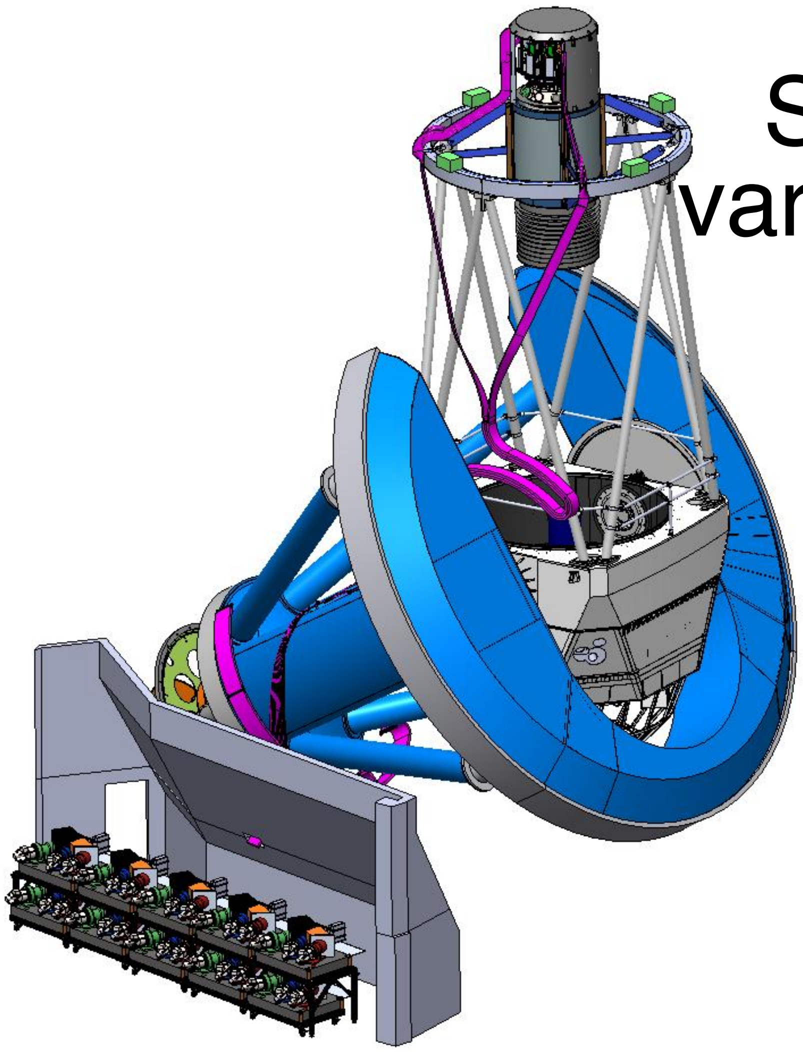


Suppressing the sample variance of DESI-like galaxy clustering with fast simulations

Zhejie Ding, Andrei Variu, Shadab Alam, Yu Yu, Chia-Hsun Chuang

Zhejie Ding
zhejied@sjtu.edu.cn
Postdoc, Shanghai Jiao Tong University
On behalf of the DESI Collaboration

Backgroud and Motivation

- DESI will observe different types of tracers, i.e. BGS, LRG, ELG and QSO. BGS: 0.1<z<0.4; LRG: 0.4<z<1.1; ELG: 0.8<z<1.6; QSO: 0.8<z<2.1

- AbacusSummit N-body simulations are used to model the DESI clustering and to calibrate different systematics.

> the number of simulations is limited, e.g. there are only 25 realizations with the base cosmology.

> DESI Y5 survey volume is huge, e.g. for DESI Y5 ELG sample volume, it is ~24 (Gpc/h)^3 which is 3 times of AbacusSummit simulation box size 8(Gpc/h)^3.

> reducing the sample variance of simulations is a challenge.

> it is very costly to generate more N-body simulation with large volume.

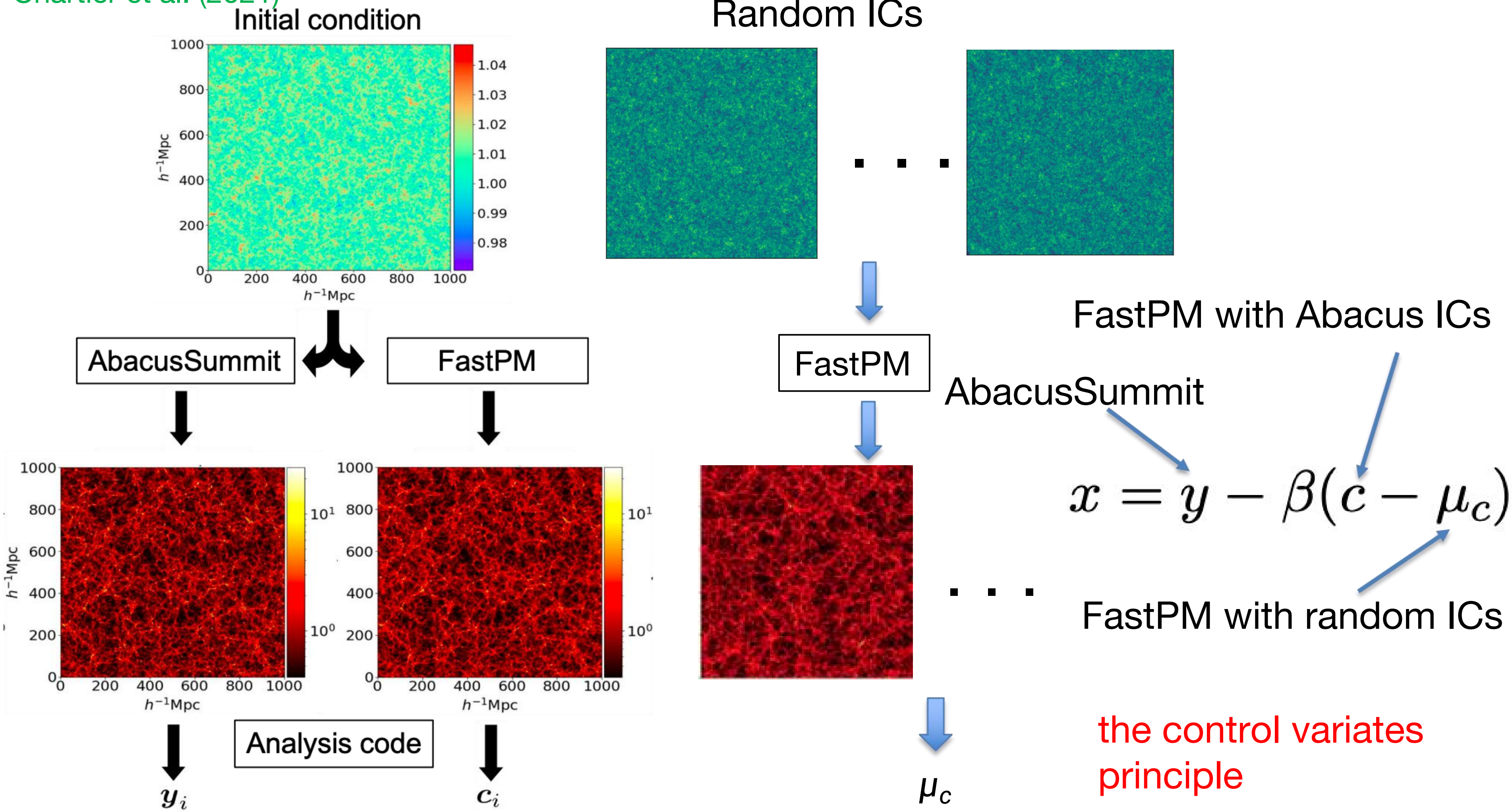
❖ Goal: using fast simulations to reduce the sample variance of galaxy clustering from N-body simulations.

Methods to suppress sample variance

- Fixed-and-paired method (Angulo & Pontzen 2016, Villaescusa-Navarro et al. 2018, Klypin et al. 2020)
- Theoretical control variates (Kokron et al. 2022, DeRose et al. 2023, Hadzhiyska et al. 2023)
- CARPool (Chariter et al. 2021, Chariter et al. 2022, Ding et al. 2022)

Convergence Acceleration by Regression and Pooling (CARPool)

Chartier et al. (2021)



CARPool method

- For a scalar observable, e.g. y, we construct x via

Chartier et al. (2021)

$$x = y - \beta(c - \mu_c),$$

Minimize the variance of x, we can find

$$\beta = \frac{\mathrm{cov}(y,c)}{\sigma_c^2}$$

$$\sigma_x^2 = \sigma_y^2(1-\rho_{y,c}^2), \text{ where } \rho_{y,c} = \mathrm{cov}(y,c)/(\sigma_y\sigma_c)$$

- If observable are vectors, we have

$$\Sigma_{xx}(\beta) = \Sigma_{yy} - \beta\Sigma_{yc}^T - \Sigma_{yc}\beta^T + \beta\Sigma_{cc}\beta^T$$

Minimize $\|\Sigma_{xx}\|$ ⟹ $\beta = \Sigma_{yc}\Sigma_{cc}^{-1}$

With a limitted number of simulation, the inverse of Σ_{cc} may not be stable, hence, we set a simple case:

$$\beta^{\mathrm{diag}} = \begin{pmatrix} \mathrm{cov}(y_1,c_1)/\sigma_{c_1}^2 & 0 & \dots & 0 \\ 0 & \mathrm{cov}(y_2,c_2)/\sigma_{c_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathrm{cov}(y_N,c_N)/\sigma_{c_N}^2 \end{pmatrix}$$

AbacusSummit galaxy mocks

- For the base cosmology Planck2018, AbacusSummit has 25 realizations ph0[00 — 24].
- Box size=2Gpc/h, N_p=6912^3, m_p=2.1e9 Msun/h.
- Galaxy catalogs from HOD; HOD is calibrated based on the DESI SV3 clustering at small scales.
- Typical redshifts for DESI dark-time tracers:

  z=0.8 for LRG; z=1.1 for ELG; z=1.4 for QSO.

	redshift	n $[10^{-4}\, h^3\mathrm{Mpc}^{-3}]$	bias	f
LRG	0.8	10	2.0	0.838
ELG	1.1	30	1.2	0.888
QSO	1.4	1.3	2.0	0.920

Table 1. The redshift, galaxy number density, galaxy bias and growth rate of the cubic AbacusSummit galaxy mocks that are used in this study.

FastPM simulations

- Box size 2Gpc/h, N_p=5184^3, m_p=5.e9 Msun/h.
- In this study, we have used 25 FastPM realizations with AbacusSummit ICs, and 313 realizations with random ICs.
- FoF Halo catalogs are available at the same redshifts of the AbacusSummit.

Ding et al. (2022)

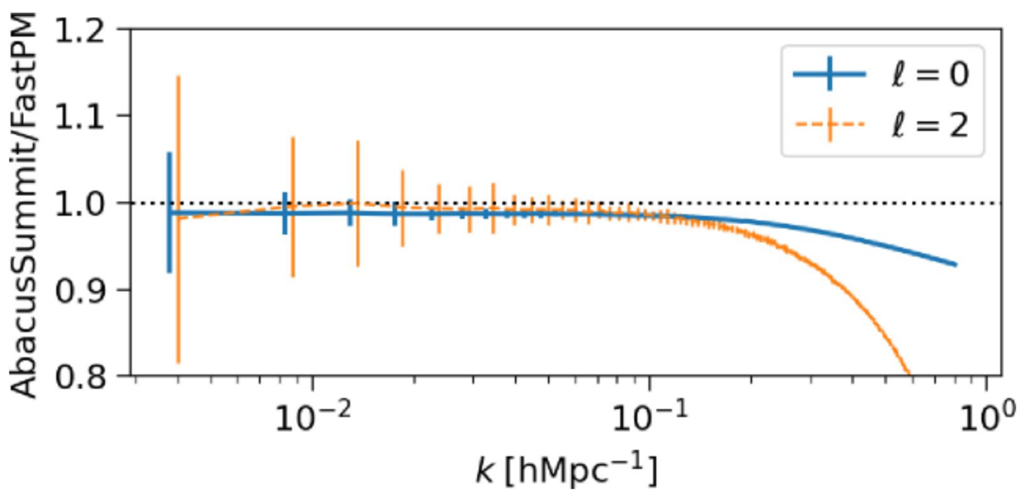
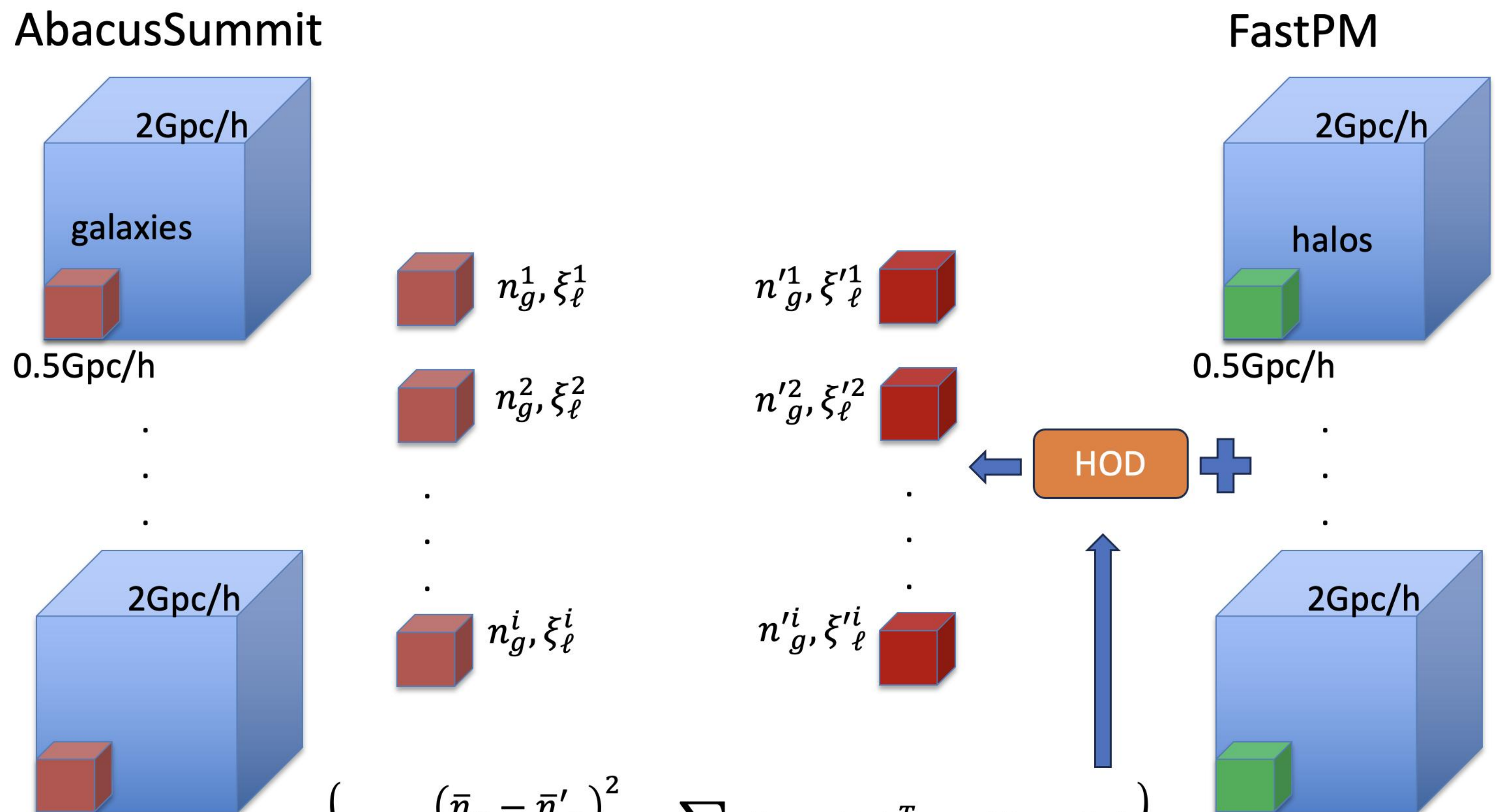


Figure B2. The ratio of the mean power spectrum monopoles ($\ell = 0$) and quadrupoles ($\ell = 2$) between the cleaned ABACUSSUMMIT and FASTPM catalogues with mass cut $10^{11}\ h^{-1}\ \mathrm{M}_{\odot}$. We slightly shift the k coordinates of the quadrupole for clarity.

HOD fitting process



$$Min\left\{\chi^2 = \frac{(\bar{n}_g - \bar{n}'_g)^2}{\sigma^2_{n_g}} + \sum_{\ell=0,2} (\bar{\xi}_\ell - \bar{\xi}'_\ell)^T \Sigma^{-1} (\bar{\xi}_\ell - \bar{\xi}'_\ell)\right\}$$

Based on Andrei Variu's pipeline: https://github.com/Andrei-EPFL/HODOR Variu et al. (2023)

HOD models

- LRG (BGS, QSO):

Zheng et al. (2005)

$$\langle N_{\rm cent}^{\rm LRG}\rangle = \frac{1}{2}\left[1+{\rm erf}\left(\frac{\log M_{\rm h}-\log M_{\rm cut}}{\sqrt{2}\sigma_{\log M_{\rm h}}}\right)\right]$$

$$\langle N_{\rm sat}^{\rm LRG}\rangle = \left(\frac{M_{\rm h}-\kappa M_{\rm cut}}{M_1}\right)^{\alpha}$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt.$$

$$v_{\parallel}^{\rm sat,\ modified} = (v_{\parallel}^{\rm sat,\ default} - v_{\parallel}^{\rm halo}) \times v_{\rm disp} + v_{\parallel}^{\rm halo}$$ (for FastPM HOD)

- ELG

Avila et al. (2020)
Alam et al. (2020)

$$\langle N_{\rm cen}^{\rm ELG}\rangle = 2A\phi(M)\Phi(\gamma M) + \frac{1}{2Q}\left[1+{\rm erf}\left(\frac{\log M_h - \log M_{\rm cut}}{0.01}\right)\right]$$

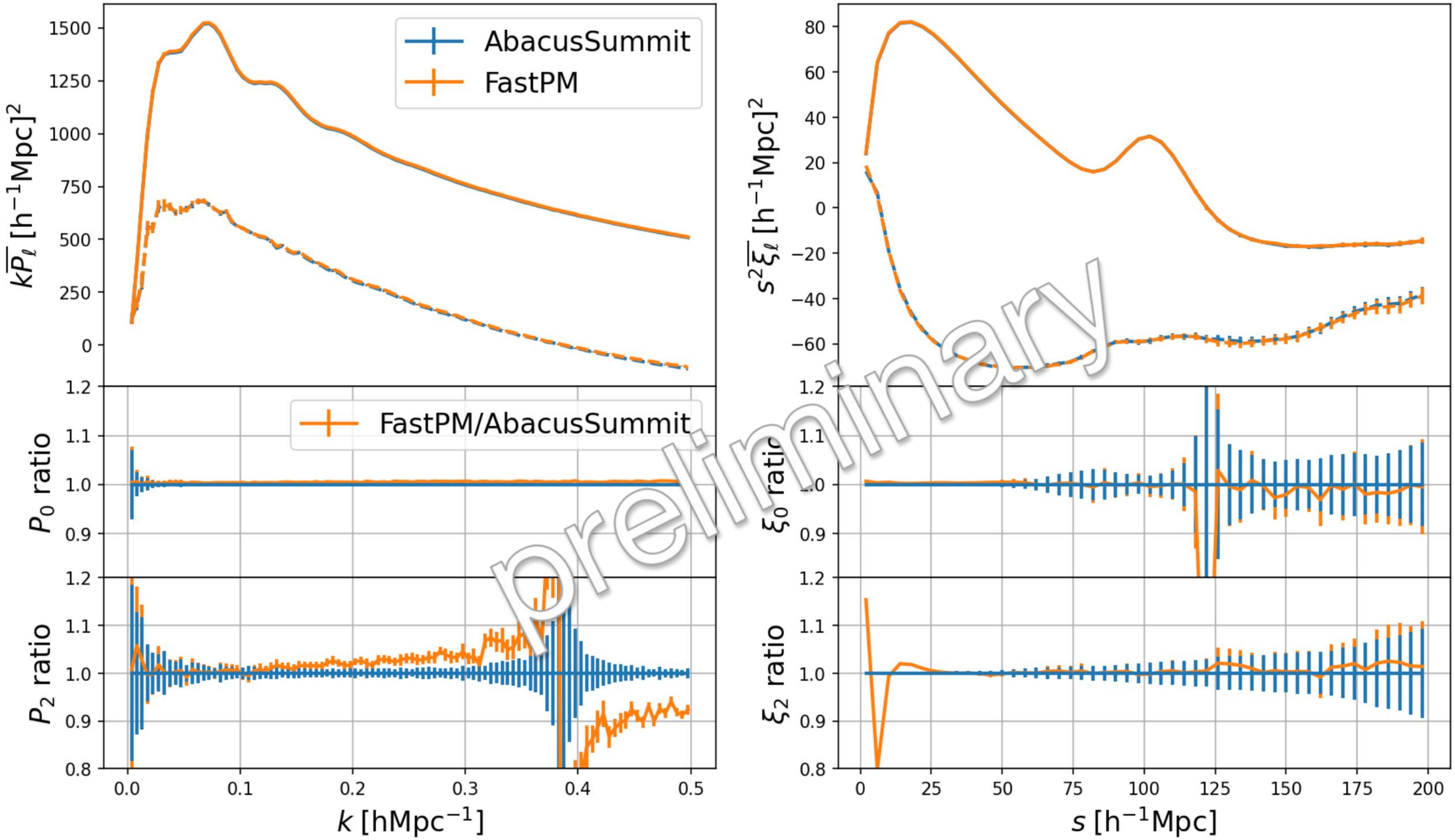
where

$$\phi(x) = \mathcal{N}(\log M_{\rm cut}, \sigma_{\log M_{\rm h}}),$$

$$\Phi(x) = \int_{-\infty}^{x}\phi(x')dx' = \frac{1}{2}\left[1+{\rm erf}\left(\frac{x}{\sqrt{2}}\right)\right],$$

$$A = p_{\rm max} - 1/Q,$$

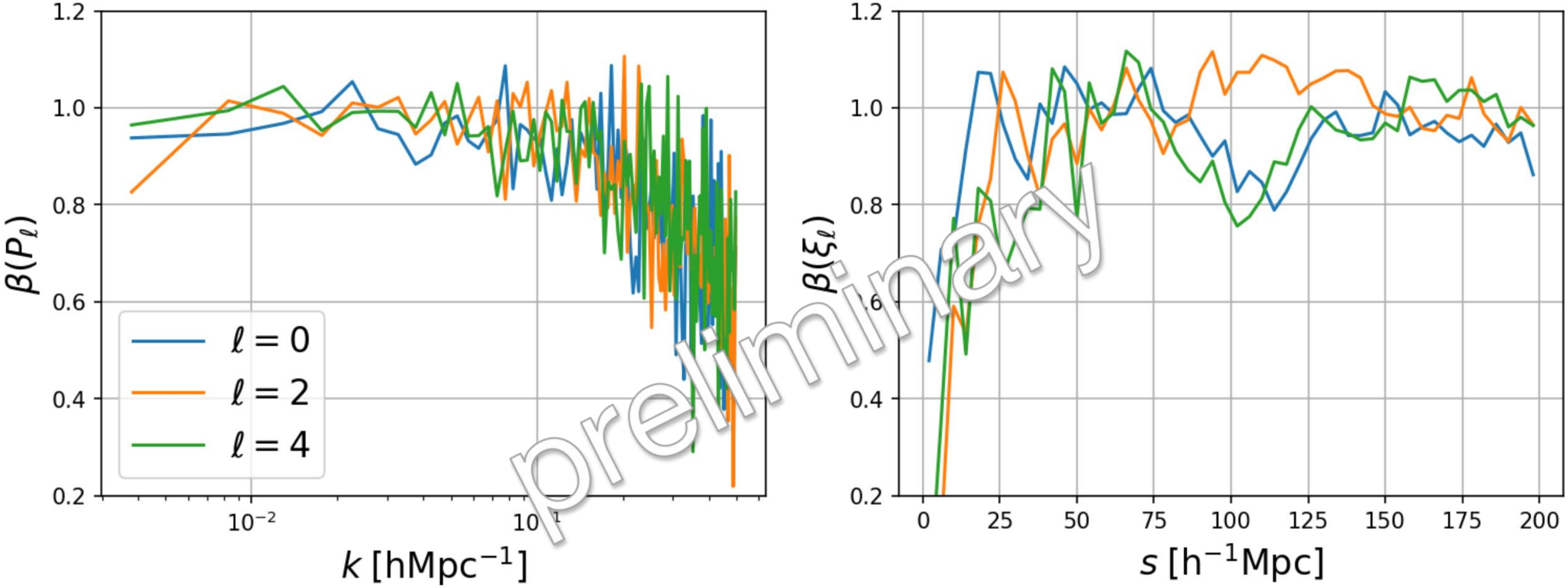
Fitted FastPM LRG clustering



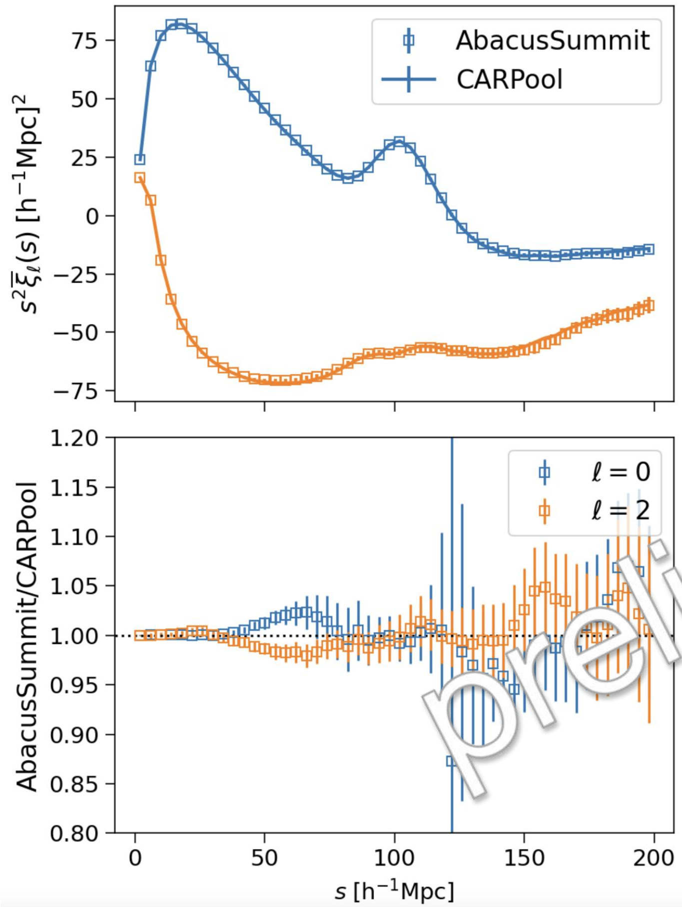
We fit 16 sub-boxes and obtain the fitted HOD parameters.

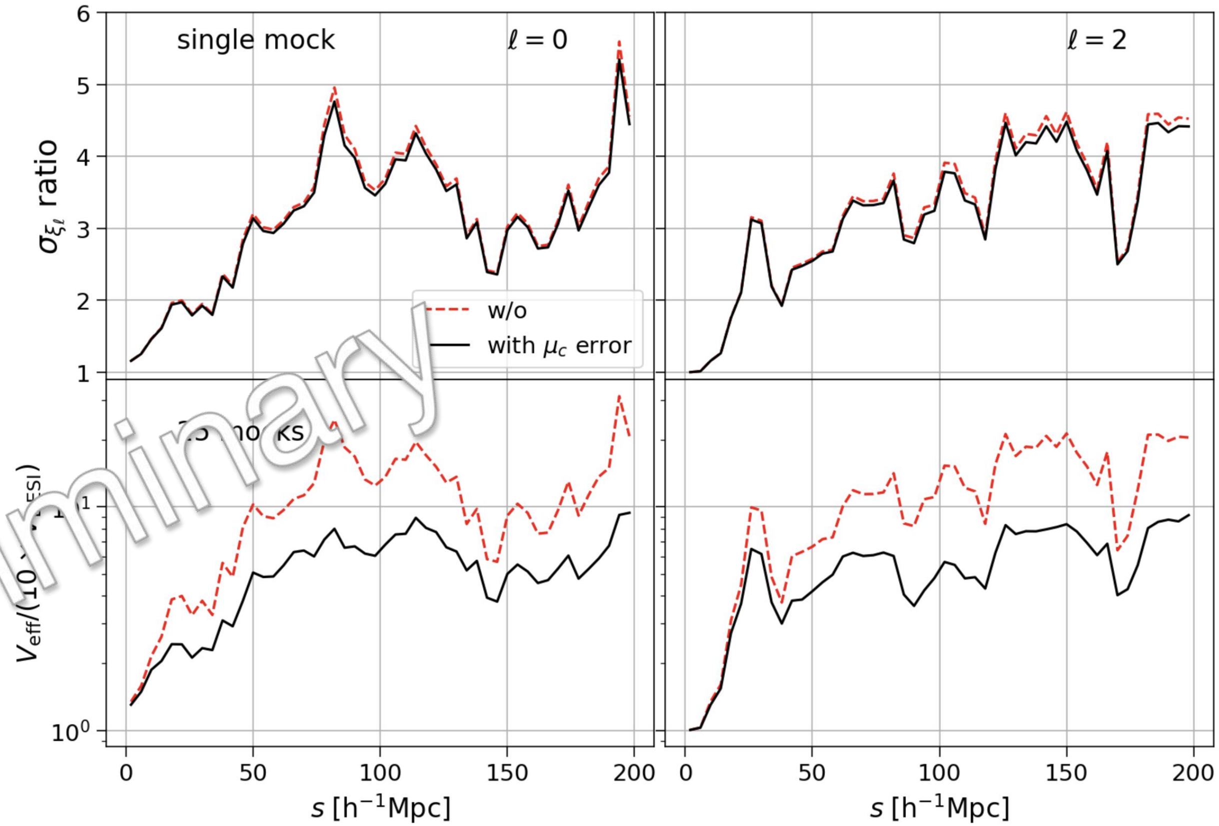
Cross-correlation between AbacusSummit and FastPM galaxy clustering

$$\beta = \frac{\mathrm{cov}(y, c)}{\sigma_c^2}$$



CARPool reduces the sample variance and increases the effective volume of the AbacusSummit clustering

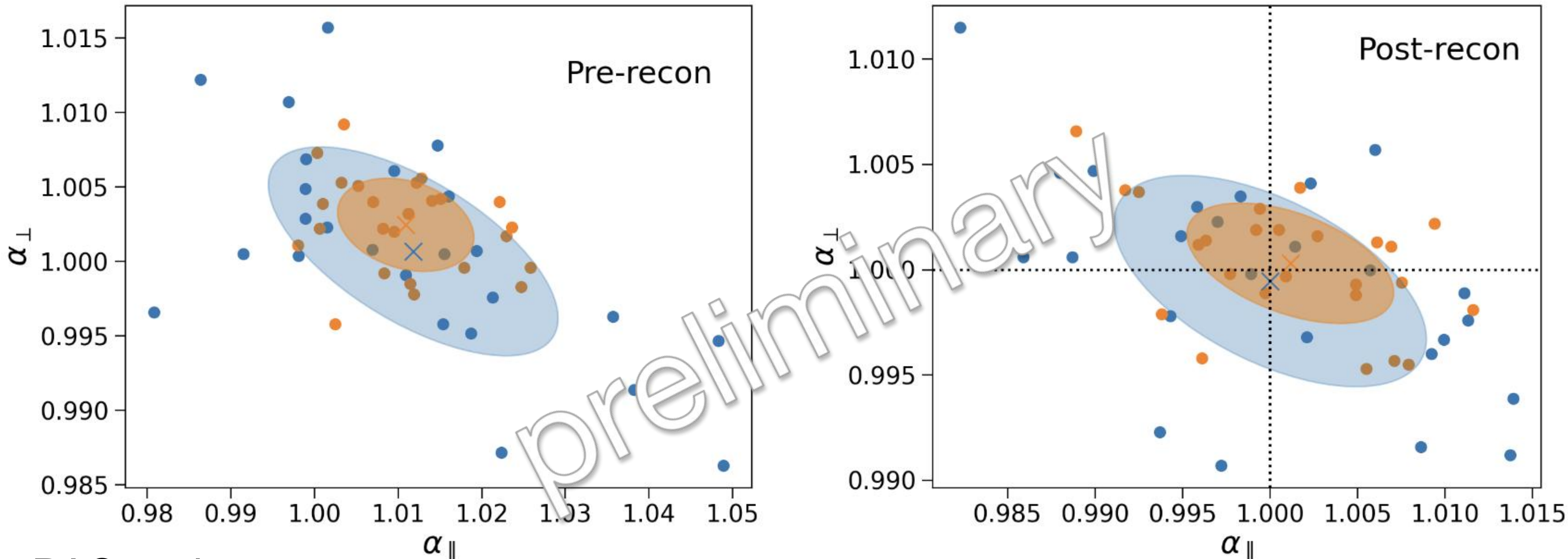




Increasing the BAO constraints with CARPool

For LRG

Blue: pre-CARPool; orange: post-CARPool



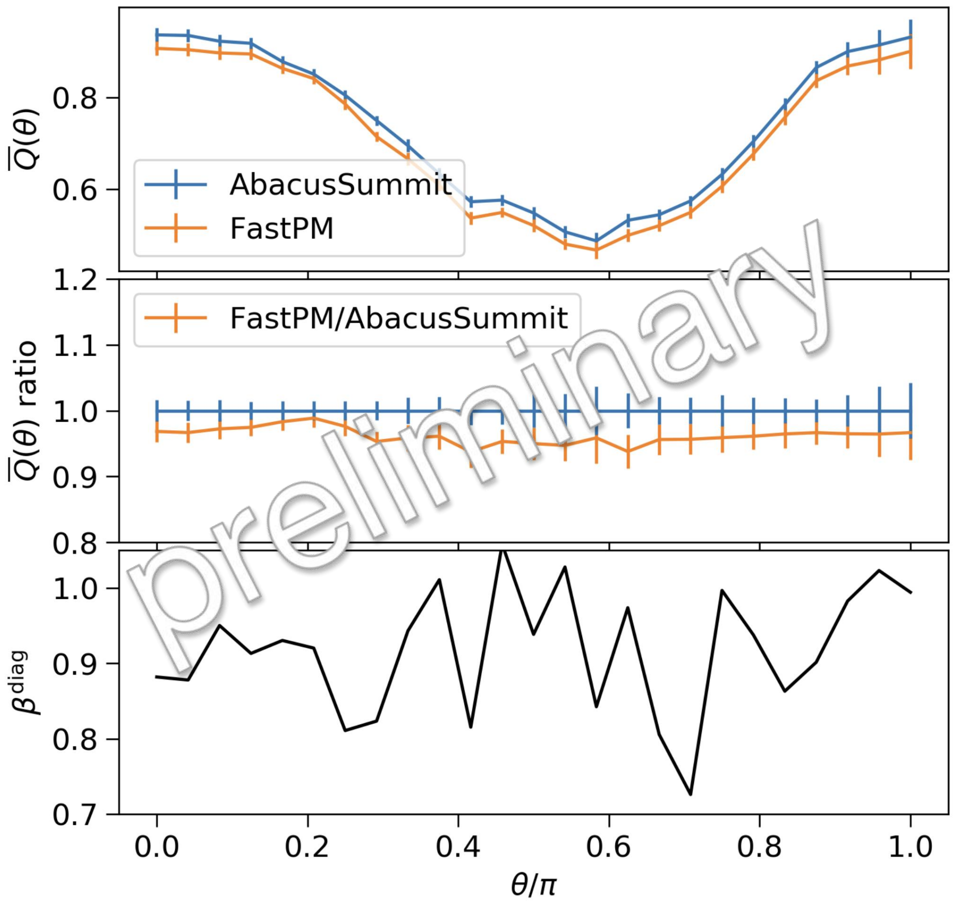
BAO scale parameters:

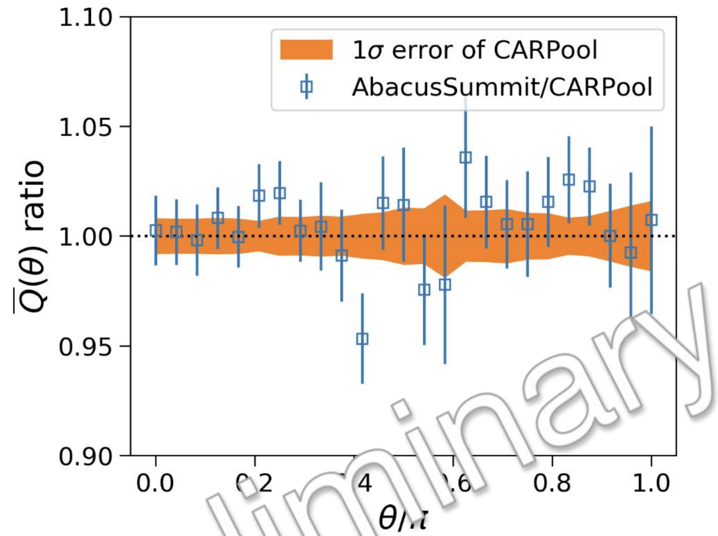
$$\alpha_\parallel = \frac{H^{\rm fid}(z) r_s^{\rm fid}(z_d)}{H(z) r_s(z_d)} \qquad \alpha_\perp = \frac{D_{\rm A}(z) r_s^{\rm fid}(z_d)}{D_{\rm A}^{\rm fid}(z) r_s(z_d)}$$

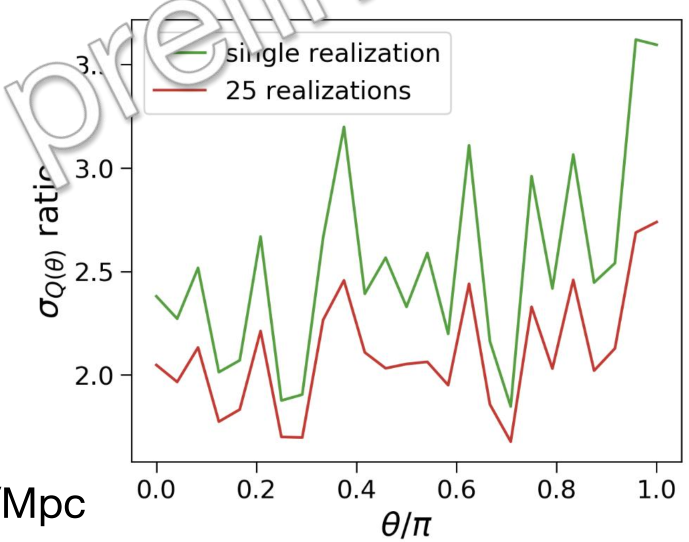
Increase the BAO constraints:

- >2 times per-recon
- >1.5 times post-recon

Reducing the sample variance of bispectrum





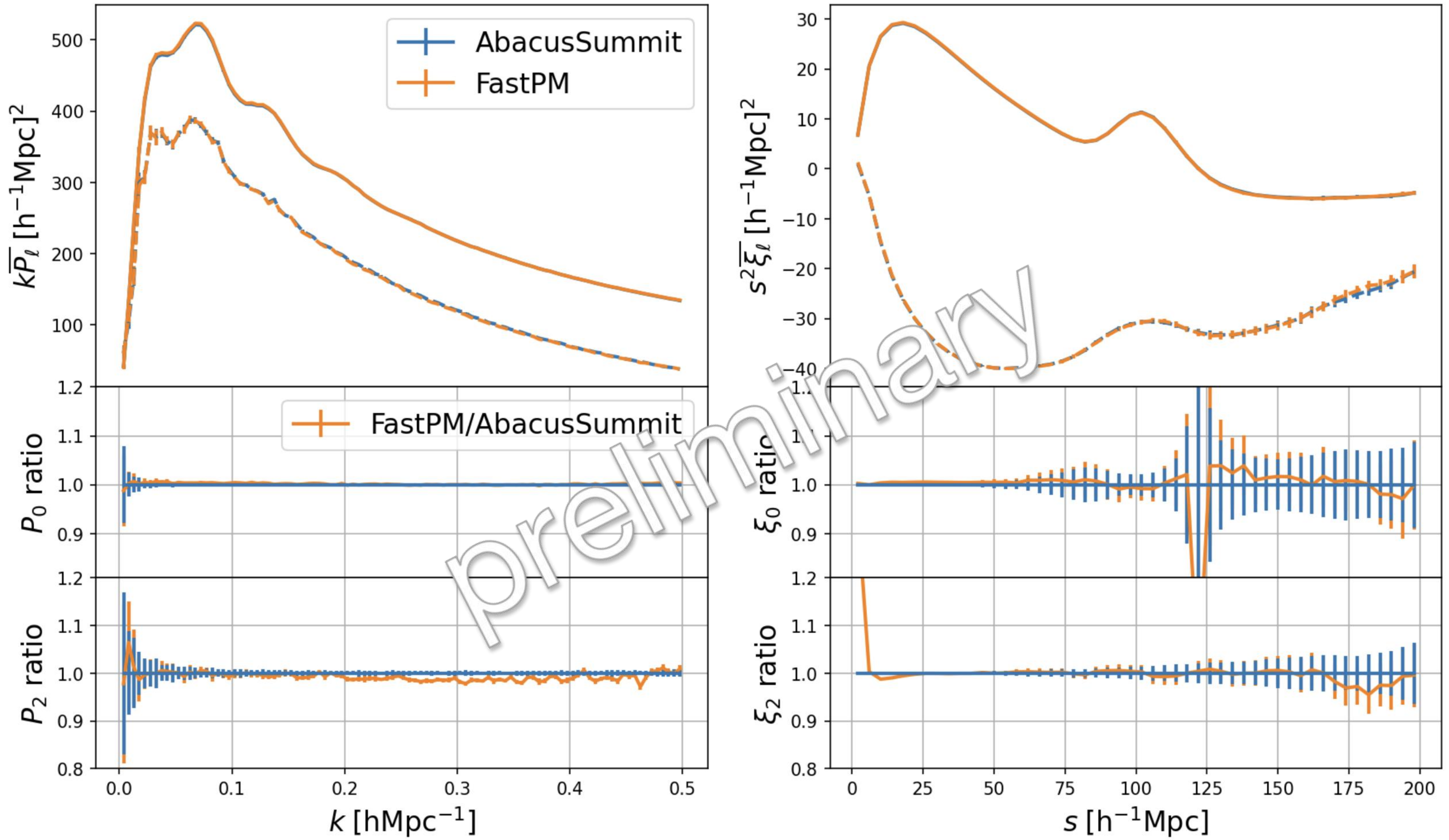


$$\langle \delta(k_1)\delta(k_2)\delta(k_3)\rangle = \delta_D(k_1+k_2+k_3)B(k_1,k_2,k_3),$$

$$Q(\theta) = \frac{B(k_1,k_2,k_3)}{P(k_1)P(k_2)+P(k_1)P(k_3)+P(k_2)P(k_3)}$$

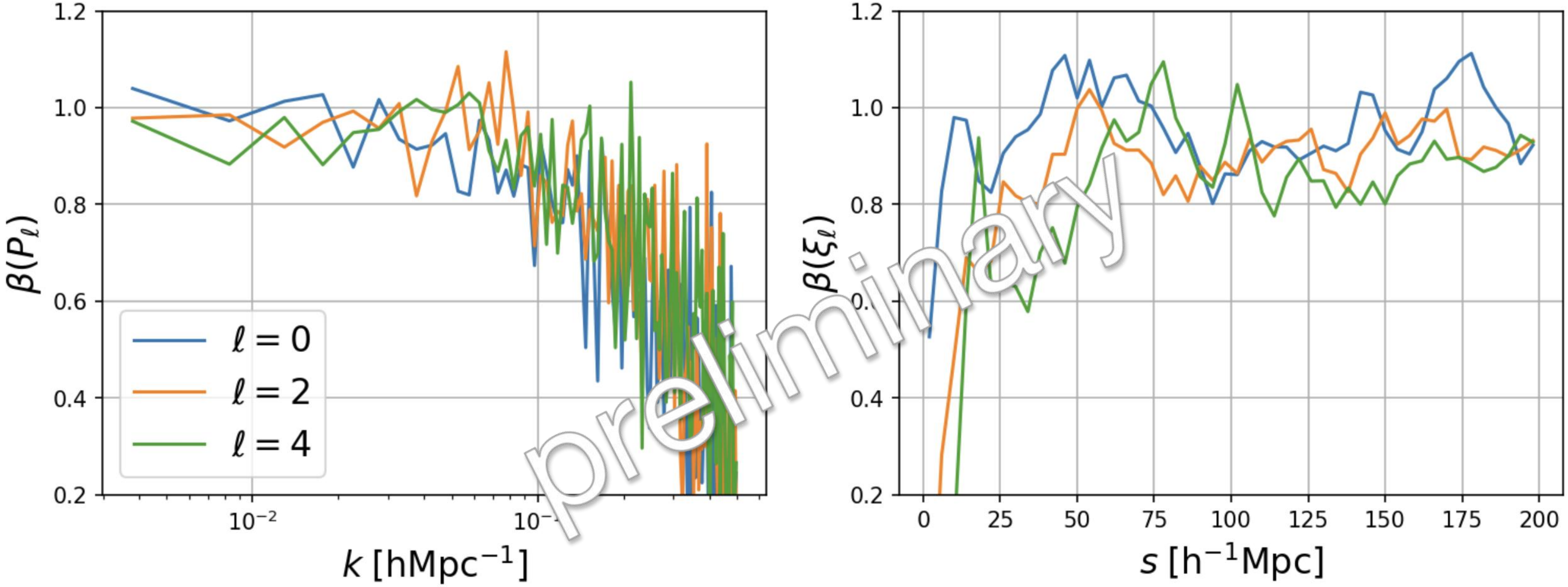
$k_2=2k_1=0.1$ h/Mpc

ELG result



Similar as LRG, we can match FastPM ELG 2pt clustering quite well to that of AbacusSummit.

For ELG


1.2
1.0
0.8
0.6
0.4
0.2
$\beta(P_\ell)$
$\ell = 0$
$\ell = 2$
$\ell = 4$
10^{-2}
k [hMpc^{-1}]
$\beta(\xi_\ell)$
0
25
50
75
100
125
150
175
200
s [h^{-1}Mpc]
preliminary

Conclusions

- Matching to the AbacusSummit n_{gal} and 2pt clustering, we find the best-fit HOD parameters for FastPM halos.
- We demonstrate that the CARPool method can effectively reduce the sample variance of the 2pt and 3pt galaxy clustering from N-body simulations.
- For LRG, we can reduce the standard deviation of correlation function by 3 – 4 times at large scales.
- With CARPool, we can increase the BAO constraints larger than 2 (1.5) times before (after) the density field reconstruction for LRG from AbacusSummit.
- We will perform similar analysis for ELG and QSO.

DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

DEPARTMENT OF ENERGY · UNITED STATES OF AMERICA

NSF

NERSC

cea

GOBIERNO DE ESPAÑA · MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

CONACYT

HEISING-SIMONS FOUNDATION

GORDON AND BETTY MOORE FOUNDATION

Science & Technology Facilities Council

Thanks to our sponsors and 69 Participating Institutions!