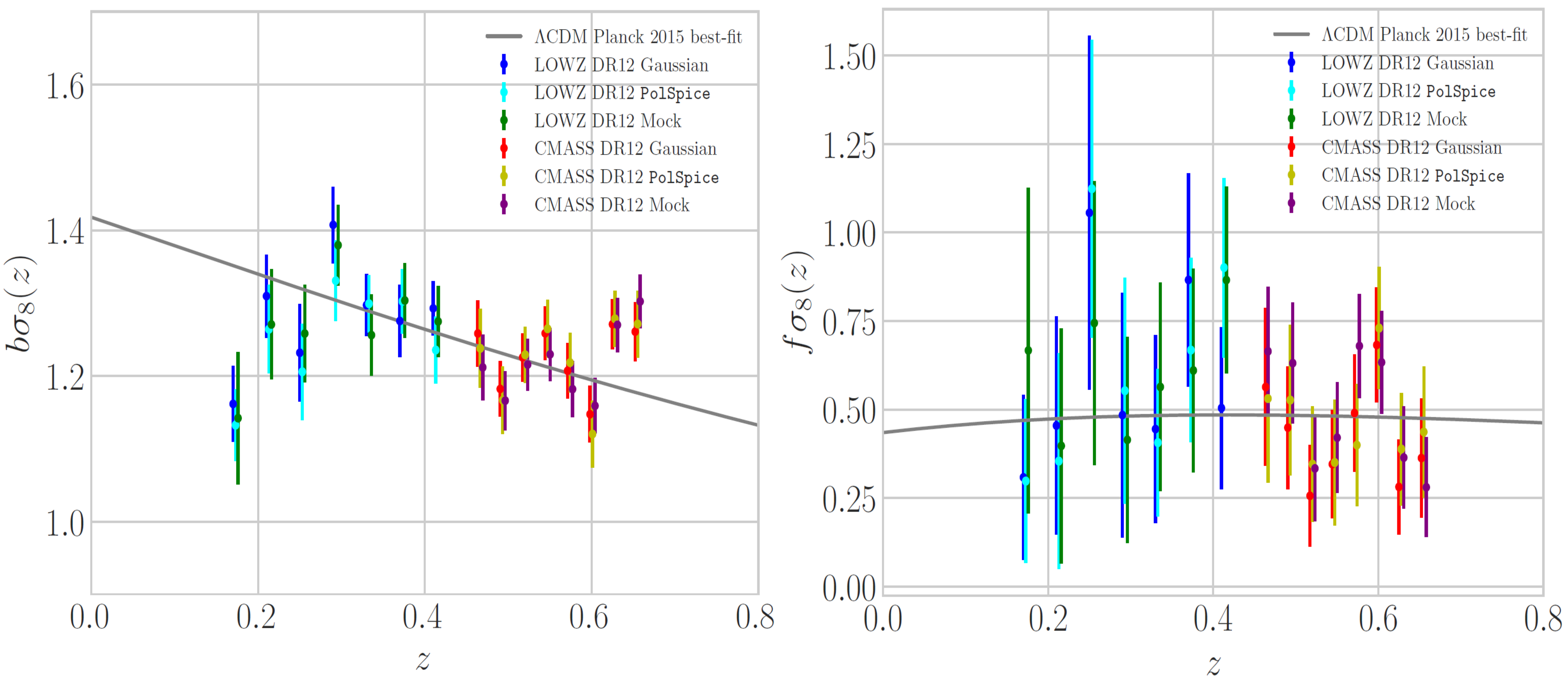
$\mathcal{C}^{-1} \rightarrow \frac{\mathrm{N_m - N_d - 2}}{\mathrm{N_m - 2}}\, \mathcal{C}^{-1}$ and keep the Gaussian likelihood, **OR**

- Replace the Gaussian likelihood with a t-Student

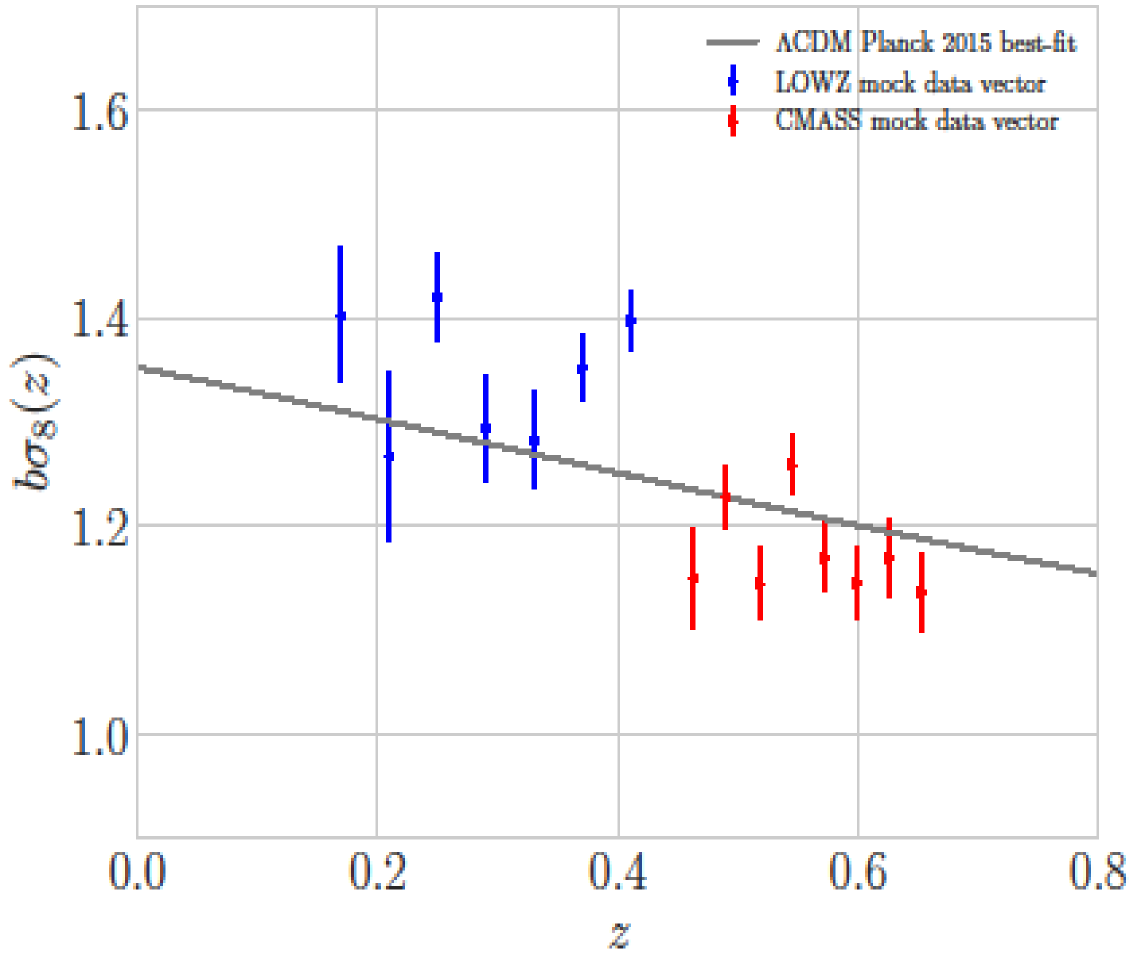
$\mathcal{L} \propto \left[1 + \frac{\chi^2}{N_m}\right]^{-N_m/2}$ and normalising with $\frac{\Gamma(N_m/2)[\det(\mathcal{C})]^{-1/2}}{[\pi(N_m-1)]^{N_d/2} - \Gamma[(N_m-N_d)/2]}$ assuming $N_m > N_d$

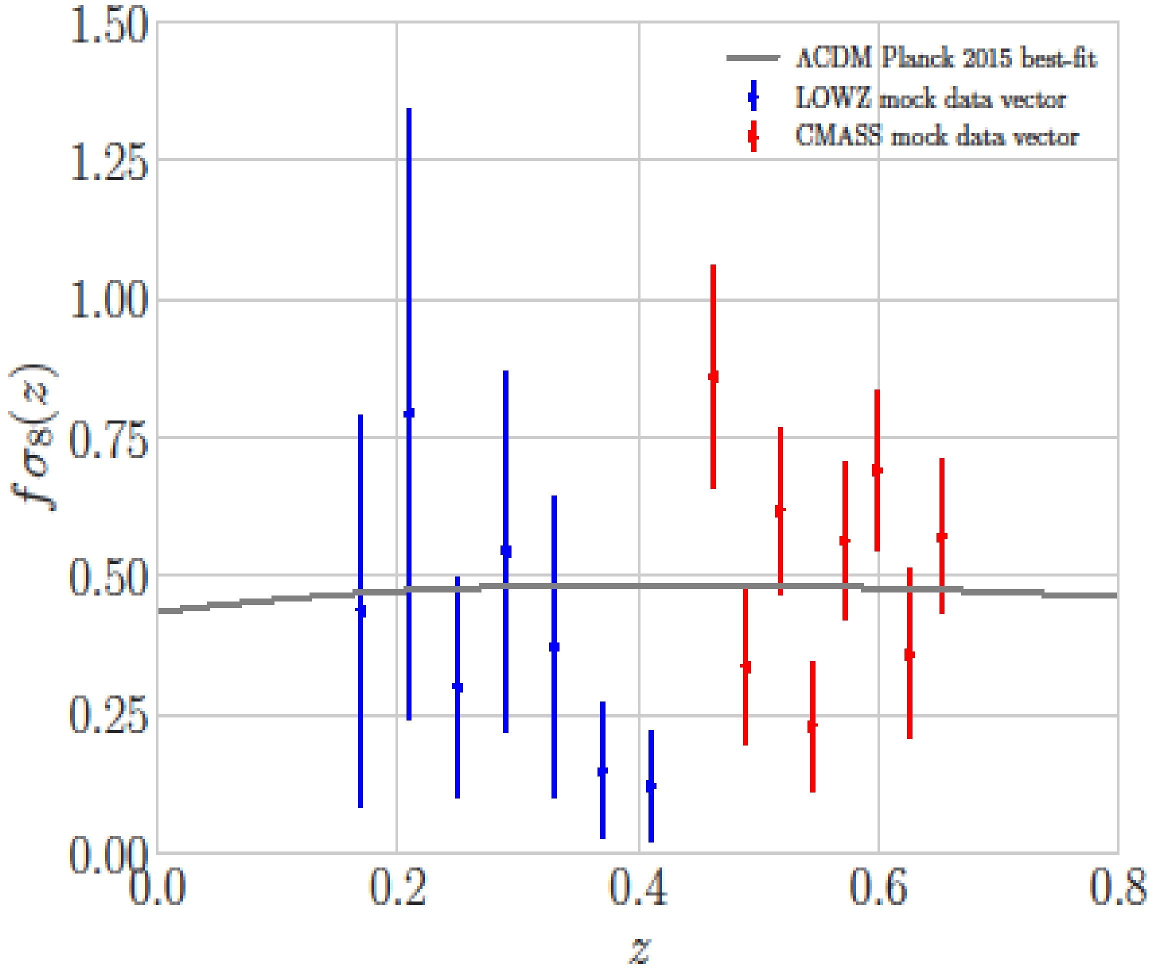
Both methods equivalent

Results: sampling with emcee, equal bin correlations only

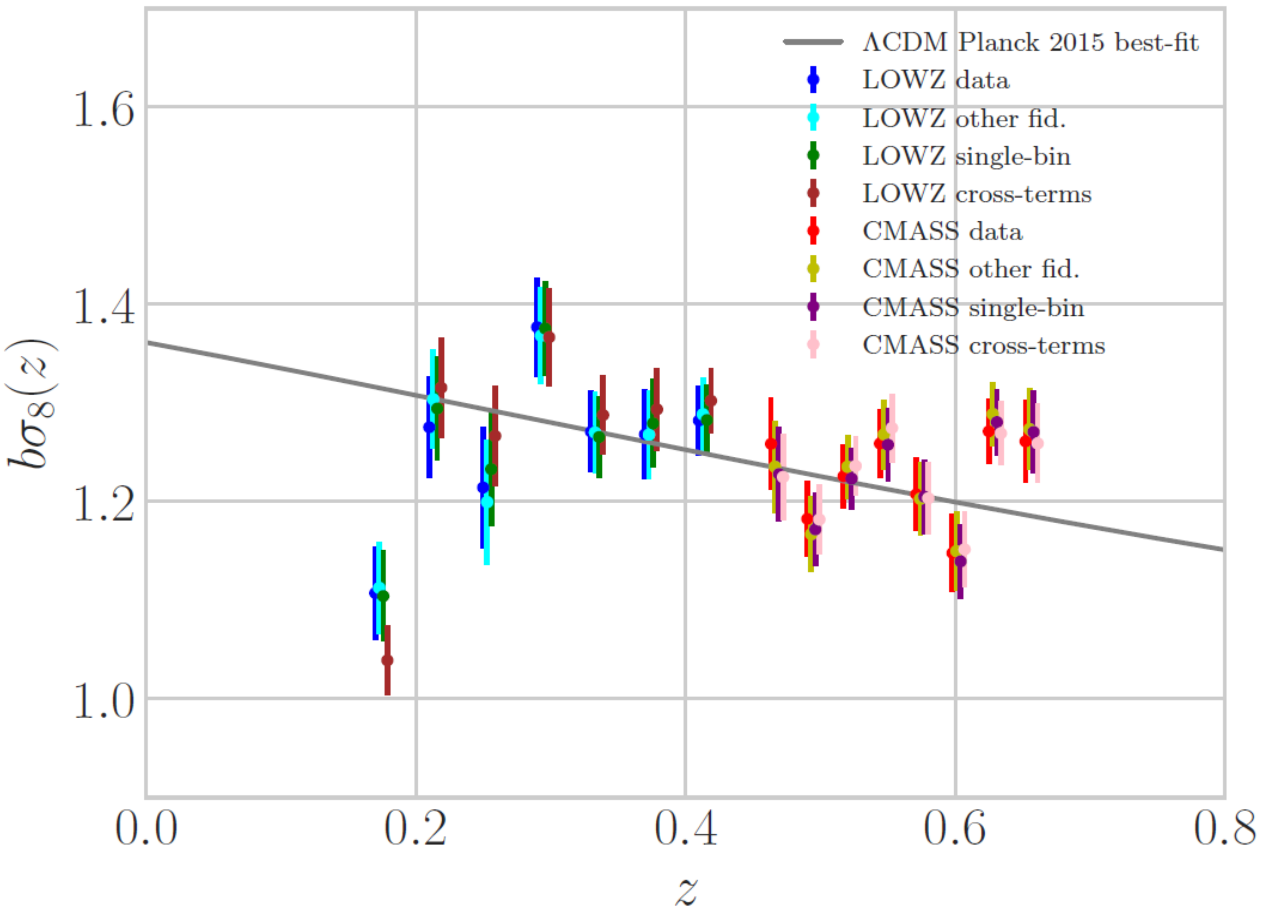


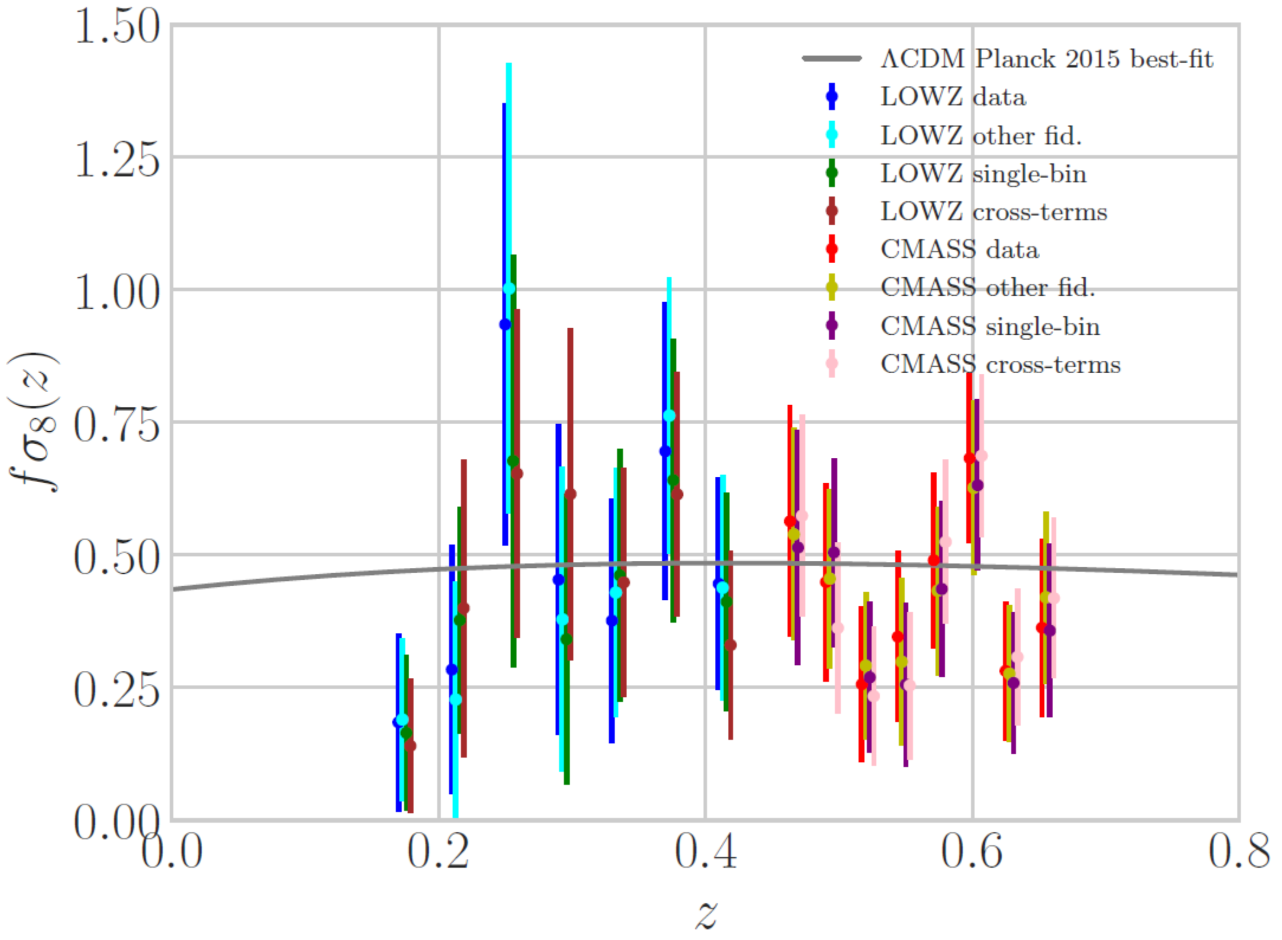
Consistency checks: Replacing data vectors with mock data vectors





Consistency checks: 1) Different fiducial cosmology (change parameters at the 95% C.L of the Planck best-fit values), 2) Single bin estimation, 3) Include cross-bin correlations in the covariance





Consistency checks: Residual distribution

- Diagonalise the covariance matrix:

$$\mathcal{C} = \mathcal{Q}\mathcal{K}\mathcal{Q}^{-1}$$

with $\mathcal{Q}$ the eigenvector matrix, and $\mathcal{K}$ the diagonal eigenvalue matrix as from $\mathcal{C}$

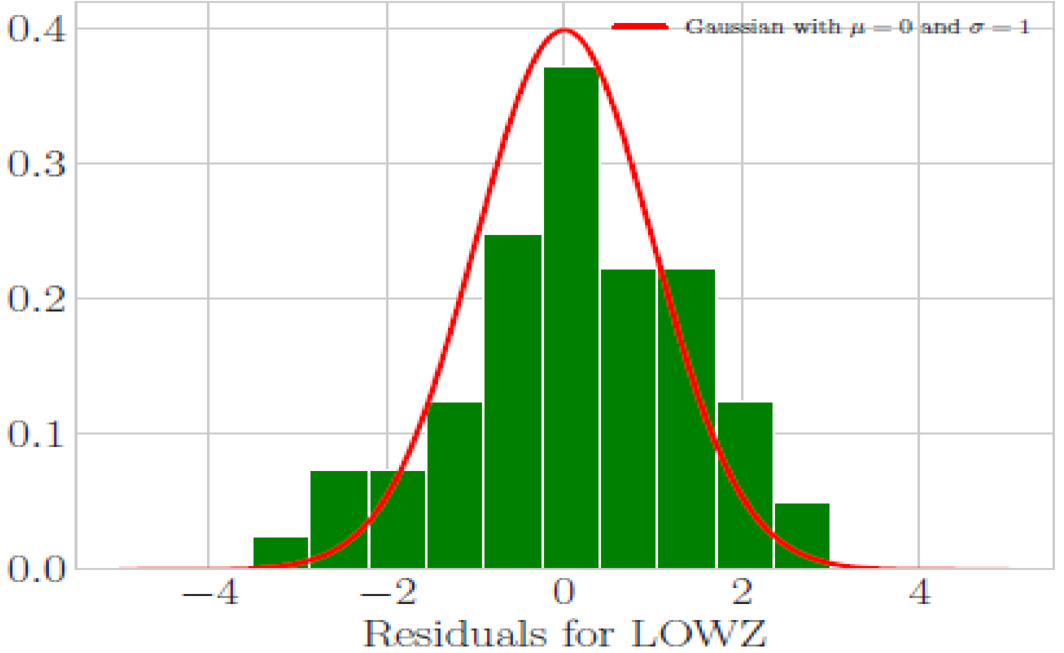
- Now the residual reads:

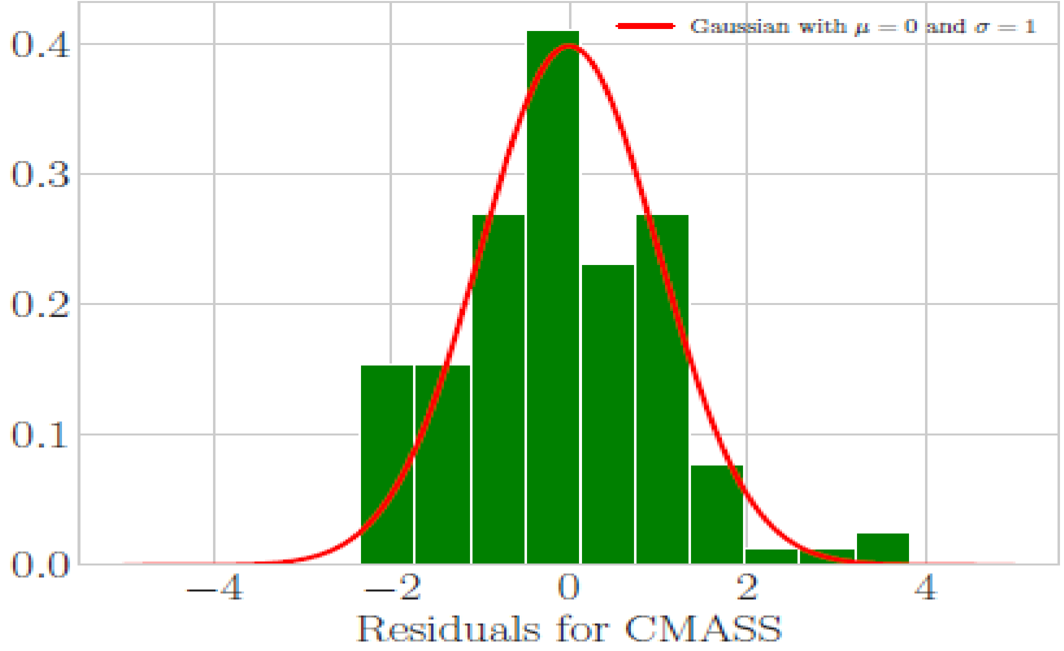
$$R = \mathcal{O}^{-1}\mathcal{Q}^{-1}[\boldsymbol{d} - \boldsymbol{t}(\boldsymbol{\theta}_{best-fit})]$$

and $\mathcal{O} = diag\left(\sqrt{\mathcal{K}_{\alpha\alpha}}\right)$

- If residuals follow the normal distribution $(\mu = 0, \sigma = 1)$ then the model is the 'true' one or the data are not good enough to show model preference, otherwise reject the model

- KS-tests accept null hypothesis





Conclusions

- Novel method to put constraints on the amplitude of the galaxy clustering $\mathrm{b}\sigma_8$ and the growth rate $f\sigma_8$ using the harmonic space (angular) PS.
- Test it against synthetic data and simulations using the BOSS DR12 galaxy subsamples LOWZ and CMASS
- Take into account observational and mask effects in a pseudo-C_ℓ approach
- Construct data covariance with three different implementations (Gaussian, PolSpice, Mocks) yielding consistent results
- Considerable independence from the theory model; passes successfully a series of sanity and systematics tests
- Findings agree very well with literature
- Potential of improving constraints with forthcoming datasets (Euclid, SKA,..) to shed light on the physics of the history of the Universe

Thank you for your attention!