arXiv:2404.03008v2 [astro-ph.CO] 5 Apr 2024

HOD-Dependent Systematics for Luminous Red Galaxies in the DESI 2024 BAO Analysis

J. Mena-Fernández,[1,2] **C. Garcia-Quintero**,[3] **S. Yuan**,[4]
B. Hadzhiyska,[5,6] **O. Alves**,[7] **M. Rashkovetskyi**,[8] **H. Seo**,[9]
N. Padmanabhan,[10] **S. Nadathur**,[11] **C. Howlett**,[12] **S. Alam**,[13]
A. Rocher,[14,15] **A. J. Ross**,[16,17,18] **E. Sanchez**,[2] **M. Ishak**,[3]
J. Aguilar,[5] **S. Ahlen**,[19] **U. Andrade**,[20,7] **S. BenZvi**,[21] **D. Brooks**,[22]
E. Burtin,[15] **S. Chen**,[23] **X. Chen**,[10] **T. Claybaugh**,[5] **S. Cole**,[24] **A. de
la Macorra**,[25] **A. de Mattia**,[15] **Arjun Dey**,[26] **B. Dey**,[27] **Z. Ding**,[28]
P. Doel,[22] **K. Fanning**,[29,4] **J. E. Forero-Romero**,[30,31]
E. Gaztañaga,[32,11,33] **H. Gil-Marín**,[34,32,35] **S. Gontcho A Gontcho**,[5]
G. Gutierrez,[36] **J. Guy**,[5] **C. Hahn**,[37] **K. Honscheid**,[16,38,18]
S. Juneau,[26] **A. Kremin**,[5] **M. Landriau**,[5] **L. Le Guillou**,[39]
M. E. Levi,[5] **M. Manera**,[40,41] **P. Martini**,[16,17,18] **L. Medina-Varela**,[3]
A. Meisner,[26] **R. Miquel**,[42,41] **J. Moustakas**,[43] **E. Mueller**,[44]
A. Muñoz-Gutiérrez,[25] **A. D. Myers**,[45] **J. A. Newman**,[27] **J. Nie**,[46]
G. Niz,[47,48] **E. Paillas**,[49,50] **N. Palanque-Delabrouille**,[15,5]
W. J. Percival,[49,51,50] **C. Poppett**,[5,52,6] **A. Pérez-Fernández**,[25,53]
A. Rosado-Marin,[9] **G. Rossi**,[54] **R. Ruggeri**,[55,12] **C. Saulder**,[53]
D. Schlegel,[5] **M. Schubnell**,[56,7] **D. Sprayberry**,[26] **G. Tarlé**,[7]
M. Vargas-Magaña,[25] **B. A. Weaver**,[26] **J. Yu**,[14] **H. Zhang**,[49,50] **and
H. Zou**[46]

Affiliations are in Appendix C

E-mail: juan.menafernandez@lpsc.in2p3.fr

Abstract. In this paper, we present the estimation of systematics related to the halo occupation distribution (HOD) modeling in the baryon acoustic oscillations (BAO) distance measurement of the Dark Energy Spectroscopic Instrument (DESI) 2024 analysis. This paper focuses on the study of HOD systematics for luminous red galaxies (LRG). We consider three different HOD models for LRGs, including the base 5-parameter vanilla model and two extensions to it, that we refer to as baseline and extended models. The baseline model is described by the 5 vanilla HOD parameters, an incompleteness factor and a velocity bias parameter, whereas the extended one also includes a galaxy assembly bias and a satellite profile parameter. We utilize the 25 dark matter simulations available in the `AbacusSummit` simulation suite at $z = 0.8$ and generate mock catalogs for our different HOD models. To test the impact of the HOD modeling in the position of the BAO peak, we run BAO fits for all these sets of simulations and compare the best-fit BAO-scaling parameters α_{iso} and α_{AP} between every pair of HOD models. We do this for both Fourier and configuration spaces independently, using post-reconstruction measurements. We find a 3.3σ detection of HOD systematic for α_{AP} in configuration space with an amplitude of 0.19%. For the other cases, we did not find a 3σ detection, and we decided to compute a conservative estimation of the systematic using the ensemble of shifts between all pairs of HOD models. By doing this, we quote a systematic with an amplitude of 0.07% in α_{iso} for both Fourier and configuration spaces; and of 0.09% in α_{AP} for Fourier space.

Contents

1 Introduction

The large-scale structure of the Universe holds valuable clues about the nature of dark energy, the force driving its accelerated expansion. Baryon acoustic oscillations (BAO) provide a powerful probe for studying this large-scale structure, offering precise measurements of the cosmic distance scale as a function of redshift. The BAO signal arises from primordial density fluctuations that have evolved over cosmic time. In the early Universe, acoustic waves propagated through the primordial plasma, leaving a characteristic scale imprinted on the distribution of matter. These acoustic waves created a preferred

distance scale, known as the sound horizon, which can be measured in the large-scale clustering of galaxies. In this context, the Dark Energy Spectroscopic Instrument (DESI) [1–3] represents a major leap forward in our ability to map the Universe and investigate the imprint of BAO.

Previous and ongoing surveys have explored various cosmological probes to gain insights into dark energy. Stage-III experiments, combining efforts like Planck [4], Pantheon [5], the Sloan Digital Sky Survey (SDSS) [6], and the Dark Energy Survey (DES) [7], have significantly advanced our understanding, achieving measurements of cosmological parameters with unprecedented precision. In the pursuit of answers about the nature of dark energy, researchers are now moving towards the next generation of experiments, or Stage-IV experiments such as DESI, which are expected to further enhance the figure of merit for dark energy studies. DESI is a next-generation spectroscopic survey designed to map millions of galaxies and quasars over a large cosmological volume, and it is the only Stage-IV dark energy experiment that is currently taking data.

DESI has been meticulously designed to perform, over 5 years, a comprehensive survey covering, approximately, 14,000 square degrees of the sky, with regions in both the south and the north galactic caps [8]. During its operations, DESI aims to measure the redshifts of, approximately, 40 million galaxies, ranging from 0.05 to 3.5 (up to 2.1 for clustering analysis, and up to 3.5 for Lyman-α forest analysis). DESI has already completed its survey validation stage [9] and released its early data to the public [10]. The DESI target selection program categorizes its tracers into four distinct types within these redshift ranges. In increasing order of redshift, these are: bright galaxies (observed during the Bright Galaxy Survey, or BGS) [11], luminous red galaxies (LRG) [12], emission line galaxies (ELG) [13] and quasars (QSO) [14] (see the references for the description of their corresponding target selections).

Systematics are a major concern in the analyses of large galaxy surveys, but more in particular when we enter the very-high-precision era of Stage-IV surveys. Systematic errors can arise from a variety of sources, including instrumental effects, calibration uncertainties and astrophysical effects, and they can impact the accuracy and precision of the measured BAO signal [15]. DESI is designed to mitigate systematics through different techniques, including careful instrument calibration, robust data reduction and analysis procedures, and the use of multiple complementary probes of cosmology. Given the exquisite precision achievable by DESI, it is crucial to have all sources of systematics well under control.

In this paper, we study the effect of systematics in the measurement of the BAO signal related to the modeling of the galaxy-halo connection [16]. In particular, we use halo occupation distribution (HOD) [17] models to describe this galaxy-halo connection, which yields to the study of what we refer to as HOD systematics. In DESI, we use algorithms designed to minimize the impact of non-linear scales in the measurement of the BAO signal, and what we aim to do here is to test if this is sufficient for viable HOD models, i.e., we want to test how well the fitting model can adapt to a given realistic galaxy bias model[1]. Any difference between these two is a systematic error. In the pre-DESI era, similar studies were carried out for SDSS, in particular in [18] for LRGs, in [19–21] for ELGs and in [22] for QSO.

The work presented here is part of the DESI 2024 analysis, which corresponds to the analysis of the first year of DESI data. Here, we focus on the study of HOD systematics in the BAO measurement for the LRG tracer, whereas in [23] we address the case of ELGs. A similar study is presented in [24] for the full-shape analysis. The catalogs and the data release are described in [25]. The clustering measurements for the different tracers can be found in [26]. The measurement of the BAO signal using galaxies and quasars is performed in [27], where the systematics estimated in this work are accounted

---

[1]It is worth mentioning here that the systematics quoted in this paper are obtained using a specific BAO-fitting methodology. If a different algorithm or BAO model was used, the results could slightly differ.

for in the calculation of the final error budget. On the other hand, the measurement of the BAO using the Lyman-α forest is carried out in [28]. The analysis of the full-shape measurements is performed in [29]. Finally, the cosmological constraints are described in three papers: in [30] using BAO, in [31] using full-shape measurements, and in [32] we obtain constraints on primordial non-Gaussianities.

Other sources of systematics in the measurement of the BAO signal and optimal settings for the different codes used in our analyses are studied in several companion papers. The optimal settings for the density-field reconstruction of the catalogs are investigated in [33], whereas [34] shows tests to unblinded simulations and data catalogs with these optimal settings. Different tests to validate the analytical covariance matrices are performed in [35]. [36] estimates the systematic error budget due to theory and modeling. Potential systematic errors due to the assumption of the fiducial cosmology are studied in [37]. Finally, [38] performs consistency tests between the LRG and ELG tracers in the overlapping redshift range.

The structure of the paper is as follows: in Section 2, we summarize the statistical uncertainties obtained for the BAO distance measurement in the context of the DESI 2024 analysis; in Section 3, we describe the methodology used to perform the clustering measurements, the reconstruction of the catalogs and the BAO-fitting code; in Section 4, we introduce the HOD models considered in this paper, together with their optimization against the DESI's One-Percent Survey data; in Section 5, we describe the generation of the `AbacusSummit` simulations used for the analysis; in Section 6, we describe the control variates (or CV) technique, which we use to obtain noise-reduced clustering measurements; in Section 7, we show the BAO-fit results and the estimate of the amplitude of the HOD systematics; and in Section 8, we show our conclusions to this analysis.

Throughout this paper, we adopt the Planck-2018 flat ΛCDM cosmology, specifically the mean estimates of the Planck TT,TE,EE+lowE+lensing likelihood chains shown in Table 2 of [4]: $\Omega_c h^2 = 0.1200$, $\Omega_b h^2 = 0.02237$, $h = 0.6736$, $\sigma_8 = 0.8111$, $n_s = 0.9649$, $w_0 = -1$, and $w_a = 0$.

2 Statistical Uncertainties for DESI 2024

In this section, we summarize the statistical uncertainties obtained in the measurement of the BAO signal using the LRG tracer in the context of the DESI 2024 analysis (see [27] for a full description of the measurements). These are displayed in Table 1, where we include the results for each redshift bin and for the different BAO scaling parameters considered in this paper: α_{iso}, α_{AP}, $\alpha_{\parallel}$ and $\alpha_{\perp}$ (see Table 2 for further details on them), for both Fourier and configuration spaces[2]. We also include the combined results, which were obtained from the individual redshift bins as

$$\sigma = \left(\sum_{\text{zbin}} \frac{1}{\sigma_{\text{zbin}}^2} \right)^{-1/2}. \tag{2.1}$$

We obtain a redshift-combined statistical uncertainty at the level of 0.5%-1.9%, depending on the parameter, which makes it extremely important to carefully account for any kind of systematics. Besides the systematics related to the HOD modeling that we describe in this paper, we also consider the contribution to the final error bars of our measurements of several additional sources of systematics, e.g., other theoretical systematics related to the non-linear modeling, such as the mode-coupling terms, see [36] for further details. Thus, the effect of each source of systematics has to be small enough not to inflate significantly our final error when combined with all the others.

---

[2]The DESI 2024 analysis is performed in both Fourier and configuration spaces independently. Therefore, the analysis carried out in this paper is also split into these two.

$z_{\min}$	$z_{\max}$	Fourier space				Configuration space			
		$\sigma_{\mathrm{stat,iso}}$	$\sigma_{\mathrm{stat,AP}}$	$\sigma_{\mathrm{stat,\parallel}}$	$\sigma_{\mathrm{stat,\perp}}$	$\sigma_{\mathrm{stat,iso}}$	$\sigma_{\mathrm{stat,AP}}$	$\sigma_{\mathrm{stat,\parallel}}$	$\sigma_{\mathrm{stat,\perp}}$
0.4	0.6	0.011	0.033	0.023	0.018	0.011	0.034	0.023	0.018
0.6	0.8	0.009	0.033	0.022	0.014	0.011	0.040	0.027	0.017
0.8	1.1	0.008	0.025	0.018	0.012	0.009	0.029	0.021	0.013
Combined		**0.005**	**0.017**	**0.012**	**0.008**	**0.006**	**0.019**	**0.013**	**0.009**

Table 1. Measured statistical errors for the DESI 2024 BAO analysis, taken from [27]. The table is split into Fourier and configuration spaces. We include the results for the BAO-scaling parameters α_{iso}, α_{AP}, $\alpha_{\parallel}$, and $\alpha_{\perp}$ (see Table 2 for further details on these parameters), for each redshift bin individually, and also for the combined case using Eq. (2.1).

We remark here that, since the Fourier and configuration space analyses are performed independently, all the results quoted in this paper use a different statistical error for these two analyses, as displayed in Table 1.

3 Methods

In this section, we describe the algorithms and codes used to compute the clustering measurements, in both Fourier and configuration spaces; the ones used to run the reconstruction of the catalogs; and the ones used to perform the BAO fits.

3.1 Clustering Estimators and Reconstruction

As mentioned earlier, we perform analyses in both Fourier and configuration spaces. In terms of Fourier space measurements, we adopt the periodic box power spectrum estimator introduced by [39] within the DESI package `pypower`[3]. If we denote the density contrast as $\delta_{\mathrm{gal}}(\boldsymbol{r}) = n_{\mathrm{gal}}(\boldsymbol{r})/\bar{n}_{\mathrm{gal}} - 1$, then the power spectrum multipoles can be calculated as

$$P_\ell(k) = \frac{2\ell+1}{V} \int \frac{d\Omega_k}{4\pi} \delta_{\mathrm{gal}}(\boldsymbol{k})\delta_{\mathrm{gal}}(-\boldsymbol{k})\mathcal{L}_\ell(\hat{\boldsymbol{k}}\cdot\hat{\eta}) - P_\ell^{\mathrm{shot-noise}}(k). \tag{3.1}$$

Here, V denotes the volume of the box, $\boldsymbol{k}$ is the wave-number vector, $\mathcal{L}_\ell$ are the Legendre polynomials and η is the line-of-sight vector. The shot-noise term is only subtracted for the monopole. We employ a triangular-shaped cloud prescription to interpolate the density field on 512^3 meshes. We originally use bins of size $\Delta k = 0.001\ h/\mathrm{Mpc}$, beginning at $k = 0\ h/\mathrm{Mpc}$, and later re-bin the measurements to $\Delta k = 0.005\ h/\mathrm{Mpc}$.

In the context of configuration space, we evaluate the galaxy two-point correlation function (2PCF) using the approach described in [40]. However, the RR (random-random) pair counts can be deduced via analytical means, given that we use periodic cubic boxes in this analysis. Consequently, we use the natural estimator, which is expressed as

$$\xi(s,\mu) = \frac{DD(s,\mu) - RR(s,\mu)}{RR(s,\mu)}. \tag{3.2}$$

In the previous equation, DD and RR represent the count of galaxy pairs and random pairs, respectively. These pair-counts are dependent on the distance between galaxies, denoted as s, and the cosine

---

[3] https://github.com/cosmodesi/pypower

of the angle formed between the line of sight[4] and the galaxy pair, denoted as μ. To compute the 2PCF, we make use of the publicly available DESI package `pycorr`[5].

To enhance the detection of the BAO in terms of both precision and significance, we use a technique known as density-field reconstruction [41]. Due to the growth of structures during the evolution of the Universe, the observed position x of a galaxy deviates from its initial position q in Lagrangian space by a displacement field $\Psi(x)$,

$$x = q + \Psi(x). \tag{3.3}$$

This displacement shifts and smears the BAO signal, biasing the redshift-distance relation and leading to a degradation in the significance of its detection. The objective of density-field reconstruction is to bring galaxies back to the position where they would be in the absence of such non-linear effects. Reconstruction achieves this by computing the displacement field based on an estimate of the galaxy velocity field, derived from the galaxy density field. The first-order term in a perturbative expansion of $\Psi(x)$ using Lagrangian perturbation theory, known as the Zeldovich approximation [42], can be expressed as

$$\Psi(x) = \int \frac{d^3 k}{(2\pi)^3} \frac{i\boldsymbol{k}}{k^2} \delta(\boldsymbol{k}) \exp\left(i\boldsymbol{k} \cdot \boldsymbol{q}\right). \tag{3.4}$$

Additionally, the displacement field is related to the observed redshift-space galaxy field δ_{gal} through the linear continuity equation [43, 44],

$$\boldsymbol{\nabla} \cdot \boldsymbol{\Psi} + \frac{f}{B} \boldsymbol{\nabla} \cdot \left[(\boldsymbol{\Psi} \cdot \hat{\boldsymbol{r}}) \hat{\boldsymbol{r}} \right] = -\frac{\delta_{\mathrm{gal}}}{B}, \tag{3.5}$$

where f is the linear growth rate, $\hat{\boldsymbol{r}}$ is the line-of-sight direction, and B is the linear galaxy bias.

Various approaches exist for solving Eq. (3.5). For example, [44] employs a grid-based solution in configuration space, calculating gradients using finite difference approximations; [45] assumes that $(\boldsymbol{\Psi} \cdot \hat{\boldsymbol{r}}) \hat{\boldsymbol{r}}$ is an irrotational field and employs an iterative fast Fourier transform (FFT) scheme for efficient displacement computation; and [46] utilizes a relaxation technique with a full multi-grid V-cycle based on a damped Jacobi iteration. We run the reconstruction using the `pyrecon`[6] tool, which includes some of these implementations. The one employed in our case is `MultiGrid`[7], see [46] for further details on the code. Throughout the reconstruction process, we assume a linear galaxy bias of $B = 1.99$ for our LRG tracer. The optimal settings used for reconstruction are detailed in [33].

The 2PCF for post-reconstruction measurements is determined using the Landy-Szalay estimator, since we need to include the shifted randoms (that come from the reconstruction process) in the calculation. The Landy-Szalay estimator for post-reconstruction datasets is given by

$$\xi(s, \mu) = \frac{DD(s, \mu) - 2DS(s, \mu) + SS(s, \mu)}{RR(s, \mu)}. \tag{3.6}$$

Here, SS represents the count of galaxy pairs in the shifted random catalog. For both pre- and post-reconstruction measurements, we convert the 2PCF into correlation function multipoles using the formula

$$\xi_\ell(s) = \frac{2\ell + 1}{2} \int_{-1}^{1} d\mu \, \xi(s, \mu) \mathcal{L}_\ell(\mu), \tag{3.7}$$

---

[4]For the box simulations used in our analysis, which are described in Section 5, the line of sight is considered as the z coordinate.

[5]https://github.com/cosmodesi/pycorr

[6]https://github.com/cosmodesi/pyrecon

[7]https://github.com/martinjameswhite/recon_code

where ℓ takes the values 0, 2, and 4. The hexadecapole is excluded from the BAO fits due to its low signal-to-noise ratio. We calculate $\xi(s, \mu)$ for equidistant bins of 1 Mpc/h width for the separation s, ranging from 0 Mpc/h to 200 Mpc/h (200 bins in total). For μ, the range is from -1 to 1 (240 bins in total). However, for the BAO fits, we ultimately re-bin the results for $\xi_\ell(s)$ with a bin size of 4 Mpc/h.

The procedures described here match the methods adopted for the final DESI 2024 results.

3.2 BAO Fits

Here, we summarize the methodology used to run the BAO fits, for both Fourier and configuration spaces. In Section 3.2.1, we describe the template-based modeling used; in Section 3.2.2, how the BAO chains are run; and in Section 3.2.3, how the covariance matrices are estimated.

3.2.1 The BAO-Fitting Model

We conduct an analysis to extract information about the BAO feature in an anisotropic way. Our approach involves an improved version of the model described in [47], which is presented and tested in [36]. The BAO template is constructed as follows:

1. The process begins with splitting the linear matter power spectrum into a "wiggle" and a "no-wiggle" part, denoted as $P_w(k)$ and $P_{nw}(k)$ respectively. To do this, we use the method as described in [48]. $P_w(k)$ encapsulates all the oscillatory behavior of the matter power spectrum.

2. A second step consists of constructing a template for the broadband of the *galactic* power spectrum. First, the random peculiar velocities of galaxies on small scales cause an elongation of positions in redshift space, resulting in a damping of the power spectrum due to the non-linear velocity field. This effect is parameterized by a factor $\mathcal{D}_{FoG}$, explicitly given by

$$\mathcal{D}_{FoG}(k, \mu) = \frac{1}{(1 + k^2 \mu^2 \Sigma_s^2/2)^2}, \tag{3.8}$$

where Σ_s is the damping scale parameter for the Fingers-of-God (FoG) term.

3. Second, terms are introduced to account for galaxy bias, represented by B, and the coherent infall of galaxies at large scales, represented by $(1 + \beta\mu^2)^2$, see [49]. β is one of the free parameters of our model, and it is defined as

$$\beta = f/B. \tag{3.9}$$

4. To address the potential washout of the BAO feature due to the growth of non-linear structures separated by a distance close to the BAO scale, an additional damping factor, $C(k, \mu)$[8], given by

$$C(k, \mu) = \exp\left[-k^2\left(\frac{(1 - \mu^2)\Sigma_{nl,\perp}^2}{2} + \frac{\mu^2 \Sigma_{nl,\parallel}^2}{2}\right)\right], \tag{3.10}$$

is introduced. The parameters $\Sigma_{nl,\parallel}$ and $\Sigma_{nl,\perp}$ in Eq. (3.10) account for damping along and across the line of sight, respectively. These two damping parameters, together with Σ_s, are allowed to vary within reasonably tight Gaussian priors (2 Mpc/h), while their central values are obtained from fits to the means of the simulations, as presented in [36].

---

[8]We use the observed k and μ for the RSD factor within $C(k, \mu)$ instead of the true k' and μ', which is slightly different from [36] and [27], but equally acceptable. [36] finds that this makes a negligible difference in the BAO-fit results.

The resulting power spectrum, including peculiar velocity effects, galaxy bias, and non-linear damping, is given by

$$P(k, \mu) = B^2(1 + \beta\mu^2)^2 \mathcal{D}_{\text{FoG}}(k, \mu) \left[P_{\text{nw}}(k) + P_{\text{w}}(k)C(k, \mu) \right]. \tag{3.11}$$

5. Finally, we also include BAO-scaling parameters $\alpha_\parallel$ and $\alpha_\perp$, which are given by

$$\alpha_\parallel(z) = \left[H(z)r_d \right]^{-1} \left[H^{\text{ref}}(z)r_d^{\text{ref}} \right] \tag{3.12}$$

and

$$\alpha_\perp(z) = \left[\frac{D_M(z)}{r_d} \right] \left[\frac{D_M^{\text{ref}}(z)}{r_d^{\text{ref}}} \right]^{-1}, \tag{3.13}$$

respectively, to account for coordinate dilation effects related to redshift measurements and assumptions about the cosmological model. $H(z)$ is the Hubble parameter, $D_M(z)$ is the comoving angular diameter distance and r_d is the sound horizon scale. The superscript "ref" refers to quantities evaluated at the reference cosmology used for the template. These $\alpha_\parallel$ and $\alpha_\perp$ parameters are used to transform the observed wave numbers and line-of-sight coordinates to the true values, hereafter denoted with a prime ($\prime$). The scaling of the wave numbers across and along the line of sight ($k_\perp$ and $k_\parallel$) as a function of these two parameters is, explicitly, given by

$$k_\perp \rightarrow k_\perp' = k_\perp/\alpha_\perp, \quad k_\parallel \rightarrow k_\parallel' = k_\parallel/\alpha_\parallel. \tag{3.14}$$

Following the previous procedure, for the analysis in Fourier space we calculate the power spectrum multipoles, $P_\ell(k)$, using the Legendre polynomials, $\mathcal{L}_\ell$, as

$$P_\ell(k) = \frac{2\ell + 1}{2} \int_{-1}^{1} d\mu \mathcal{L}_\ell(\mu) P[k'(k, \mu), \mu'(\mu)], \tag{3.15}$$

where $P(k, \mu)$ is given by Eq. (3.11). The template power spectrum for the BAO fits is, then, defined as

$$P_\ell^{\text{temp}}(k) = P_\ell(k) + \sum_{i=-1}^{4} A_\ell^{(i+1)} k^i. \tag{3.16}$$

The explicit expressions for k' and μ' as a function of k and μ are given by

$$k'(k, \mu) = \sqrt{k_\perp'^2 + k_\parallel'^2} = \frac{k}{\alpha_\perp} \sqrt{1 + \mu^2 \left(\frac{\alpha_\perp^2}{\alpha_\parallel^2} - 1 \right)} \tag{3.17}$$

and

$$\mu'(\mu) = \frac{k_\parallel'}{k'(k, \mu)} = \frac{\mu}{\frac{\alpha_\parallel}{\alpha_\perp} \sqrt{1 + \mu^2 \left(\frac{\alpha_\perp^2}{\alpha_\parallel^2} - 1 \right)}}, \tag{3.18}$$

respectively. In configuration space, the power spectrum multipoles are transformed into the correlation function ones, $\xi_\ell(s)$, using the spherical Bessel functions, $j_\ell(x)$, as

$$\xi_\ell(s) = i^\ell \int_0^\infty dk \frac{k^2}{2\pi^2} P_\ell(k) j_\ell(ks), \tag{3.19}$$

where $P_\ell(k)$ is given by Eq. (3.15). The template correlation function for the BAO fits is, then, given by

$$\xi_\ell^{\text{temp}}(s) = \xi_\ell(s) + \sum_{i=-1}^{2} A_\ell^{(i+1)} s^{-i}. \tag{3.20}$$

In Table 2 we summarize the different parameters related to BAO fitting used in this work, together with their priors and some derived parameters. The BAO fits will be run using the code Barry[9], which utilizes α_{iso} and ϵ as the free BAO scaling parameters, instead of the $\alpha_\parallel$ and $\alpha_\perp$ ones that we introduced previously. However, we also compute $\alpha_\parallel$ and $\alpha_\perp$ in each step of the chain using

$$\begin{cases} \alpha_\parallel = \alpha_{\text{iso}}(1 + \epsilon)^2, \\ \alpha_\perp = \dfrac{\alpha_{\text{iso}}}{1 + \epsilon}. \end{cases} \tag{3.21}$$

The final results for the HOD systematics will be quoted in terms of α_{iso} and α_{AP}, i.e., α_{AP} instead of ϵ. These two are related by

$$\alpha_{\text{AP}} = (1 + \epsilon)^3 = \frac{\alpha_\parallel}{\alpha_\perp}. \tag{3.22}$$

Using Eqs. (3.12) and (3.13), we can express α_{iso} and α_{AP} as a function of $H(z)$ and $D_M(z)$ via Eqs. (3.21) and (3.22). We find

$$\alpha_{\text{iso}}(z) = \left[\alpha_\parallel(z)\alpha_\perp^2(z)\right]^{1/3} = \left[\frac{D_V(z)}{r_d}\right]\left[\frac{D_V^{\text{ref}}(z)}{r_d^{\text{ref}}}\right]^{-1}, \tag{3.23}$$

where $D_V(z)$ is given by

$$D_V(z) \equiv \left[\frac{zD_M^2(z)}{H(z)}\right]^{1/3}, \tag{3.24}$$

and

$$\alpha_{\text{AP}}(z) = \frac{\alpha_\parallel(z)}{\alpha_\perp(z)} = \left[D_M(z)H(z)\right]^{-1}\left[D_M^{\text{ref}}(z)H^{\text{ref}}(z)\right], \tag{3.25}$$

respectively.

We note that, due to the parallel nature in which the two studies were performed, our BAO modeling procedure differs slightly from that recommended by [36] in that 1) we allow the FoG damping term to affect both the wiggle and no-wiggle terms; 2) we use a "polynomial"-based broadband method instead of their preferred "spline"-based method and 3) we allow the BAO dilation parameters to affect both the wiggle and no-wiggle model components. However, they conclude that the BAO-fitting methodology is robust to any of these choices, at the 0.1% level for α_{iso} and 0.2% level for α_{AP}. In any case, as demonstrated later, our method is also able to recover unbiased BAO constraints, and because we focus on comparative differences between different simulations in this work, our results are immune to these choices.

3.2.2 Running the BAO Chains

For both Fourier and configuration space analyses, we run a Markov chain Monte Carlo (MCMC) to fit the BAO template to the clustering measurements of our simulations, which are described in Section 5. We use the code Barry to run our BAO chains, see [50] for further details on the code.

---

[9]https://github.com/Samreay/Barry

Parameter	Description	Prior
$\alpha_{\rm iso}$	Isotropic shift in the BAO scale.	$(0.8, 1.0)$
ϵ	Anisotropic warping of the BAO signal.	$(-0.2, 0.2)$
$\Sigma_{\rm nl, \parallel}$	Non-linear damping of the BAO feature along the line of sight.	$\begin{cases} \mathcal{N}(9.6, 2.0)\ \text{(pre-recon)} \\ \mathcal{N}(5.4, 2.0)\ \text{(post-recon)} \end{cases}$
$\Sigma_{\rm nl, \perp}$	Non-linear damping of the BAO feature across the line of sight.	$\begin{cases} \mathcal{N}(5.0, 2.0)\ \text{(pre-recon)} \\ \mathcal{N}(1.8, 2.0)\ \text{(post-recon)} \end{cases}$
$\Sigma_{\rm s}$	Fingers-of-God parameter for the velocity dispersion (Lorentzian form).	$\mathcal{N}(0.0, 2.0)$
β	Kaiser term parameter ($\beta = f/B$).	$(0.01, 4.0)$
B	Linear galaxy bias. Obtained by analytical marginalization.	$(0.1, 10.0)$
$A_\ell^{(i)}$	Broadband-term parameters. Obtained by analytical marginalization.	$(-20000.0, 20000.0)$
$\alpha_{\rm AP}$	Alcock-Paczynski scale distortion parameter. Derived as $\alpha_{\rm AP} = (1 + \epsilon)^3$.	
$\alpha_\parallel$	BAO scaling parameter along the line of sight. Derived from $\alpha_{\rm iso}$ and ϵ.	
$\alpha_\perp$	BAO scaling parameter across the line of sight. Derived from $\alpha_{\rm iso}$ and ϵ.	

Table 2. Description of the parameters related to BAO fitting used in this paper. The BAO template parameters are described in the first few entries, along with the priors used during the fitting stage. The parameters without priors below the horizontal line are derived parameters. The values used as centrals for the Gaussian priors on $\Sigma_{\rm nl, \parallel}$, $\Sigma_{\rm nl, \perp}$ and $\Sigma_{\rm s}$ were optimized in [36].

We explore the parameter space using the nested sampling algorithm DYNESTY, see [51] for reference. The χ^2 to be minimized is given by

$$\chi^2 = (M - D)^T \, \text{COV}^{-1} \, (M - D) \,, \tag{3.26}$$

where M is given by Eq. (3.16) for Fourier space and Eq. (3.20) for configuration space. D represents the corresponding data vector, which can be either the correlation function multipoles or the power spectrum ones. COV is the covariance matrix, which includes both covariances between scales (i.e., different k/s bins for Fourier/configuration spaces), but also cross-covariances between multipoles. We rely on analytical estimations of the covariance matrices, as described in Section 3.2.3.

In Table 3 we display the settings used for Barry to run the BAO fits in this paper. We include the edges, bin sizes, multipoles fitted, cosmology for the template, and the number of broadband-term parameters, for both Fourier and configuration spaces.

Setting	Fourier space - $P_\ell(k)$	Configuration space - $\xi_\ell(s)$
Edges	$k \in [0.02, 0.3]$ h/Mpc	$s \in [50, 150]$ Mpc/h
Bin size	$\Delta k = 0.005$ h/Mpc	$\Delta s = 4$ Mpc/h
Multipoles	$\ell = 0,\ 2$	$\ell = 0,\ 2$
Cosmology	Planck-2018	Planck-2018
# of bb-term par.	$i = -1, 0, 1, 2, 3, 4$	$i = -1, 0, 1, 2$

Table 3. Default settings to perform the BAO fits with `Barry` in Fourier and configuration spaces. The scales for the fits are well within the limits in which we computed the clustering measurements; however, the original clustering measurements were re-binned (for both Fourier and configuration spaces). As mentioned in the text, Planck-2018 corresponds to flat ΛCDM with $\Omega_c h^2 = 0.1200$, $\Omega_b h^2 = 0.02237$, $h = 0.6736$, $\sigma_8 = 0.8111$, $n_s = 0.9649$, $w_0 = -1$, and $w_a = 0$.

3.2.3 Covariance Matrix

In order to run the BAO fits we just described, we need accurate estimations for the covariance matrices for both $\xi_\ell(s)$ and $P_\ell(k)$. In this work, we rely on analytical covariance matrices, whose amplitudes are tuned to the clustering of the 25 `AbacusSummit` realizations for each HOD model. These covariances are validated in [35], where their accuracy is compared to those computed from mock simulations.

1. In the Fourier space case, we estimate Gaussian analytical covariances using the code `thecov`[10] [52, 53]. The covariance is calculated as

$$C_{\ell_1\ell_2}(k_i, k_j) = \frac{\delta_{ij}}{N_{\mathrm{modes}}(k_i)} \sum_{\ell_3, \ell_4} L_{\ell_1\ell_2}^{\ell_3\ell_4} P_{\ell_3}(k_i) P_{\ell_4}(k_j), \tag{3.27}$$

where

$$N_{\mathrm{modes}}(k_i) = \frac{1}{2} \frac{L_{\mathrm{box}}^3}{(2\pi)^3} \frac{4\pi}{3} \left[(k_i + \Delta k/2)^3 - (k_i - \Delta k/2)^3 \right] \tag{3.28}$$

and

$$L_{\ell_1\ell_2}^{\ell_3\ell_4} = \frac{(2\ell_1 + 1)(2\ell_2 + 1)}{2} \int_{-1}^{1} d\mu\, \mathcal{L}_{\ell_1}(\mu) \mathcal{L}_{\ell_2}(\mu) \mathcal{L}_{\ell_3}(\mu) \mathcal{L}_{\ell_4}(\mu). \tag{3.29}$$

The monopole, $P_0(k)$, includes the shot-noise contribution. We generate one covariance matrix for each HOD model considered in this paper. The input $P_\ell(k)$ used in the computation is the average of the power spectrum multipoles measured from each of the 25 `AbacusSummit` realizations described in Section 5 (and $L_{\mathrm{box}} = 2$ Gpc/h is the size of these cubic simulations). The code and its usage for the DESI 2024 analysis are validated in [54].

2. In the configuration space case, we use the code `RascalC`[11] (following the Legendre multipole covariance procedure described in [55]; see also [56, 57] for further details on the code). We generate one covariance matrix tuned to the pre-reconstruction clustering of HOD A0 (that we use for HODs A0-A3), another one tuned to HOD B0 (that we use for HODs B0-B3) and another one tuned to FirstGen (that we use for FirstGen). All these different HOD models are described in Section 5. Besides these covariance matrices, we also generate another set for the

---

[10]https://github.com/cosmodesi/thecov
[11]https://github.com/oliverphilcox/RascalC/tree/v2.2.1

post-reconstruction case. The code and its usage for the DESI 2024 analysis are validated in [58].

The same methodology described here is applied to obtain the final DESI 2024 covariance matrices.

4 The Galaxy-Halo Connection

In this section, we introduce the formalism used to establish the connection between galaxies and dark matter halos, the so-called galaxy-halo connection.

4.1 HOD Models for LRGs

To propagate the simulated matter density field to galaxy distributions, we adopt the Halo Occupation Distribution (HOD) model, which probabilistically populates dark matter halos with galaxies according to a set of halo properties. The HOD can be summarized as a probabilistic distribution $P(n_{\mathrm{gal}}|X_h)$, where n_{gal} is the number of galaxies of the given halo, and X_h is some set of halo properties.

In the vanilla HOD model, the halo mass is assumed to be the only relevant halo property, i.e., $X_h = \{M_h\}$ [59, 60]. In the vanilla scheme, galaxies are separated into centrals and satellites. For LRGs, the mean central and satellite HODs are well approximated by

$$\bar{n}_{\mathrm{cent}}^{\mathrm{LRG}}(M_h) = \frac{f_{\mathrm{ic}}}{2}\mathrm{erfc}\left[\frac{\log_{10}(M_{\mathrm{cut}}/M_h)}{\sqrt{2}\sigma}\right],\tag{4.1}$$

$$\bar{n}_{\mathrm{sat}}^{\mathrm{LRG}}(M_h) = \left[\frac{M_h - \kappa M_{\mathrm{cut}}}{M_1}\right]^{\alpha}\bar{n}_{\mathrm{cent}}^{\mathrm{LRG}}(M_h),\tag{4.2}$$

where the five vanilla parameters are $M_{\mathrm{cut}}, M_1, \sigma, \alpha, \kappa$ (the previous expressions are originally shown in [60]). M_{cut} sets the minimum halo mass to host a central galaxy. M_1 roughly sets the typical halo mass that hosts one satellite galaxy. σ controls the steepness of the transition from 0 to 1 in the number of central galaxies. α is the power law index on the number of satellite galaxies. κM_{cut} gives the minimum halo mass to host a satellite galaxy. We have also included an incompleteness parameter f_{ic}, which is a downsampling factor controlling the overall number density of the mock galaxies. This parameter is relevant when trying to match the observed mean density of the galaxies in addition to clustering measurements. By definition, $0 < f_{\mathrm{ic}} \leq 1$.

We use the `AbacusHOD` package to implement our HODs [61]. In `AbacusHOD`, once the mean occupation is determined, the actual number of central galaxies is sampled from a Bernoulli distribution, whereas the satellites are sampled from a Poisson distribution. Additionally, the position and velocity of the central galaxies are set to be the same as those of the halo center, whereas the satellite galaxies are randomly assigned to halo particles with uniform weights, each satellite inheriting the position and velocity of its host particle.

In our fiducial HOD model, which we refer to as the baseline model, we also have a velocity bias, which essentially modifies the vanilla velocities by introducing a parametrized bias to the velocities of the central and satellite galaxies relative to their respective host halos and particles. The central galaxy velocity is modified relative to the host halo velocity, which is computed as the mean velocity of the halo parties, whereas the satellite velocity is modified relative to the dark matter particle, see [61] for further details. This is a necessary ingredient in modeling the LRG redshift-space clustering (see [62] for justification and detailed description). The two parameters governing velocity bias are the central velocity bias parameter, α_c, and the satellite velocity bias parameter, α_s.

Beyond the vanilla model, galaxy occupation can also depend on secondary halo properties besides the halo mass, a phenomenon commonly referred to as galaxy assembly bias or galaxy

secondary bias (see [63] for a review). We include galaxy assembly bias in an extended HOD model. Our assembly bias prescription is parametrized with two parameters, B_c and B_s, which are the environment-based secondary bias parameters for centrals and satellites, respectively. The environment is defined as the local overdensity over a 5 Mpc/h top-hat filter. Positive/negative B_c and B_s preferentially assign galaxies to halos in higher/lower density environments, whereas $B_c = 0$ and $B_s = 0$ correspond to no assembly bias (see [61] for further details). Finally, we also add a satellite profile parameter, s, to the extended model. This parameter modulates how the radial distribution of satellite galaxies within halos deviates from the radial profile of the halo (potentially due to baryonic effects).

Thus, to summarize, the baseline LRG model is described by 8 parameters: (1) 5 vanilla HOD parameters M_{cut}, M_1, σ, α, κ; (2) an incompleteness parameter f_{ic}; (3) velocity bias parameters α_c and α_s. The extended LRG model introduces three additional parameters: (4) galaxy assembly bias parameters B_c and B_s; (5) satellite profile parameter s. Again, we refer the readers to [62] for additional details. Besides the baseline and extended models we just mentioned, we also consider the vanilla model as our third HOD model, i.e., the one described by Eqs. (4.1) and (4.2) (without including any velocity biases, galaxy assembly bias or satellite profile parameter).

4.2 Fits to the DESI Data

The fits are done by optimizing the HOD models against the DESI's One-Percent Survey two-point correlation function, $\xi(r_\perp, r_\parallel)$, measured on small scales. DESI observed its One-Percent Survey as the third and final phase of its Survey Validation (SV) program[12] in April and May of 2021. Observation fields were chosen to be in twenty non-overlapping "rosettes", where a high completeness was obtained by observing in each rosette at least 12 times. See [9] and [10] for further details.

In this paper, we examine the LRG sample in the redshift range $0.6 < z < 0.8$. In order to run the fits, it is necessary to construct the likelihood function: we assume a simple Gaussian likelihood, for which the χ^2 is given by

$$\chi^2 = \chi_\xi^2 + \chi_{n_{\mathrm{gal}}}^2. \tag{4.3}$$

The term χ_ξ^2 is computed as

$$\chi_\xi^2 = (\xi_{\mathrm{model}} - \xi_{\mathrm{data}})^T C^{-1} (\xi_{\mathrm{model}} - \xi_{\mathrm{data}}), \tag{4.4}$$

where the ξ_{model} is the model-predicted ξ, ξ_{data} is the DESI measurement and C is the covariance matrix. On the other hand, the term $\chi_{n_{\mathrm{gal}}}^2$ is defined as

$$\chi_{n_{\mathrm{gal}}}^2 = \begin{cases} \left(\dfrac{n_{\mathrm{mock}} - n_{\mathrm{data}}}{\sigma_n} \right)^2 & (n_{\mathrm{mock}} < n_{\mathrm{data}}), \\ 0 & (n_{\mathrm{mock}} \geq n_{\mathrm{data}}), \end{cases} \tag{4.5}$$

where σ_n is the uncertainty of the galaxy number density, hereafter $\sigma_n = 0.1 n_{\mathrm{data}}$. The $\chi_{n_{\mathrm{gal}}}^2$ is a half normal around the observed number density, n_{data}. When $n_{\mathrm{mock}} < n_{\mathrm{data}}$, we set $f_{\mathrm{ic}} = 1$ and give a Gaussian-type penalty on the difference between n_{mock} and n_{data}. Otherwise, we set $f_{\mathrm{ic}} = n_{\mathrm{data}}/n_{\mathrm{mock}}$ such that the mock catalog is uniformly downsampled to match the data number density, and we impose no penalty. This definition of $\chi_{n_{\mathrm{gal}}}^2$ allows for incompleteness in the observed galaxy sample while penalizing HOD models that produce insufficient galaxy number densities.

---

[12]There were three main phases of operations during DESI's Survey Validation: the Target Selection Validation phase, the Operations Development stage, and the One-Percent Survey, being the latter the one that corresponds to the data used in this work.
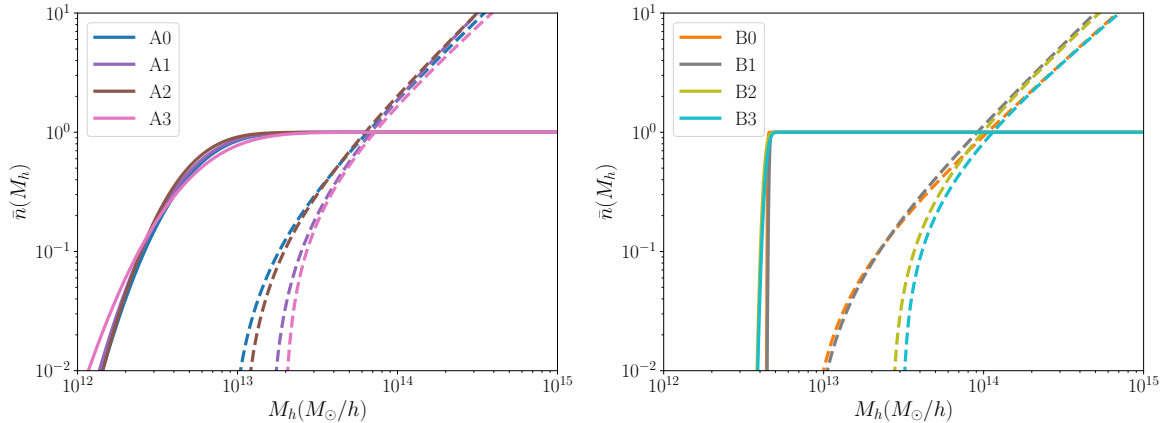
Figure 1. Mean number of galaxies as a function of the halo mass for the different HOD models considered (see Table 4). Solid lines represent central galaxies, whereas dashed lines represent satellites. In the plot on the left, we show the baseline HOD model, whereas in the one on the right, we show the extended one.

The HOD fits to the DESI's One-Percent Survey were performed in [62] using the `AbacusSummit` simulation suite, which we describe in Section 5. The best-fit parameters for our two alternative HOD models, baseline and extended, are listed in Table 4, in the columns labeled as "A0" and "B0", respectively. We find that the best-fit parameters for both models are, in general, quite similar (for those parameters that are common to both models). However, $\log\sigma$ is much smaller for the extended model compared to the baseline one (-2.52 vs -0.60), which makes σ ~100 times smaller for the extended model. Since σ controls the steepness of the transition from 0 to 1 in the number of central galaxies, we find that this transition is much more abrupt for the extended model. We can observe this in Figure 1, in which we show the mean number of central and satellite galaxies as a function of the halo mass for both the baseline and the extended models, respectively. For the extended model, the mean number of central galaxies is a step function.

Besides the baseline and the extended models, the vanilla model was also optimized against the DESI early data, but to a slightly earlier version of the LRG main sample compared to the previous models we just discussed. In order to get such an analysis going, we performed an earlier version of the HOD fits where we did not use the density as the constraint, but only as a lower limit. We fit a combination of wedges and multipoles of the correlation function using scales up to 30 Mpc/h, as described in [64]. Also, in this analysis, we only found a best-fit model rather than exploring the full parameter space to determine the posterior. The motivation for these simpler steps was to produce a large suite of mocks that was close enough to the true DESI sample as fast as possible so that all the different components of the analysis could be developed simultaneously.

5 The AbacusSummit Simulation Suite

The fiducial DESI mocks are built on top of the `AbacusSummit` simulation suite, which is a set of large, high-accuracy cosmological *N*-body simulations using the `Abacus` *N*-body code [65–67]. The entire suite consists of over 150 simulations, containing approximately 60 trillion particles at 97 different cosmologies. This study makes use of the "base" boxes, each of which contains 6912^3 particles within a $(2\,\text{Gpc}/h)^3$ volume corresponding to a particle mass of $2.1 \times 10^9\,M_\odot/h$. [13]

---

[13]For more details, see https://abacussummit.readthedocs.io/en/latest/abacussummit.html

The simulation output is organized into discrete redshift snapshots. For this analysis, we only use redshift snapshot $z = 0.8$ at Planck-2018 cosmology. The suite contains 25 different base boxes with different phases in the initial conditions.

The dark matter halos are identified with the `CompaSO` halo finder, which is a highly efficient on-the-fly group finder specifically designed for the `AbacusSummit` simulations [68]. `CompaSO` builds on the existing spherical overdensity (SO) algorithm by taking into consideration the tidal radius around a smaller halo before competitively assigning halo membership to the particles in an effort to more effectively deblend halos. We also run a post-processing "cleaning" procedure that leverages the halo merger trees to "re-merge" a subset of halos. This is done both to remove over-deblended halos in the spherical overdensity finder and to intentionally merge physically-associated halos that have merged and then physically separated [69].

5.1   Generation of the Mock Catalogs

We apply the best-fit HODs obtained to all 25 base boxes available in `AbacusSummit` at Planck-2018 cosmology to create high-fidelity mocks. In addition to using just the best-fit HOD parameters, we also create additional mocks where we perturb the HOD parameters around the best-fit to generate mocks that share the same cosmology but differ in bias prescriptions. The perturbations are sampled from the 3σ region of the parameter space around the best-fit values. We repeat this procedure for both the baseline HOD model and the extended HOD model, resulting in a set of mocks that encompass a diverse range of possible HODs. All the different HOD models considered and their labels are displayed in Table 4. As mentioned earlier, we denote the best-fit HOD baseline model as A0, and the best-fit extended one as B0. HODs A1, A2 and A3 are random variations within 3σ of the best-fit parameters of A0, whereas HODs B1, B2 and B3 are analogous variations but of B0. Since we have 8 HOD models, we end up with a total of 25×8 `AbacusSummit` simulations; however, each set of 25 is analyzed independently.

In Figure 2 we show the clustering measurements for the LRG One-Percent Survey data in the redshift range from 0.6 to 0.8 (black points with error bars), together with the mean correlation functions of the `AbacusSummit` simulations. We include the results for all the different HOD models considered, all of them listed in Table 4. Even though the HOD fits were only performed in the range $0.1 \, \mathrm{Mpc}/h < s < 30 \, \mathrm{Mpc}/h$, we also include the correlation functions up to 200 Mpc/h to show the excellent agreement between data and simulations.

We also generated mock catalogs populating the 25 `AbacusSummit` boxes using the vanilla best-fit HOD parameters (hereafter, "FirstGen" mocks). These simulations are the ones used to study the impact of theoretical systematics in [36], which is one of the main motivations for including them in this analysis as well. There are two kinds of FirstGen mocks: the first one uses the `AbacusSummit` N-body simulations, which means we have a total of 25 realizations; the second one uses the Zeldovich approximation, and we call them `EZmock`. Throughout this paper, we will refer as FirstGen to the former. The N-body mocks are produced using a similar procedure to the one we described previously (the one used for HODs A0-A3 and B0-B3). One critical difference is that for this case we perform a fit of the density profiles of halos to the Navarro-Frenk-White (NFW) model [70], and then use the best-fit NFW parameters to populate the dark matter halos with satellite galaxies instead of dark matter particles as previously.

It is worth mentioning here that the clustering to which the FirstGen mocks were tuned is an earlier version of the One-Percent Survey. However, we still expect these simulations to be useful for our analysis since the clustering signal is not that different from the one used to optimize the baseline and extended models. Also, the number density of the FirstGen mocks is $10^{-3} h^3 / \mathrm{Mpc}^3$, which is
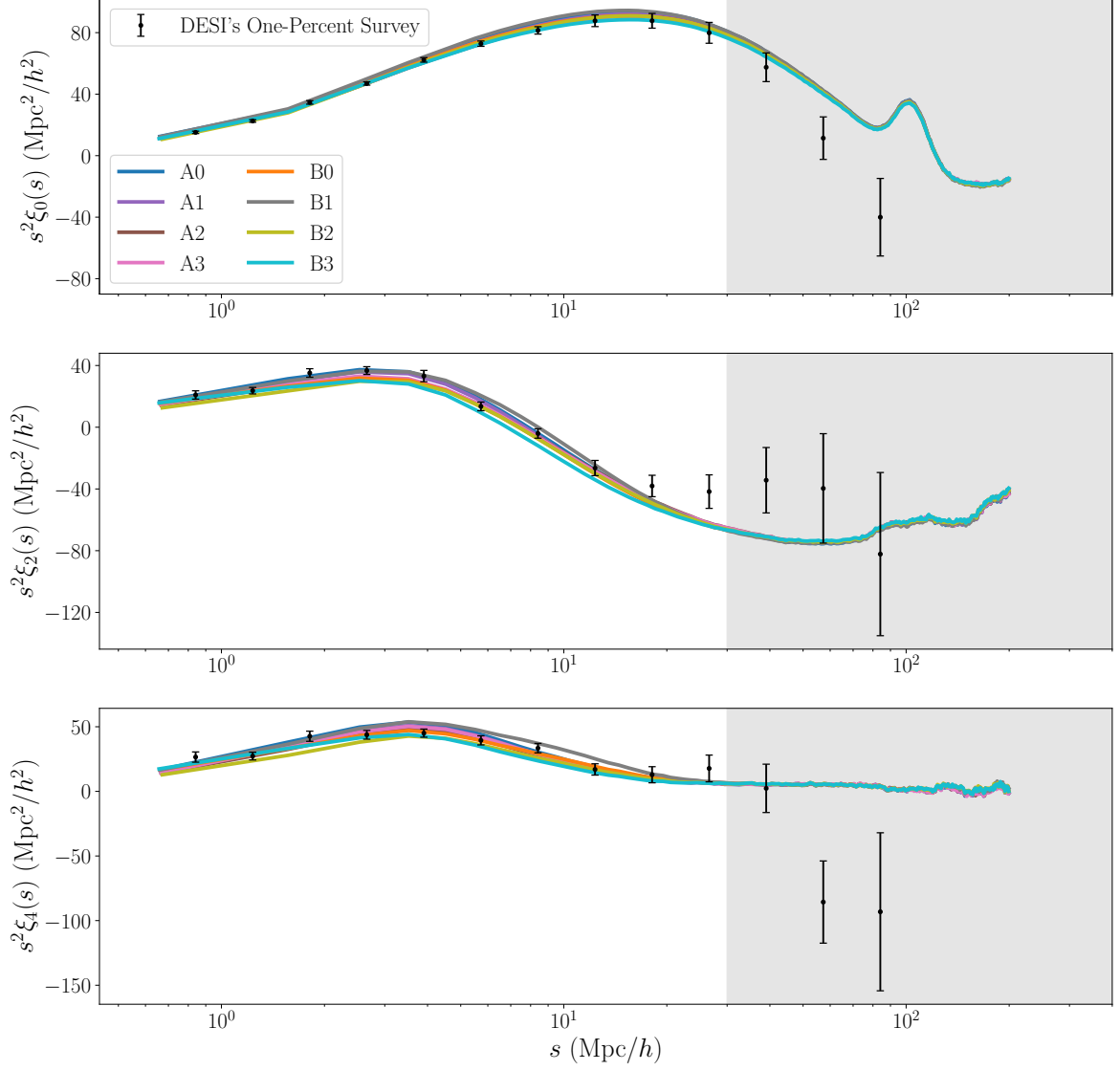
Figure 2. Clustering measurements for the DESI's One-Percent Survey data for LRGs in the redshift range $0.6 < z < 0.8$ (black points with error bars), together with the mean of the `AbacusSummit` simulations for the different HOD models considered (from top to bottom, $\ell = 0, 2, 4$). The scales used for the HOD fits are 0.1 Mpc/$h < s < 30$ Mpc/h, but here we include distances up to 200 Mpc/h (the region between 30 and 200 Mpc/h is shown as a grey band to highlight the fact that those data points were not used for the fits. However, it is worth mentioning that what we used for the fits was $\xi(r_\perp, r_\parallel)$ and not the multipoles, as detailed in the text. The HOD parameters used to generate each set of `AbacusSummit` simulations are displayed in Table 4.

Tracer	LRG $0.6 < z < 0.8$							
Model	Baseline model: Zheng07 $+ f_{ic} + \alpha_c + \alpha_s$				Fully extended model: Zheng07 $+ f_{ic} + \alpha_c + \alpha_s$ $+B_c + B_s + s$			
Label	A0	A1	A2	A3	B0	B1	B2	B3
$\log_{10} M_{\mathrm{cut}}$	12.74	12.72	12.69	12.77	12.65	12.66	12.63	12.64
$\log_{10} M_1$	13.75	13.72	13.72	13.72	13.99	13.92	13.86	13.92
$\log_{10} \sigma$	-0.60	-0.60	-0.64	-0.52	-2.52	-2.23	-1.70	-1.65
α	1.28	1.32	1.31	1.18	1.19	1.30	1.19	1.10
κ	0.20	0.48	0.33	0.52	0.25	0.25	0.79	0.85
α_c	0.17	0.21	0.15	0.12	0.16	0.13	0.21	0.15
α_s	0.91	0.83	0.93	0.93	0.94	0.90	0.91	0.79
B_c	-	-	-	-	0.12	0.15	0.08	0.12
B_s	-	-	-	-	-0.90	-0.97	-0.65	-0.83
s	-	-	-	-	0.11	0.05	-0.02	-0.36

Table 4. Values used for the HOD parameters when generating the mocks. HOD A0 corresponds to the best-fit baseline model, whereas A1, A2 and A3 are random variations within 3σ of the best-fit parameters of HOD A0. On the other hand, HOD B0 corresponds to the best-fit extended model, whereas B1, B2 and B3 are random variations within 3σ of the best-fit parameters of HOD B0. The number density of the mocks is $3.73 \times 10^{-4} h^3/\mathrm{Mpc}^3$ for all the different HODs.

about three times larger compared to that of HODs A0-A3 and B0-B3 (which have a number density of $3.73 \times 10^{-4} h^3/\mathrm{Mpc}^3$).

6 Control Variates Technique

In this section, we briefly describe the control variates (CV) technique used to obtain noise-reduced clustering measurements for our `AbacusSummit` simulations, i.e., this technique allows us to use low-noise datavectors for the estimation of HOD systematics, in both Fourier and configuration spaces. Further details are given in [71], together with the actual application of the code to our simulations.

6.1 Basic Concepts

Here, we briefly summarize the formalism of [71], which builds upon [72] and [73]. The CV technique is a powerful tool to reduce the variance of a random variable, X, given another random variable, C, which is correlated with X and has a mean, $\mu_C = \langle C \rangle$. We write the estimator for a new variable, Y, as

$$Y = X - \beta(C - \mu_C), \tag{6.1}$$

where β is an arbitrary coefficient. Y is unbiased provided that $\langle C - \mu_C \rangle = 0$ regardless of the value of β, and its optimal value can be found to be given by

$$\beta = \frac{\mathrm{Cov}[X, C]}{\mathrm{Var}[C]}. \tag{6.2}$$

For this particular value of β, the variance of Y is given by

$$\mathrm{Var}[Y] = \mathrm{Var}[X] \left(1 - \frac{\mathrm{Cov}^2[X, C]}{\mathrm{Var}[X]\mathrm{Var}[C]} + \beta^2 \mathrm{Var}[\mu_C] \right), \tag{6.3}$$

where we assumed that μ_C, X and β are not correlated with each other. If μ_C is analytically known, then

$$\mathrm{Var}[Y] = \mathrm{Var}[X] \left(1 - \rho_{XC}^2 \right), \tag{6.4}$$

where ρ_{XC} is the Pearson correlation coefficient between X and C,

$$\rho_{XC} \equiv \frac{\mathrm{Cov}[X, C]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[C]}}. \tag{6.5}$$

Therefore, the reduction in the noise after applying the CV technique is given by $\sqrt{1 - \rho_{XC}^2}$, i.e., it increases as the correlation between X and C increases.

6.2 Zeldovich and Linear Control Variates

Relevant to cosmology is the case in which X is a measurement of a summary statistic, such as the power spectrum or the correlation function, calculated from an N-body simulation, such as our `AbacusSummit` mock catalogs. Using the Zeldovich approximation (ZA), [72] and [73] noticed that the Zeldovich displacement fields calculated for the initial conditions are an excellent choice for C, since they exhibit a high correlation with the late-time density field and have an analytically known mean, in both real and redshift spaces. This yields to what is known as Zeldovich control variates (ZCV).

ZCV has several benefits, such as its mean prediction being known analytically and also the fact that the ZA matter density fields for the same seed of initial conditions are strongly correlated with the late-time N-body density fields [74–76]. Novel in [71] is the development of the ZCV method for the configuration space two-point correlation function, as well as the development of the adjacent linear control variates (LCV) method, used to reduce the noise of two-point statistics computed with the reconstructed density fields. The main reason for adopting the ZA instead of linear theory when applied to the pre-reconstruction galaxy catalogs is that much of the decorrelation between initial and final density fields comes from large-scale displacements, which the ZA models well, rather than, e.g., the growth of structure. However, reconstruction aims to remove precisely these displacements, which also removes many of the advantages of ZA over linear theory in this case. Therefore, the LCV provides a very good approximation for the post-reconstruction samples on large scales, see [71]

Following [73], we can write the redshift-space ZCV-reduced power spectrum, $\hat{P}_\ell^{*,tt}(k)$, as

$$\hat{P}_\ell^{*,tt}(k) = \hat{P}_\ell^{tt}(k) - \beta_\ell(k) \left(\hat{P}_\ell^{ZZ}(k) - P_\ell^{ZZ}(k) \right), \tag{6.6}$$

where $\hat{P}_\ell^{tt}(k)$ is the power spectrum measured in the N-body simulation, $\hat{P}_\ell^{ZZ}(k)$ is the ZA power spectrum and $P_\ell^{ZZ}(k)$ is the ZA ensemble-average power spectrum. It can be shown that

$$\beta_\ell(k) = \left[\frac{\hat{P}_\ell^{tZ}(k)}{\hat{P}_\ell^{ZZ}(k)} \right]^2, \tag{6.7}$$

where $\hat{P}_\ell^{tZ}(k)$ is the measured cross-power spectrum between the tracer in question and our ZA control variates.

On the other hand, in the case of LCV, i.e., for post-reconstruction measurements, the noise-reduced power spectrum can be written as

$$\hat{P}_\ell^{*,rr}(k) = \hat{P}_\ell^{rr}(k) - \beta_\ell(k)\left(\hat{P}_\ell^{LL}(k) - P_\ell^{LL}(k)\right), \qquad (6.8)$$

where $\hat{P}_\ell^{rr}(k)$ is the power spectrum measured for a given tracer, $\hat{P}_\ell^{LL}(k)$ is the linear-theory power spectrum and $P_\ell^{LL}(k)$ is the linear-theory ensemble-average power spectrum. Similarly to the ZCV case, it can be shown that

$$\beta_\ell(k) = \left[\frac{\hat{P}_\ell^{rL}(k)}{\hat{P}_\ell^{LL}(k)}\right]^2, \qquad (6.9)$$

where $\hat{P}_\ell^{rL}(k)$ is the measured cross-power spectrum between the true and the modeled reconstructed fields.

The cross-correlation coefficient between the modeled and the measured power spectrum is given by

$$\rho_{XC} = \frac{\mathrm{Cov}[\hat{P}_\ell^{tt}(k), \hat{P}_\ell^{ZZ}(k)]}{\sqrt{\mathrm{Var}[\hat{P}_\ell^{tt}(k)]\mathrm{Var}[\hat{P}_\ell^{ZZ}(k)]}}. \qquad (6.10)$$

The previous expression applies to the ZCV case, whereas changing $t \to r$ and $Z \to L$ applies to the LCV one. Further details on Zeldovich CV and the actual implementation of the CV technique are given in [71].

7 Results

In this section, we describe and discuss the main results obtained in our analysis. It is divided into three sub-sections: in Section 7.1, we validate the CV technique at the level of the clustering signal and also at the level of the BAO scaling parameters; in Section 7.2, we discuss the results of the BAO measurements on the different sets of HOD mocks; and in Section 7.3, we estimate the amplitude of the HOD systematics from the BAO-fit results.

7.1 CV Technique

Here we show how the noise is reduced after applying the CV technique, described in Section 6, at the level of the clustering measurements, and also at the level of the BAO distance measurements.

In Figures 3 and 4 we show a comparison between non-CV and CV-reduced clustering measurements, for Fourier and configuration spaces, respectively. We compute and plot the 1σ region coming from the 25 `AbacusSummit` realizations for the two-point correlation function and the power spectrum as a function of s and k, respectively. We only show the results for the post-reconstruction case, and only for the FirstGen mocks as an example. The CV-reduced results correspond to the narrower shaded regions, which lie within the non-CV ones, which are wider. We find that, after applying the CV, the noise is reduced (the 1σ region shrinks) and the mean is preserved, as expected. This provides a visual validation of the LCV technique in Fourier and configuration spaces. The scales chosen for these figures are the ones used for the BAO fits, as detailed in Table 3. The results for pre-reconstruction are analogous to the post-reconstruction ones, even though we do not show them explicitly, i.e., the ZCV technique was also validated by testing it decreases the noise while preserving the mean in our correlation function and power spectrum measurements.
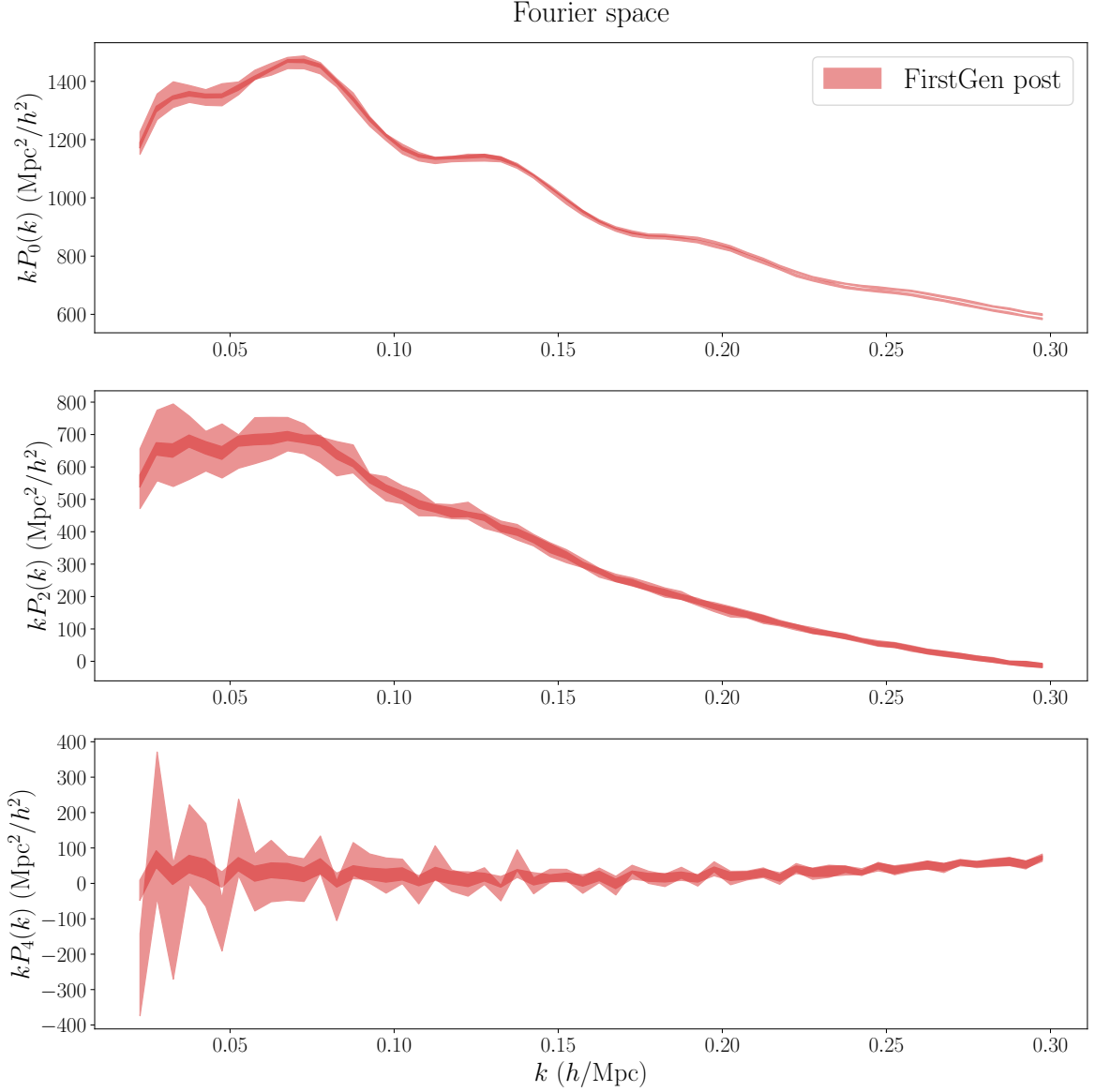
Figure 3. Error reduction at the level of the power spectrum after applying the CV technique. The plot shows the 1σ region of the power spectra computed from the 25 `AbacusSummit` boxes (from top to bottom, $\ell = 0,\ 2,\ 4$), for the non-CV and CV-reduced cases. The non-CV results correspond to the wider shaded regions, whereas the CV-reduced ones correspond to the narrower ones that lie within the former. We only include the results for the FirstGen mocks post-reconstruction as an example, but the results are similar for the other HODs. After applying the CV, the noise is reduced (the 1σ region shrinks) and the mean is preserved. This provides a visual validation of the LCV technique in Fourier space. The scales shown are the ones used for the BAO fits, as specified in Table 3.
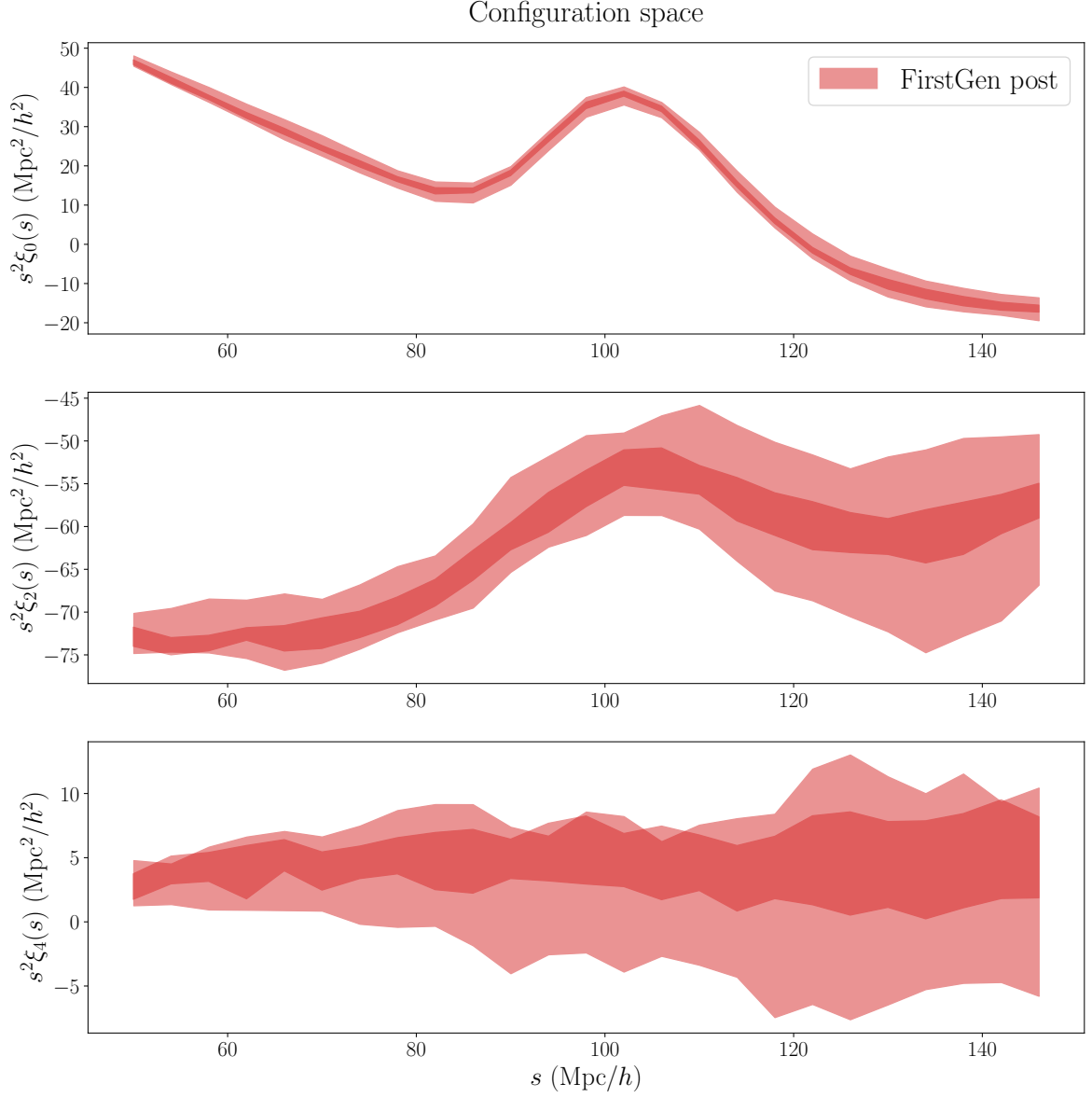
Figure 4. Error reduction at the level of the correlation function after applying the CV technique. The plot shows the 1σ region of the correlation functions computed from the 25 `AbacusSummit` boxes (from top to bottom, $\ell = 0,\ 2,\ 4$), for the non-CV and CV-reduced cases. The non-CV results correspond to the wider shaded regions, whereas the CV-reduced ones correspond to the narrower ones that lie within the former. We only include the results for the FirstGen mocks post-reconstruction as an example, but the results are similar for the other HODs. After applying the CV, the noise is reduced (the 1σ region shrinks) and the mean is preserved. This provides a visual validation of the LCV technique in configuration space. The scales shown are the ones used for the BAO fits, as specified in Table 3.
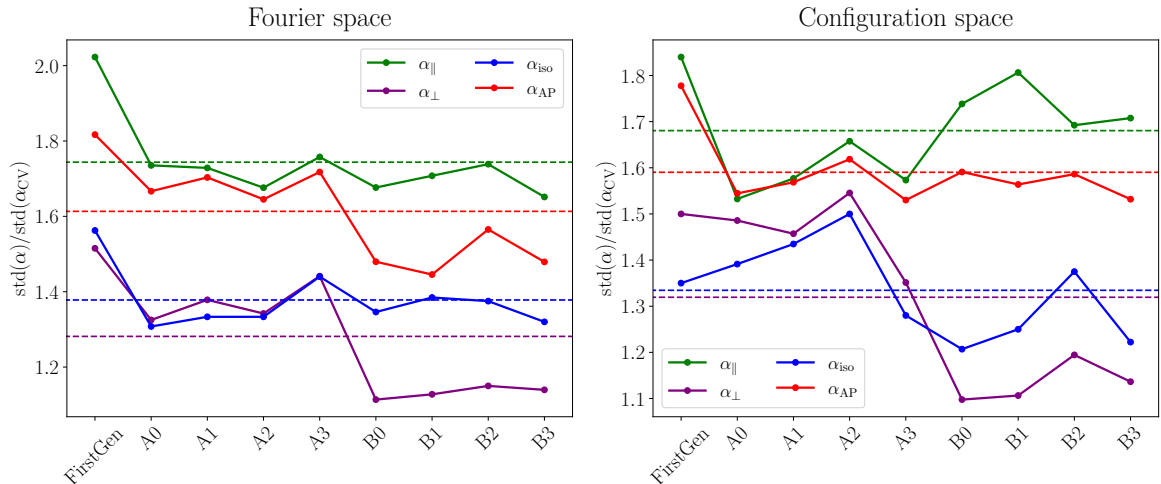
Figure 5. Error reduction factor after applying the CV technique for the Fourier space case (left) and for the configuration space one (right), for the post-reconstruction case. We show such factors for all the different BAO scaling parameters: $\alpha_\parallel$, $\alpha_\perp$, $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$. The mean values obtained from all the HOD models are shown as horizontal dashed lines. In the Fourier space case, we find a reduction in the error of 1.74, 1.28, 1.38 and 1.61 for $\alpha_\parallel$, $\alpha_\perp$, $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$, respectively, whereas in configuration space we find 1.68, 1.32, 1.33 and 1.59. The results for Fourier and configuration spaces are consistent.

We just validated the CV technique results at the level of the clustering measurements (in both Fourier and configuration spaces). Our goal is to now validate them at the level of the BAO distance scale measurement. In Figure 5 we show the error reduction factor when measuring the BAO distance scale after applying the CV technique. We do this by computing the ratio of std(α) (across the 25 AbacusSummit realizations) between the non-CV results and the CV-reduced ones. This plot only shows the results for the post-reconstruction case, for both Fourier (left) and configuration (right) spaces. In the Fourier space case, we find that the average reduction factor across all the HODs is about 1.74, 1.28, 1.38 and 1.61 for $\alpha_\parallel$, $\alpha_\perp$, $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$, respectively, whereas in configuration space we find 1.68, 1.32, 1.33 and 1.59, consistent with Fourier space. We also find that the reduction factor is, in general, larger for the FirstGen mocks, since their number density is also larger with respect to all the other HODs. We conclude that the CV technique effectively reduces the variance not only at the level of the clustering measurements, but also at the level of the BAO measurement, as expected.

7.2 BAO-Fit Results

Here, we show and analyze the BAO-fit results obtained for our simulations.

In Tables 6 and 7 of Appendix A we show the BAO-fit results for Fourier and configuration spaces, respectively, for the different sets of AbacusSummit mocks (FirstGen, A0-A3, B0-B3). We display the average values and standard deviations of $\alpha_\parallel$, $\alpha_\perp$, $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$, together with their average errors. We also include the results for $\Sigma_{\rm nl,\parallel}$, $\Sigma_{\rm nl,\perp}$ and $\Sigma_{\rm s}$, the average χ^2 and the degrees of freedom of the fits. All these different columns are shown for both pre- and post-reconstruction datasets, and for non-CV and CV and-reduced measurements. We find that the standard deviations are, in general, consistent with the average errors. We also find that Fourier and configuration space results are in quite good agreement. The average χ^2 are close to the degrees of freedom for all the different cases (besides the CV-reduced ones, since the covariance matrices are the same we use for the non-CV ones, which makes $\langle \chi^2 \rangle/{\rm dof} < 1$). Post-reconstruction results have lower standard deviations and
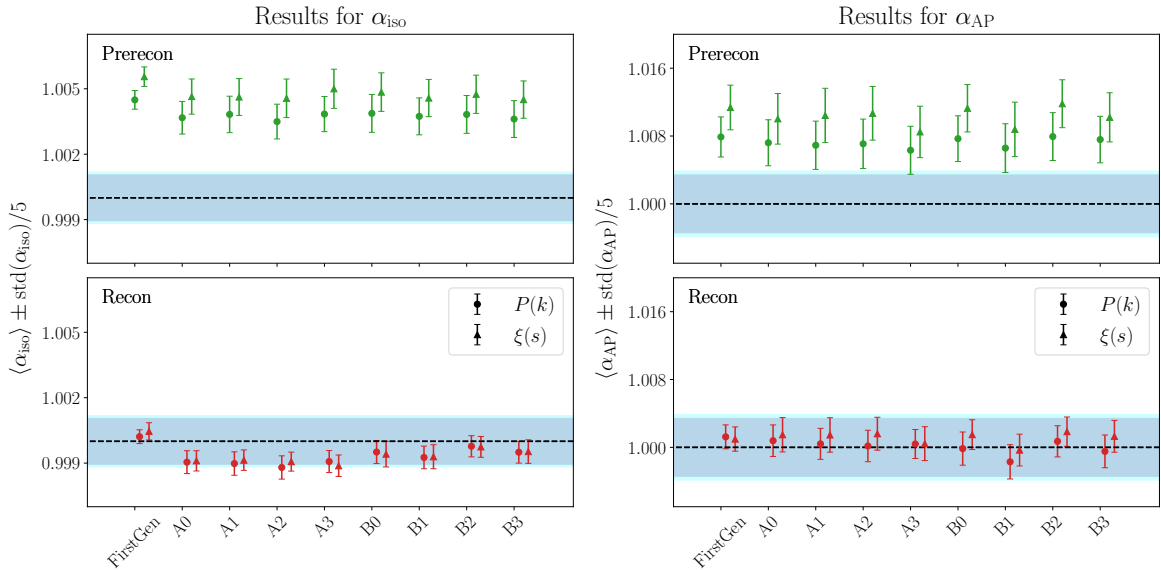
Figure 6. $\langle\alpha\rangle \pm \mathrm{std}(\alpha)/5$ for the different HODs. The plot on the left shows α_{iso}, whereas the one on the right shows α_{AP}. In the top/bottom panels, we show pre-/post-reconstruction results, and these are displayed as circles/triangles for Fourier/configuration spaces. The shaded regions represent $\pm\sigma_{\mathrm{stat,iso}}/5$ (purple/cyan for Fourier/configuration spaces), where $\sigma_{\mathrm{stat,iso}}$ is the measured statistical uncertainty for the DESI 2024 BAO analysis (displayed in Table 1 for Fourier and configuration spaces independently). All the results shown here correspond to CV-reduced measurements.

average errors than pre-reconstruction ones, as expected, for all the different cases displayed in these tables. Finally, CV-reduced results have smaller standard deviations compared to the non-CV ones, as expected for noise-reduced measurements, whereas the average errors are the same (since, as we just mentioned, we are using the same covariances).

In Figure 6 we plot the results displayed in Tables 6 and 7 for α_{iso} and α_{AP} (left and right plots, respectively). We show the average α and we estimate the error as its standard deviation divided by the square root of the number of mocks, i.e., $\langle\alpha\rangle \pm \mathrm{std}(\alpha)/\sqrt{25}$. We include the results for all the different HOD models considered, for pre-reconstruction (top panel) and post-reconstruction (bottom panel), but only for the case of CV-reduced measurements. We find that pre-reconstruction results are biased with respect to 1 (at the 0.4% level for α_{iso} and 1% level for α_{AP}), whereas in the case of post-reconstruction there is a bias at the level of 0.1% for α_{iso}, and no bias for α_{AP}. The results for the different HODs are consistent, but we find a smaller error for the FirstGen mock case (which is expected, since its number density is larger than that of the others). The shaded regions represent a fifth of the measured statistical error, which is displayed in Table 1 (purple for Fourier space, cyan for configuration space). For the case of α_{iso}, the largest difference in the average values shown in Figure 6 (left) is between FirstGen and A2 in Fourier space and between FirstGen and A3 in configuration space. For the case of α_{AP}, taking a look at Figure 6 (right) we find that the largest difference is between FirstGen and B1 in Fourier space and between HODs B1 and B2 in configuration space.

In Appendix B we run BAO fits on the average of the 25 `AbacusSummit` simulations, and analyze its results.

7.3 Estimation of HOD Systematics

The philosophy followed in this study is the following: the same underlying cosmological field can be sampled by galaxies in many different ways, i.e., assuming different HOD models, while still being consistent with measurements. Even in the absence of errors when measuring the BAO signal, these samplings may lead to different α values, and there is no way to ever know in exactly which way the field was sampled. Therefore, there is an unavoidable systematic error floor in any BAO measurement.

Here, we describe the methodology developed to estimate the amplitude of the HOD systematics from the BAO-fit results presented earlier. This methodology is the same as the one applied in [23] for the case of ELGs.

We quantify the level of systematics from the HOD variations by computing the shift

$$\langle \Delta \alpha_{ij} \rangle = \langle \alpha_i - \alpha_j \rangle \tag{7.1}$$

between any pair of HOD models i and j. The average in the previous expression is computed over the 25 `AbacusSummit` realizations. The expected error of $\langle \Delta \alpha \rangle$ can be computed as the standard deviation of $\Delta \alpha$ divided by the square root of the number of mocks. Therefore, we can estimate the significance of the shift as

$$N_\sigma(\alpha_{ij}) = \frac{\langle \Delta \alpha_{ij} \rangle}{\text{std}(\Delta \alpha_{ij})/5}. \tag{7.2}$$

We consider to have detected a HOD systematic shift if the previous quantity reaches the 3σ threshold, i.e., if $N_\sigma(\alpha_{ij}) \geq 3$. Otherwise, we consider it as a non-detection, even though the shift $\Delta \alpha$ could be large. If we find no detection, then we compute the region that encloses the 68% (1σ) of the different values of $\langle \Delta \alpha_{ij} \rangle$, and quote that number as a conservative estimate of our HOD systematic, i.e.,

$$\sigma_{\text{HOD}} = \begin{cases} \max(\langle \Delta \alpha_{ij} \rangle) & \text{if } N_\sigma(\alpha_{ij}) \geq 3 & \text{(HOD sys. detection)}, \\ \sigma_{68\%}(\langle \Delta \alpha_{ij} \rangle) & \text{if } N_\sigma(\alpha_{ij}) < 3 \; \forall i, j & \text{(no HOD sys. detection)}. \end{cases} \tag{7.3}$$

In Figures 7 and 8 we show the values of $\langle \Delta \alpha_{ij} \rangle$ between all the different HODs for α_{iso} and α_{AP}, respectively, and also the corresponding values of $N_\sigma(\alpha_{ij})$. The lower-diagonal regions of these figures represent $\langle \Delta \alpha_{ij} \rangle$ in %, whereas the upper-diagonal ones show $N_\sigma(\alpha_{ij})$, as given by Eq. (7.2). All the results shown in these figures are for CV-reduced post-reconstruction measurements. By analyzing the results on the left panels vs the ones on the right ones in both figures, we find that Fourier and configuration space results are in very good agreement. Taking a look at the results shown in the upper-diagonal regions, i.e., the values for $N_\sigma(\alpha_{ij})$, we find that most of the shifts between the different HODs are not significant enough, i.e., they do not reach the 3σ threshold:

- for the case of α_{iso}, see Figure 7, none of them reaches the 3σ threshold. This happens for both Fourier and configuration space results: we do not measure a significant shift due to the HOD modeling. The largest shift found for α_{iso} is between FirstGen and A2 in Fourier space (0.141% with a significance of 2.60σ) and between FirstGen and A3 in configuration space (0.158% with a significance of 2.70σ). These are the same pairs of HODs that showed the largest differences in the average values of α_{iso} in Figure 6.

- for the case of α_{AP}, see Figure 8, we do find a detection in configuration space, since there is one case that reaches the 3σ threshold: we find a 0.187% shift with a significance of 3.33σ between HODs B0 and B1. For Fourier space, the most significant case found is between B1 and B2: 0.239% shift with a significance of 2.91σ, which does not reach the detection threshold.
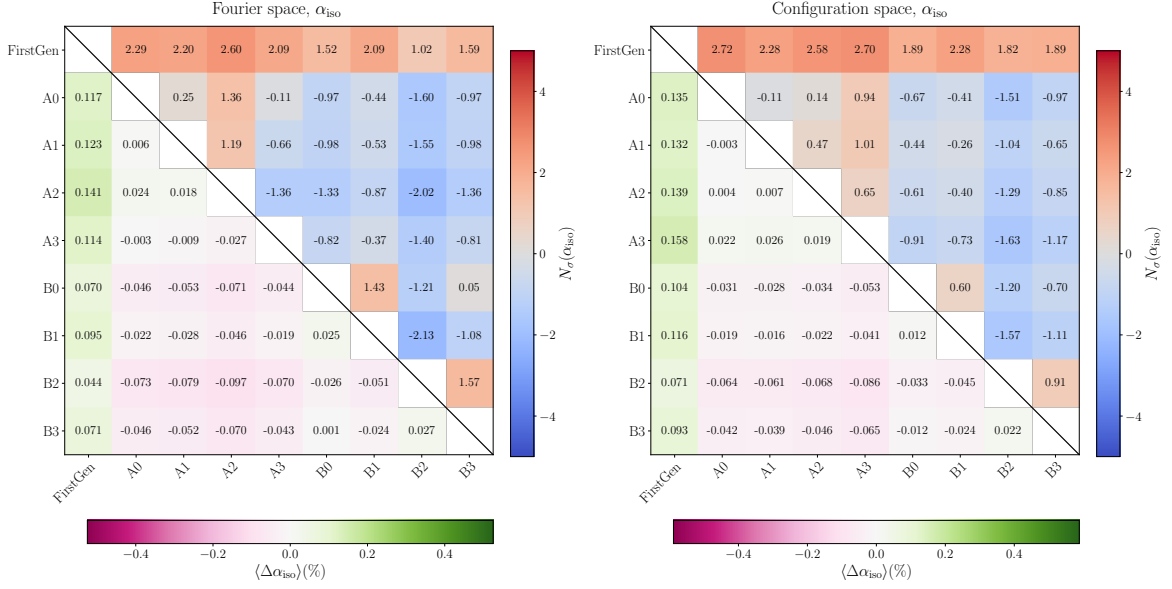
Figure 7. Heat-maps showing differences between the best-fit α_{iso} for any pair of the HOD models considered. The results are consistent between Fourier (left panel) and configuration spaces (right panel). The upper-diagonal region shows the $N_\sigma(\alpha_{\mathrm{iso}})$ as computed with Eq. (7.2), and the limits of the color-bars are ±5. The lower-diagonal region shows the values for $\langle\Delta\alpha_{\mathrm{iso}}\rangle$ in %, and the limits of the color-bars are ±1$\sigma_{\mathrm{stat,iso}}$ (also in %), where $\sigma_{\mathrm{stat,iso}}$ is the measured statistical uncertainty for the DESI 2024 BAO analysis (displayed in Table 1 for Fourier and configuration spaces independently).
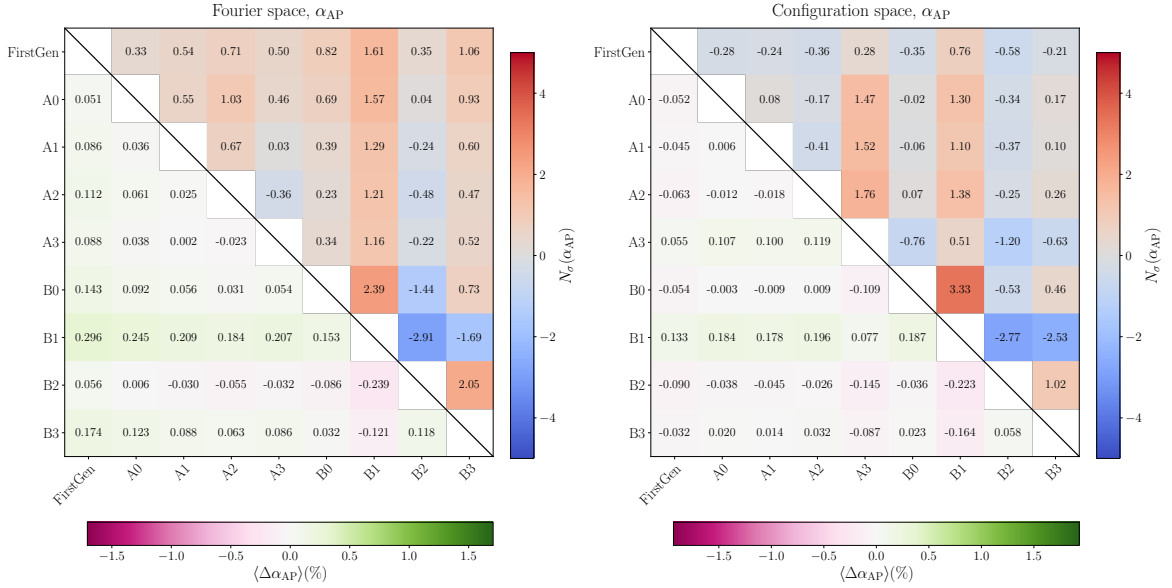
Figure 8. Heat-maps showing differences between the best-fit α_{AP} for any pair of the HOD models considered. The results are consistent between Fourier (left panel) and configuration spaces (right panel). The upper-diagonal region shows the $N_\sigma(\alpha_{\mathrm{AP}})$ as computed with Eq. (7.2), and the limits of the color-bars are ±5. The lower-diagonal region shows the values for $\langle\Delta\alpha_{\mathrm{AP}}\rangle$ in %, and the limits of the color-bars are ±1$\sigma_{\mathrm{stat,AP}}$ (also in %), where $\sigma_{\mathrm{stat,AP}}$ is the measured statistical uncertainty for the DESI 2024 BAO analysis (displayed in Table 1 for Fourier and configuration spaces independently).

We found a HOD systematic detection with an amplitude $\sigma_{\text{HOD}} = 0.187\%$ for α_{AP} in configuration space. For all the other cases, the significance did not reach the 3σ threshold. As we mentioned earlier, for these cases in which we have no detections we estimate the amplitude of the HOD systematic by using all the different values of $\langle \Delta \alpha_{ij} \rangle$: we calculate the region that encloses 68% of them, see Eq. (7.3). For the case of α_{iso}, we find that σ_{HOD} is 0.066% and 0.074% for Fourier and configuration spaces, respectively; whereas for the case of α_{AP} we find a shift of 0.094% in Fourier space. We can also compute this number in the case of α_{AP} in configuration space, for which we find a shift of 0.093%, which is consistent with the one we just quoted for Fourier space (0.094%). It can be compared with the value we obtained from the detection between HODs B0 and B1, 0.187%, which is about two times larger.

In Table 5, we show a summary of the results obtained in this work. The table is split into Fourier and configuration spaces, and for each of them we show α_{iso} and α_{AP} separately. We include a column specifying if we found a 3σ detection or not; one with the amplitude of the systematic, σ_{HOD}, computed using Eq. (7.3); another one with the significance of the systematic in terms of the measured statistical error, $N_{\sigma_{\text{stat}}}$; and a last one with the error increase when adding in quadrature such systematic to the statistical error, $\Delta \sigma_{\text{stat}}/\sigma_{\text{stat}}$. The variable $N_{\sigma_{\text{stat}}}$ is simply computed as the ratio between σ_{HOD} and σ_{stat},

$$N_{\sigma_{\text{stat}}} = \frac{\sigma_{\text{HOD}}}{\sigma_{\text{stat}}}. \tag{7.4}$$

On the other hand, since adding the systematic error in quadrature increases the statistical one as

$$\sigma_{\text{stat}} \rightarrow \sqrt{(\sigma_{\text{stat}})^2 + (\sigma_{\text{HOD}})^2}, \tag{7.5}$$

we defined $\Delta \sigma_{\text{stat}}$ as

$$\Delta \sigma_{\text{stat}} = \sqrt{(\sigma_{\text{stat}})^2 + (\sigma_{\text{HOD}})^2} - \sigma_{\text{stat}}. \tag{7.6}$$

Taking a look at the results displayed in Table 5, we find that the largest increase in the error happens for α_{iso}, for which we find an increase of 0.78% ($0.13\sigma_{\text{stat,iso}}$) for both Fourier and configuration spaces (even though σ_{HOD} is smaller in the case of Fourier space, the measured statistical error is also smaller, see Table 1, which is the reason why these two quantities have the same values for Fourier and configuration spaces). For α_{AP}, the increase in the error is smaller in the case of Fourier space compared to configuration space, 0.15% ($0.06\sigma_{\text{stat,AP}}$) vs 0.48% ($0.10\sigma_{\text{stat,AP}}$), which is due to the 3σ detection we found for the latter.

8 Discussion and Conclusions

In this paper, we have studied the effect of the HOD modeling in the measurement of the position of the BAO peak for the DESI 2024 analysis. In particular, we have focused on the LRG tracer, whereas in its companion paper [23] we focus on ELGs. The methodology followed in both studies is consistent, but the HOD models (and, subsequently, the simulations used) differ from one analysis to the other.

For the DESI 2024 analysis, we have obtained a precision of 0.5-0.6% and 1.7-1.9% in the measurement of the BAO scaling parameters α_{iso} and α_{AP} for the LRG tracer, see [27]. Because of DESI's unprecedented level of precision, it is required to keep all possible sources of systematics well under control, which is our main reason for studying the effect of HOD systematics in this work. To study these systematics, we have used a total of 9 different HOD models and generated a set of 25 `AbacusSummit` cubic boxes for each of them. These 9 HOD models include 4 for the baseline model (vanilla+velocity bias), A0-A3; 4 for the extended one (vanilla+velocity bias+assembly bias+satellite

Space	Parameter	Detection	$\sigma_{\rm HOD}$ (Eq. (7.3))	$N_{\sigma_{\rm stat}}$ (Eq. (7.4))	$\Delta\sigma_{\rm stat}/\sigma_{\rm stat}$ (Eq. (7.6))
Fourier	$\alpha_{\rm iso}$	No	0.066%	0.13	0.78%
	$\alpha_{\rm AP}$	No	0.094%	0.06	0.15%
Config.	$\alpha_{\rm iso}$	No	0.074%	0.13	0.78%
	$\alpha_{\rm AP}$	**Yes**	**0.187%**	**0.10**	**0.48%**

Table 5. Estimation of the HOD systematics: summary of the results. The table is split into Fourier and configuration space results. As columns, we include whether we had a 3σ detection or not, the amplitude of the systematic for each case ($\sigma_{\rm HOD}$), its significance in terms of the measured statistical error ($N_{\sigma_{\rm stat}}$) and the error increase when added in quadrature to the statistical error ($\Delta\sigma_{\rm stat}/\sigma_{\rm stat}$). Following Eq. (7.3), for the detection case we quote $\sigma_{\rm HOD}$ as the value of $\langle\Delta\alpha_{ij}\rangle$ that has the maximum significance, i.e., the largest $N_\sigma(\alpha_{ij})$, provided it is larger than 3σ; and for the no detection cases we compute the region that encloses 68% of the distribution of $\langle\Delta\alpha_{ij}\rangle$, whose values are displayed in Figures 7 and 8.

profile), B0-B3; and 1 for the vanilla, which we referred to as FirstGen. The `AbacusSummit` simulations labeled as A0-A3 and B0-B3 were already generated as a sub-product in [62], and we refer the reader to that paper for further details on them. However, here we performed the reconstruction of the catalogs and the measurement of the clustering signal, both in Fourier and configuration spaces, using the tools within the DESI scientific pipeline. We also produced noise-reduced clustering measurements using the CV technique, as described in [71], which are the fiducial data-vectors we use to obtain our main results (for the post-reconstruction case).

We have estimated the amplitude of the HOD systematics by comparing the BAO-fit results on each set of 25 `AbacusSummit` boxes (one per HOD model) to all the others. In particular, we have computed the average shift in $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$ and its standard deviation, in order to calculate the significance of such shift. We found a significant systematic (more than 3σ detection) between HODs B0 and B1 for $\alpha_{\rm AP}$ in configuration space, with an amplitude of 0.187%. For all the other cases, we did not find a 3σ detection, and therefore we computed a conservative estimate of the systematic using the full heat-maps. By doing this, we obtained an amplitude of the systematic of 0.066% and 0.074% in the case of $\alpha_{\rm iso}$ for Fourier and configuration spaces, respectively. In the case of $\alpha_{\rm AP}$ in Fourier space, we obtained a systematic of 0.094%, which is half of what we found for configuration space (for which we did have a detection).

Finally, we have explicitly quoted how much these systematics increase our error bars when added in quadrature to the measured statistical errors. In the case of $\alpha_{\rm iso}$, the HOD systematics increase the error by 0.78% for both Fourier and configuration spaces, reinforcing the consistency between these two analyses. In the case of $\alpha_{\rm AP}$, the increase in our error bars is of about 0.15 and 0.48%, respectively, which is larger for the latter because of the 3σ detection.

9 Data Availability

The data used in this analysis will be made public as part of DESI Data Release 1. Details can be found in `https://data.desi.lbl.gov/doc/releases/`. Also, the code to reproduce the figures is available at `https://doi.org/10.5281/zenodo.10882070`, as part of DESI's Data Management Plan.

Acknowledgments

References

[1] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L.E. Allen et al., *The DESI Experiment Part I: Science,Targeting, and Survey Design*, *arXiv e-prints* (2016) arXiv:1611.00036 [1611.00036].

[2] M.E. Levi, L.E. Allen, A. Raichoor, C. Baltay, S. BenZvi, F. Beutler et al., *The Dark Energy Spectroscopic Instrument (DESI)*, *arXiv preprint arXiv:1907.10688* (2019) .

[3] B. Abareshi, J. Aguilar, S. Ahlen, S. Alam, D.M. Alexander, R. Alfarsy et al., *Overview of the instrumentation for the Dark Energy Spectroscopic Instrument*, *The Astronomical Journal* **164** (2022) 207.

[4] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini et al., *Planck 2018 results-VI. Cosmological parameters*, *Astronomy & Astrophysics* **641** (2020) A6.

[5] D. Brout, D. Scolnic, B. Popovic, A.G. Riess, A. Carr, J. Zuntz et al., *The pantheon+ analysis: cosmological constraints*, *The Astrophysical Journal* **938** (2022) 110.

[6] S. Alam, M. Aubert, S. Avila, C. Balland, J.E. Bautista, M.A. Bershady et al., *Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory*, *Physical Review D* **103** (2021) 083533.

[7] T. Abbott, M. Aguena, A. Alarcon, S. Allam, O. Alves, A. Amon et al., *Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing*, *Physical Review D* **105** (2022) 023520.

[8] A. Dey, D.J. Schlegel, D. Lang, R. Blum, K. Burleigh, X. Fan et al., *Overview of the DESI legacy imaging surveys*, *The Astronomical Journal* **157** (2019) 168.

[9] D. Collaboration, A. Adame, J. Aguilar, S. Ahlen, S. Alam, G. Aldering et al., *Validation of the scientific program for the Dark Energy Spectroscopic Instrument*, *The Astronomical Journal* **167** (2024) 33pp.

[10] A. Adame, J. Aguilar, S. Ahlen, S. Alam, G. Aldering, D. Alexander et al., *The Early Data Release of the Dark Energy Spectroscopic Instrument*, *arXiv preprint arXiv:2306.06308* (2023) .

[11] C. Hahn, M.J. Wilson, O. Ruiz-Macias, S. Cole, D.H. Weinberg, J. Moustakas et al., *The DESI Bright Galaxy Survey: Final Target Selection, Design, and Validation*, *AJ* **165** (2023) 253 [2208.08512].

[12] R. Zhou, B. Dey, J.A. Newman, D.J. Eisenstein, K. Dawson, S. Bailey et al., *Target Selection and Validation of DESI Luminous Red Galaxies*, *AJ* **165** (2023) 58 [2208.08515].

[13] A. Raichoor, J. Moustakas, J.A. Newman, T. Karim, S. Ahlen, S. Alam et al., *Target Selection and Validation of DESI Emission Line Galaxies*, *AJ* **165** (2023) 126 [2208.08513].

[14] E. Chaussidon, C. Yèche, N. Palanque-Delabrouille, D.M. Alexander, J. Yang, S. Ahlen et al., *Target Selection and Validation of DESI Quasars*, *ApJ* **944** (2023) 107 [2208.08511].

[15] Z. Ding, H.-J. Seo, Z. Vlah, Y. Feng, M. Schmittfull and F. Beutler, *Theoretical systematics of future baryon acoustic oscillation surveys*, *Monthly Notices of the Royal Astronomical Society* **479** (2018) 1021.

[16] R.H. Wechsler and J.L. Tinker, *The connection between galaxies and their dark matter halos*, *Annual Review of Astronomy and Astrophysics* **56** (2018) 435.

[17] A.A. Berlind and D.H. Weinberg, *The halo occupation distribution: Toward an empirical determination of the relation between galaxies and mass*, *The Astrophysical Journal* **575** (2002) 587.

[18] G. Rossi, P.D. Choi, J. Moon, J.E. Bautista, H. Gil-Marín, R. Paviot et al., *The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: N-body mock challenge for galaxy clustering measurements*, *Monthly Notices of the Royal Astronomical Society* **505** (2021) 377.

[19] S. Avila, V. Gonzalez-Perez, F.G. Mohammad, A. de Mattia, C. Zhao, A. Raichoor et al., *The completed sdss-iv extended baryon oscillation spectroscopic survey: exploring the halo occupation distribution model for emission line galaxies*, *Monthly Notices of the Royal Astronomical Society* **499** (2020) 5486.

[20] S. Lin, J.L. Tinker, A. Klypin, F. Prada, M.R. Blanton, J. Comparat et al., *The completed sdss-iv extended baryon oscillation spectroscopic survey: Glam-qpm mock galaxy catalogues for the emission line galaxy sample*, *Monthly Notices of the Royal Astronomical Society* **498** (2020) 5251.

[21] S. Alam, A. De Mattia, A. Tamone, S. Avila, J.A. Peacock, V. Gonzalez-Perez et al., *The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: N-body mock challenge for the eBOSS emission line galaxy sample*, *Monthly Notices of the Royal Astronomical Society* **504** (2021) 4667.

[22] A. Smith, E. Burtin, J. Hou, R. Neveux, A.J. Ross, S. Alam et al., *The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: N-body mock challenge for the quasar sample*, *Monthly Notices of the Royal Astronomical Society* **499** (2020) 269.

[23] C. Garcia-Quintero, J. Mena-Fernández, A. Rocher, S. Yuan, B. Hadzhiyska, O. Alves et al., *Hod-dependent systematics in emission line galaxies for the desi 2024 bao analysis*, *arXiv e-prints* (2024) arXiv:2404.03009 [2404.03009].

[24] N. Findlay et al., *Exploring HOD-dependent systematics for DESI 2024 full shape analysis*, *in preparation* (2024) .

[25] DESI Collaboration, *DESI 2024 I: Data Release 1 of the Dark Energy Spectroscopic Instrument*, *in preparation* (2024) .

[26] DESI Collaboration, *DESI 2024 II: Sample definitions, characteristics and two-point clustering statistics*, *in preparation* (2024) .

[27] DESI Collaboration, A.G. Adame, J. Aguilar, S. Ahlen, S. Alam, D.M. Alexander et al., *DESI 2024 III: Baryon Acoustic Oscillations from Galaxies and Quasars*, *arXiv e-prints* (2024) arXiv:2404.03000 [`2404.03000`].

[28] DESI Collaboration, A.G. Adame, J. Aguilar, S. Ahlen, S. Alam, D.M. Alexander et al., *DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman Alpha Forest*, *arXiv e-prints* (2024) arXiv:2404.03001 [`2404.03001`].

[29] DESI Collaboration, *DESI 2024 V: Analysis of the full shape of two-point clustering statistics from galaxies and quasars*, in preparation (2024) .

[30] DESI Collaboration, A.G. Adame, J. Aguilar, S. Ahlen, S. Alam, D.M. Alexander et al., *DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations*, *arXiv e-prints* (2024) arXiv:2404.03002 [`2404.03002`].

[31] DESI Collaboration, *DESI 2024 VII: Cosmological constraints from full-shape analyses of the two-point clustering statistics measurements*, in preparation (2024) .

[32] DESI Collaboration, *DESI 2024 VIII: Constraints on Primordial Non-Gaussianities*, in preparation (2024) .

[33] X. Chen, Z. Ding, E. Paillas et al., *Extensive analysis of reconstruction algorithms for DESI 2024 baryon acoustic oscillations*, in preparation (2024) .

[34] E. Paillas, Z. Ding, X. Chen, H. Seo, N. Padmanabhan, A. de Mattia et al., *Optimal Reconstruction of Baryon Acoustic Oscillations for DESI 2024*, *arXiv e-prints* (2024) arXiv:2404.03005 [`2404.03005`].

[35] D. Forero-Sanchez et al., *Analytical and EZmock covariance validation for the DESI 2024 results*, in preparation (2024) .

[36] S.-F. Chen, C. Howlett, M. White, P. McDonald, A.J. Ross, H.-J. Seo et al., *Baryon Acoustic Oscillation Theory and Modelling Systematics for the DESI 2024 results*, *arXiv e-prints* (2024) arXiv:2402.14070 [`2402.14070`].

[37] A. Perez-Fernandez, R. Ruggeri et al., *Fiducial Cosmology systematics for DESI 2024 BAO Analysis*, in preparation (2024) .

[38] D. Valcin et al., *Combined tracer analysis for DESI 2024 BAO analysis*, in preparation (2024) .

[39] N. Hand, Y. Li, Z. Slepian and U. Seljak, *An optimal FFT-based anisotropic power spectrum estimator*, *J. Cosmology Astropart. Phys.* **2017** (2017) 002 [`1704.02357`].

[40] S.D. Landy and A.S. Szalay, *Bias and Variance of Angular Correlation Functions*, *ApJ* **412** (1993) 64.

[41] D.J. Eisenstein, H.-J. Seo, E. Sirko and D.N. Spergel, *Improving Cosmological Distance Measurements by Reconstruction of the Baryon Acoustic Peak*, *ApJ* **664** (2007) 675 [`astro-ph/0604362`].

[42] Y.B. Zeldovich, *Gravitational instability: An approximate theory for large density perturbations.*, *A&A* **5** (1970) 84.

[43] A. Nusser and M. Davis, *On the Prediction of Velocity Fields from Redshift Space Galaxy Samples*, *ApJ* **421** (1994) L1 [`astro-ph/9309009`].

[44] N. Padmanabhan, X. Xu, D.J. Eisenstein, R. Scalzo, A.J. Cuesta, K.T. Mehta et al., *A 2 per cent distance to z = 0.35 by reconstructing baryon acoustic oscillations - I. Methods and application to the Sloan Digital Sky Survey*, *MNRAS* **427** (2012) 2132 [`1202.0090`].

[45] A. Burden, W.J. Percival and C. Howlett, *Reconstruction in Fourier space*, *MNRAS* **453** (2015) 456 [`1504.02591`].

[46] M. White, *Reconstruction within the Zeldovich approximation*, *MNRAS* **450** (2015) 3822 [`1504.03677`].

[47] F. Beutler, H.-J. Seo, A.J. Ross, P. McDonald, S. Saito, A.S. Bolton et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Fourier space*, *MNRAS* **464** (2017) 3409 [`1607.03149`].

[48] S. Hinton, *Extraction of Cosmological Information from WiggleZ*, *arXiv e-prints* (2016) arXiv:1604.01830 [1604.01830].

[49] N. Kaiser, *Clustering in real space and in redshift space*, *Monthly Notices of the Royal Astronomical Society* **227** (1987) 1 [https://academic.oup.com/mnras/article-pdf/227/1/1/18522208/mnras227-0001.pdf].

[50] S.R. Hinton, C. Howlett and T.M. Davis, *Barry and the BAO Model Comparison*, *Mon. Not. Roy. Astron. Soc.* **493** (2020) 4078 [1912.01175].

[51] J.S. Speagle, *DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences*, *MNRAS* **493** (2020) 3132 [1904.02180].

[52] O. Alves and DESI Collaboration, "Analytic covariance matrix of DESI galaxy power spectrum multipoles." 2024.

[53] D. Wadekar and R. Scoccimarro, *Galaxy power spectrum multipoles covariance in perturbation theory*, *Phys. Rev. D* **102** (2020) 123517 [1910.02914].

[54] O. Alves et al., *Analytical covariance matrices of DESI galaxy power spectra*, *in preparation* (2024) .

[55] O.H.E. Philcox and D.J. Eisenstein, *Estimating covariance matrices for two- and three-point correlation function moments in Arbitrary Survey Geometries*, *MNRAS* **490** (2019) 5931 [1910.04764].

[56] O.H.E. Philcox, D.J. Eisenstein, R. O'Connell and A. Wiegand, *RASCALC: a jackknife approach to estimating single- and multitracer galaxy covariance matrices*, *MNRAS* **491** (2020) 3290 [1904.11070].

[57] M. Rashkovetskyi, D.J. Eisenstein, J.N. Aguilar, D. Brooks, T. Claybaugh, S. Cole et al., *Validation of semi-analytical, semi-empirical covariance matrices for two-point correlation function for early DESI data*, *MNRAS* **524** (2023) 3894 [2306.06320].

[58] M. Rashkovetskyi, D. Forero-Sánchez, A. de Mattia, D.J. Eisenstein, N. Padmanabhan, H. Seo et al., *Semi-analytical covariance matrices for two-point correlation function for DESI 2024 data*, *arXiv e-prints* (2024) arXiv:2404.03007 [2404.03007].

[59] Z. Zheng, A.A. Berlind, D.H. Weinberg, A.J. Benson, C.M. Baugh, S. Cole et al., *Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies*, *ApJ* **633** (2005) 791 [astro-ph/0408564].

[60] Z. Zheng, A.L. Coil and I. Zehavi, *Galaxy Evolution from Halo Occupation Distribution Modeling of DEEP2 and SDSS Galaxy Clustering*, *ApJ* **667** (2007) 760 [astro-ph/0703457].

[61] S. Yuan, L.H. Garrison, B. Hadzhiyska, S. Bose and D.J. Eisenstein, *ABACUSHOD: A highly efficient extended multi-tracer HOD framework and its application to BOSS and eBOSS data*, *MNRAS* **510** (2021) 3301 [2110.11412].

[62] S. Yuan, H. Zhang, A.J. Ross, J. Donald-McCann, B. Hadzhiyska, R.H. Wechsler et al., *The DESI One-Percent Survey: Exploring the Halo Occupation Distribution of Luminous Red Galaxies and Quasi-Stellar Objects with AbacusSummit*, *arXiv e-prints* (2023) arXiv:2306.06314 [2306.06314].

[63] R.H. Wechsler and J.L. Tinker, *The Connection Between Galaxies and Their Dark Matter Halos*, *ARA&A* **56** (2018) 435 [1804.03097].

[64] S. Alam, A. Paranjape and J.A. Peacock, *Impact of tidal environment on galaxy clustering in GAMA*, *Monthly Notices of the Royal Astronomical Society* **527** (2023) 3771 [https://academic.oup.com/mnras/article-pdf/527/2/3771/53687659/stad3423.pdf].

[65] N.A. Maksimova, L.H. Garrison, D.J. Eisenstein, B. Hadzhiyska, S. Bose and T.P. Satterthwaite, *ABACUSSUMMIT: A Massive Set of High-Accuracy, High-Resolution N-Body Simulations*, *MNRAS* (2021) [2110.11398].

[66] L.H. Garrison, D.J. Eisenstein and P.A. Pinto, *A high-fidelity realization of the Euclid code comparison N-body simulation with ABACUS*, *MNRAS* **485** (2019) 3370 [1810.02916].

[67] L.H. Garrison, D.J. Eisenstein, D. Ferrer, N.A. Maksimova and P.A. Pinto, *The ABACUS cosmological N-body code*, *MNRAS* **508** (2021) 575 [2110.11392].

[68] B. Hadzhiyska, D. Eisenstein, S. Bose, L.H. Garrison and N. Maksimova, *COMPASO: A new halo finder for competitive assignment to spherical overdensities*, *MNRAS* **509** (2022) 501 [2110.11408].

[69] S. Bose, D.J. Eisenstein, B. Hadzhiyska, L.H. Garrison and S. Yuan, *Constructing high-fidelity halo merger trees in ABACUSSUMMIT*, *MNRAS* **512** (2022) 837 [2110.11409].

[70] J.F. Navarro, C.S. Frenk and S.D. White, *A universal density profile from hierarchical clustering*, *The Astrophysical Journal* **490** (1997) 493.

[71] B. Hadzhiyska, M.J. White, X. Chen, L.H. Garrison, J. DeRose, N. Padmanabhan et al., *Mitigating the noise of DESI mocks using analytic control variates*, *arXiv preprint arXiv:2308.12343* (2023) .

[72] N. Kokron, S.-F. Chen, M. White, J. DeRose and M. Maus, *Accurate predictions from small boxes: variance suppression via the Zel'dovich approximation*, *J. Cosmology Astropart. Phys.* **2022** (2022) 059 [2205.15327].

[73] J. DeRose, S.-F. Chen, N. Kokron and M. White, *Precision redshift-space galaxy power spectra using Zel'dovich control variates*, *J. Cosmology Astropart. Phys.* **2023** (2023) 008 [2210.14239].

[74] A.G. Doroshkevich, Y.B. Zeldovich, R.A. Syunyaev and M.Y. Khlopov, *Astrophysical implications of the neutrino rest mass. II - The density-perturbation spectrum and small-scale fluctuations in the microwave background. III - Nonlinear growth of perturbations and the missing mass*, *Pisma v Astronomicheskii Zhurnal* **6** (1980) 457.

[75] P. Coles, A.L. Melott and S.F. Shandarin, *Testing approximations for non-linear gravitational clustering*, *MNRAS* **260** (1993) 765.

[76] J.L. Pauls and A.L. Melott, *Hierarchical pancaking: why the Zel'dovich approximation describes coherent large-scale structure in N-body simulations of gravitational clustering*, *MNRAS* **274** (1995) 99 [astro-ph/9408019].

A   BAO-Fit Tables

Here we explicitly show the BAO-fit results for the different HOD mocks considered in this paper. We include the results for all the HODs displayed in Table 4, i.e., A0-A3 and B0-B3, and also including the FirstGen. In Tables 6 and 7 we display the BAO-fit results in Fourier and configuration spaces, respectively, including as columns the average values (averaged over the 25 AbacusSummit realizations) of $\alpha_\parallel$, $\alpha_\perp$, $\alpha_{\rm iso}$, $\alpha_{\rm AP}$, their standard deviations (std) and errors (σ), together with $\Sigma_{\rm nl,\parallel}$, $\Sigma_{\rm nl,\perp}$, $\Sigma_{\rm s}$ and χ^2/dof. We include results for pre-reconstructed and post-reconstructed mocks, non-CV and CV-reduced. The fits were run using the code Barry, with the methodology described in Section 3.2.

B   BAO Fits to the Average of the Simulations

Here, we analyze the BAO-fit results when fitting the average of the 25 AbacusSummit realizations for each HOD model (rather than the average of the results fitting the individual realizations). The results are shown in Figure 9, where we plot the 68% confidence regions for $\alpha_\parallel$ and $\alpha_\perp$ as obtained from the fit to the average of the 25 mocks (Fourier space in the panel on the left and configuration space in the one on the right). All the results shown in this figure correspond to CV-reduced measurements. We find that most of the HOD models are consistent with the fiducial result of $\alpha_\parallel = 1$ and $\alpha_\perp = 1$ within the 1σ region. It must be noted that the size of the ellipses comes directly from the covariance matrices used for the fits, rather than the dispersion over the 25 realizations (however, it was re-scaled

								Fourier space results								
case	$\langle\alpha_\parallel\rangle$	std($\alpha_\parallel$)	$\langle\sigma_\parallel\rangle$	$\langle\alpha_\perp\rangle$	std($\alpha_\perp$)	$\langle\sigma_\perp\rangle$	$\langle\alpha_{\rm iso}\rangle$	std($\alpha_{\rm iso}$)	$\langle\sigma_{\rm iso}\rangle$	$\langle\alpha_{\rm AP}\rangle$	std($\alpha_{\rm AP}$)	$\langle\sigma_{\rm AP}\rangle$	$\langle\Sigma_{\rm nl,\parallel}\rangle$	$\langle\Sigma_{\rm nl,\perp}\rangle$	$\langle\Sigma_{\rm s}\rangle$	$\langle\chi^2\rangle$/dof
							non-CV pre-recon									
FirstGen	1.0124	0.0174	0.0175	1.0013	0.0071	0.0076	1.0049	0.0042	0.0059	1.0113	0.0231	0.0220	9.18	4.83	1.20	95.5/93
A0	1.0104	0.0186	0.0198	1.0006	0.0083	0.0092	1.0037	0.0063	0.0068	1.0099	0.0236	0.0252	9.24	4.79	1.18	94.6/93
A1	1.0101	0.0184	0.0199	1.0009	0.0081	0.0091	1.0038	0.0065	0.0068	1.0093	0.0228	0.0252	9.25	4.70	1.18	94.4/93
A2	1.0099	0.0186	0.0198	1.0005	0.0083	0.0092	1.0035	0.0064	0.0068	1.0094	0.0234	0.0252	9.24	4.76	1.19	94.6/93
A3	1.0099	0.0187	0.0200	1.0013	0.0078	0.0091	1.0040	0.0064	0.0068	1.0086	0.0228	0.0253	9.29	4.72	1.19	94.8/93
B0	1.0110	0.0181	0.0202	1.0007	0.0079	0.0093	1.0040	0.0065	0.0069	1.0103	0.0222	0.0258	9.22	4.84	1.17	95.5/93
B1	1.0095	0.0195	0.0203	1.0009	0.0083	0.0094	1.0037	0.0068	0.0070	1.0087	0.0239	0.0258	9.32	4.86	1.18	94.6/93
B2	1.0113	0.0184	0.0200	1.0003	0.0084	0.0092	1.0038	0.0065	0.0068	1.0110	0.0231	0.0255	9.29	4.78	1.19	94.4/93
B3	1.0105	0.0179	0.0201	1.0005	0.0083	0.0092	1.0037	0.0062	0.0069	1.0101	0.0229	0.0255	9.25	4.76	1.17	95.6/93
							non-CV post-recon									
FirstGen	1.0003	0.0089	0.0078	0.9996	0.0050	0.0045	0.9998	0.0025	0.0030	1.0008	0.0129	0.0106	5.52	1.73	1.11	93.4/93
A0	0.9985	0.0118	0.0096	0.9986	0.0053	0.0056	0.9986	0.0034	0.0038	1.0000	0.0155	0.0129	5.75	1.81	1.13	93.5/93
A1	0.9982	0.0121	0.0095	0.9987	0.0051	0.0056	0.9985	0.0036	0.0037	0.9996	0.0155	0.0129	5.74	1.79	1.11	93.8/93
A2	0.9979	0.0119	0.0096	0.9986	0.0051	0.0056	0.9984	0.0036	0.0037	0.9994	0.0153	0.0130	5.70	1.78	1.12	94.1/93
A3	0.9985	0.0116	0.0097	0.9988	0.0049	0.0056	0.9986	0.0036	0.0038	0.9997	0.0146	0.0131	5.75	1.79	1.13	93.0/93
B0	0.9987	0.0114	0.0097	0.9994	0.0049	0.0056	0.9991	0.0035	0.0038	0.9994	0.0145	0.0131	5.73	1.82	1.11	93.9/93
B1	0.9974	0.0111	0.0096	0.9997	0.0053	0.0057	0.9989	0.0036	0.0038	0.9977	0.0146	0.0131	5.91	1.87	1.11	94.0/93
B2	0.9996	0.0113	0.0095	0.9993	0.0046	0.0056	0.9994	0.0033	0.0038	1.0004	0.0144	0.0129	5.74	1.82	1.12	93.6/93
B3	0.9985	0.0109	0.0096	0.9995	0.0049	0.0056	0.9991	0.0033	0.0038	0.9990	0.0142	0.0130	5.61	1.80	1.12	94.8/93
							CV-reduced pre-recon									
FirstGen	1.0099	0.0083	0.0173	1.0019	0.0043	0.0076	1.0045	0.0021	0.0058	1.0079	0.0118	0.0217	9.14	4.75	1.20	23.2/93
A0	1.0085	0.0093	0.0195	1.0014	0.0062	0.0090	1.0037	0.0037	0.0067	1.0072	0.0136	0.0247	9.22	4.73	1.19	39.3/93
A1	1.0085	0.0103	0.0194	1.0016	0.0064	0.0090	1.0038	0.0042	0.0067	1.0069	0.0142	0.0246	9.23	4.65	1.18	39.2/93
A2	1.0083	0.0103	0.0194	1.0012	0.0065	0.0090	1.0035	0.0040	0.0067	1.0071	0.0146	0.0247	9.21	4.70	1.19	40.2/93
A3	1.0081	0.0106	0.0194	1.0018	0.0060	0.0090	1.0038	0.0040	0.0067	1.0063	0.0142	0.0246	9.26	4.67	1.18	40.4/93
B0	1.0091	0.0100	0.0196	1.0014	0.0062	0.0092	1.0039	0.0043	0.0067	1.0077	0.0135	0.0250	9.20	4.78	1.18	38.3/93
B1	1.0082	0.0107	0.0200	1.0017	0.0062	0.0093	1.0037	0.0042	0.0069	1.0066	0.0144	0.0254	9.31	4.82	1.18	38.1/93
B2	1.0092	0.0104	0.0194	1.0013	0.0064	0.0090	1.0038	0.0043	0.0067	1.0079	0.0141	0.0247	9.25	4.73	1.17	38.1/93
B3	1.0087	0.0097	0.0196	1.0012	0.0064	0.0090	1.0036	0.0042	0.0067	1.0076	0.0137	0.0248	9.22	4.71	1.17	38.8/93
							CV-reduced post-recon									
FirstGen	1.0011	0.0044	0.0076	0.9998	0.0033	0.0044	1.0002	0.0016	0.0030	1.0012	0.0071	0.0102	5.36	1.66	1.11	39.0/93
A0	0.9996	0.0068	0.0093	0.9988	0.0040	0.0056	0.9990	0.0026	0.0037	1.0008	0.0093	0.0127	5.64	1.76	1.12	52.9/93
A1	0.9993	0.0070	0.0093	0.9989	0.0037	0.0055	0.9990	0.0027	0.0037	1.0004	0.0091	0.0126	5.63	1.72	1.12	52.8/93
A2	0.9989	0.0071	0.0093	0.9988	0.0038	0.0055	0.9988	0.0027	0.0037	1.0002	0.0093	0.0126	5.59	1.73	1.13	53.2/93
A3	0.9993	0.0066	0.0093	0.9990	0.0034	0.0055	0.9991	0.0025	0.0037	1.0004	0.0085	0.0127	5.63	1.73	1.12	52.5/93
B0	0.9994	0.0068	0.0093	0.9996	0.0044	0.0055	0.9995	0.0026	0.0037	0.9999	0.0098	0.0126	5.61	1.76	1.12	52.3/93
B1	0.9981	0.0065	0.0094	0.9999	0.0047	0.0055	0.9993	0.0026	0.0037	0.9983	0.0101	0.0128	5.79	1.80	1.12	52.9/93
B2	1.0002	0.0065	0.0093	0.9996	0.0040	0.0055	0.9998	0.0024	0.0037	1.0007	0.0092	0.0126	5.64	1.76	1.13	51.6/93
B3	0.9992	0.0066	0.0093	0.9997	0.0043	0.0054	0.9995	0.0025	0.0037	0.9995	0.0096	0.0126	5.50	1.73	1.12	53.0/93

Table 6. BAO-fit results in Fourier space: 25 `AbacusSummit` cubic boxes. Results shown for the different HOD models listed in Table 4 and also for FirstGen, all of them pre- and post-reconstruction. We included the results for the non-CV and the CV-reduced measurements. The different columns show the average values (averaged over the 25 `AbacusSummit` realizations) for $\alpha_\parallel$, $\alpha_\perp$, $\alpha_{\rm iso}$, $\alpha_{\rm AP}$, their standard deviations (std) and errors (σ), together with the average values for $\Sigma_{\rm nl,\parallel}$, $\Sigma_{\rm nl,\perp}$, $\Sigma_{\rm s}$ and χ^2/dof. The fits were run using the code `Barry`, with the methodology described in Section 3.2.

	Configuration space results															
case	$\langle\alpha_\parallel\rangle$	$\mathrm{std}(\alpha_\parallel)$	$\langle\sigma_\parallel\rangle$	$\langle\alpha_\perp\rangle$	$\mathrm{std}(\alpha_\perp)$	$\langle\sigma_\perp\rangle$	$\langle\alpha_{\mathrm{iso}}\rangle$	$\mathrm{std}(\alpha_{\mathrm{iso}})$	$\langle\sigma_{\mathrm{iso}}\rangle$	$\langle\alpha_{\mathrm{AP}}\rangle$	$\mathrm{std}(\alpha_{\mathrm{AP}})$	$\langle\sigma_{\mathrm{AP}}\rangle$	$\langle\Sigma_{\mathrm{nl},\parallel}\rangle$	$\langle\Sigma_{\mathrm{nl},\perp}\rangle$	$\langle\Sigma_{\mathrm{s}}\rangle$	$\langle\chi^2\rangle/\mathrm{dof}$
	non-CV pre-recon															
FirstGen	1.0140	0.0179	0.0193	1.0008	0.0072	0.0082	1.0051	0.0041	0.0063	1.0133	0.0238	0.0243	9.66	4.94	1.18	33.2/37
A0	1.0124	0.0185	0.0220	1.0001	0.0088	0.0099	1.0040	0.0060	0.0074	1.0124	0.0243	0.0279	9.67	4.89	1.16	37.3/37
A1	1.0118	0.0183	0.0222	1.0004	0.0084	0.0099	1.0040	0.0060	0.0075	1.0115	0.0237	0.0281	9.69	4.85	1.17	36.3/37
A2	1.0121	0.0187	0.0228	1.0000	0.0085	0.0103	1.0039	0.0060	0.0076	1.0122	0.0243	0.0290	9.67	4.87	1.16	36.5/37
A3	1.0117	0.0182	0.0226	1.0010	0.0079	0.0102	1.0044	0.0059	0.0076	1.0107	0.0230	0.0286	9.71	4.89	1.17	36.5/37
B0	1.0134	0.0180	0.0224	1.0000	0.0080	0.0101	1.0043	0.0062	0.0075	1.0134	0.0226	0.0285	9.66	4.98	1.16	37.0/37
B1	1.0121	0.0189	0.0220	1.0002	0.0084	0.0098	1.0040	0.0062	0.0074	1.0120	0.0241	0.0278	9.73	5.00	1.16	37.7/37
B2	1.0137	0.0184	0.0223	0.9995	0.0084	0.0100	1.0041	0.0062	0.0074	1.0143	0.0236	0.0284	9.70	4.93	1.16	37.4/37
B3	1.0130	0.0181	0.0226	0.9996	0.0086	0.0101	1.0039	0.0059	0.0076	1.0135	0.0237	0.0287	9.68	4.94	1.16	37.0/37
	non-CV post-recon															
FirstGen	0.9998	0.0092	0.0080	0.9996	0.0048	0.0046	0.9997	0.0027	0.0031	1.0002	0.0128	0.0108	5.45	1.84	1.12	39.3/37
A0	0.9991	0.0118	0.0100	0.9981	0.0052	0.0058	0.9984	0.0032	0.0039	1.0011	0.0156	0.0136	5.77	2.06	1.12	35.5/37
A1	0.9984	0.0123	0.0100	0.9982	0.0051	0.0058	0.9983	0.0033	0.0039	1.0003	0.0160	0.0135	5.71	2.04	1.12	36.7/37
A2	0.9982	0.0121	0.0102	0.9982	0.0051	0.0060	0.9982	0.0033	0.0040	1.0000	0.0157	0.0138	5.69	2.05	1.12	33.7/37
A3	0.9986	0.0118	0.0102	0.9983	0.0050	0.0059	0.9983	0.0032	0.0040	1.0004	0.0153	0.0139	5.74	2.08	1.12	34.5/37
B0	0.9992	0.0113	0.0101	0.9989	0.0045	0.0058	0.9989	0.0035	0.0039	1.0004	0.0140	0.0136	5.69	1.98	1.13	38.7/37
B1	0.9978	0.0112	0.0101	0.9992	0.0052	0.0057	0.9987	0.0035	0.0039	0.9986	0.0147	0.0136	5.81	2.00	1.11	38.5/37
B2	0.9999	0.0110	0.0101	0.9987	0.0043	0.0058	0.9991	0.0033	0.0039	1.0013	0.0138	0.0136	5.73	2.00	1.12	38.5/37
B3	0.9987	0.0111	0.0100	0.9990	0.0050	0.0058	0.9989	0.0033	0.0039	0.9997	0.0144	0.0135	5.63	1.98	1.12	39.6/37
	CV-reduced pre-recon															
FirstGen	1.0132	0.0091	0.0194	1.0019	0.0048	0.0083	1.0056	0.0022	0.0063	1.0114	0.0132	0.0244	9.55	4.93	1.21	26.7/37
A0	1.0114	0.0101	0.0219	1.0014	0.0069	0.0100	1.0046	0.0040	0.0074	1.0100	0.0149	0.0278	9.60	4.90	1.18	33.6/37
A1	1.0117	0.0116	0.0220	1.0013	0.0067	0.0100	1.0046	0.0042	0.0074	1.0104	0.0160	0.0279	9.62	4.88	1.18	32.2/37
A2	1.0118	0.0110	0.0225	1.0011	0.0072	0.0102	1.0046	0.0044	0.0076	1.0107	0.0158	0.0285	9.62	4.90	1.19	32.6/37
A3	1.0108	0.0114	0.0225	1.0023	0.0064	0.0104	1.0050	0.0045	0.0076	1.0085	0.0152	0.0286	9.61	4.91	1.19	32.3/37
B0	1.0125	0.0101	0.0221	1.0012	0.0065	0.0101	1.0048	0.0044	0.0075	1.0113	0.0140	0.0282	9.60	4.99	1.17	32.7/37
B1	1.0105	0.0112	0.0219	1.0018	0.0071	0.0100	1.0046	0.0043	0.0074	1.0088	0.0160	0.0279	9.66	5.04	1.19	33.3/37
B2	1.0128	0.0107	0.0223	1.0009	0.0062	0.0100	1.0047	0.0044	0.0074	1.0118	0.0141	0.0282	9.62	4.94	1.17	32.5/37
B3	1.0114	0.0101	0.0225	1.0012	0.0067	0.0103	1.0045	0.0043	0.0076	1.0102	0.0145	0.0286	9.61	4.97	1.18	32.5/37
	CV-reduced post-recon															
FirstGen	1.0011	0.0050	0.0079	1.0002	0.0032	0.0045	1.0005	0.0020	0.0031	1.0010	0.0072	0.0106	5.32	1.84	1.17	41.1/37
A0	1.0001	0.0077	0.0099	0.9987	0.0035	0.0058	0.9991	0.0023	0.0039	1.0015	0.0101	0.0135	5.69	2.05	1.15	33.8/37
A1	1.0001	0.0078	0.0100	0.9987	0.0035	0.0058	0.9991	0.0023	0.0039	1.0015	0.0102	0.0135	5.65	2.00	1.15	36.9/37
A2	1.0002	0.0073	0.0101	0.9986	0.0033	0.0059	0.9991	0.0022	0.0039	1.0016	0.0097	0.0138	5.64	1.99	1.15	35.4/37
A3	0.9992	0.0075	0.0101	0.9988	0.0037	0.0059	0.9989	0.0025	0.0040	1.0004	0.0100	0.0138	5.65	2.02	1.15	34.6/37
B0	1.0004	0.0065	0.0100	0.9989	0.0041	0.0058	0.9994	0.0029	0.0039	1.0015	0.0088	0.0135	5.65	2.00	1.15	38.0/37
B1	0.9991	0.0062	0.0101	0.9994	0.0047	0.0057	0.9993	0.0028	0.0039	0.9997	0.0094	0.0136	5.75	2.02	1.14	37.8/37
B2	1.0010	0.0065	0.0099	0.9992	0.0036	0.0058	0.9997	0.0024	0.0039	1.0019	0.0087	0.0135	5.65	1.96	1.15	37.3/37
B3	1.0004	0.0065	0.0100	0.9991	0.0044	0.0059	0.9995	0.0027	0.0039	1.0013	0.0094	0.0136	5.59	2.00	1.15	38.3/37

Table 7. BAO-fit results in configuration space: 25 `AbacusSummit` cubic boxes. Results shown for the different HOD models listed in Table 4 and also for FirstGen, all of them pre- and post-reconstruction. We included the results for the non-CV and the CV-reduced measurements. The different columns show the average values (averaged over the 25 `AbacusSummit` realizations) for $\alpha_\parallel$, $\alpha_\perp$, α_{iso}, α_{AP}, their standard deviations (std) and errors (σ), together with the average values for $\Sigma_{\mathrm{nl},\parallel}$, $\Sigma_{\mathrm{nl},\perp}$, Σ_{s} and χ^2/dof. The fits were run using the code `Barry`, with the methodology described in Section 3.2.

Figure 9. 68% confidence regions, or 1σ contours, for $\alpha_\parallel$ and $\alpha_\perp$, as obtained from the fit to the average of the 25 `AbacusSummit` realizations. The Fourier space results are shown on the left panel, whereas the configuration space ones are shown on the right one. The stars represent the best-fit values for each type of HOD. The dashed black lines represent the fiducial value for both $\alpha_\parallel$ and $\alpha_\perp$. The fiducial value of $\alpha_{\rm iso}$ is shown in a gray dashed-dotted line ($\alpha_{\rm iso} = 1$). Similarly, the fiducial value for $\alpha_{\rm AP}$ is shown in a dotted gray line ($\alpha_{\rm AP} = 1$). All the results displayed in this figure were computed from the CV-reduced measurements.

with a factor $\sqrt{25}$). The degeneracy direction (i.e., the angle with respect to the horizontal line) was computed from the correlation of $\alpha_\parallel$ and $\alpha_\perp$ between the 25 realizations.

C   Author Affiliations

[1]Laboratoire de Physique Subatomique et de Cosmologie, 53 Avenue des Martyrs, 38000 Grenoble, France

[2]CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain

[3]Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA

[4]SLAC National Accelerator Laboratory, Menlo Park, CA 94305, USA

[5]Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

[6]University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA

[7]University of Michigan, Ann Arbor, MI 48109, USA

[8]Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

[9]Department of Physics & Astronomy, Ohio University, Athens, OH 45701, USA

[10]Physics Department, Yale University, P.O. Box 208120, New Haven, CT 06511, USA

[11]Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK

[12]School of Mathematics and Physics, University of Queensland, 4072, Australia

[13]Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400005, India

[14]Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

[15]IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

[16]Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

[17]Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, USA

[18]The Ohio State University, Columbus, 43210 OH, USA

[19]Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA

[20]Leinweber Center for Theoretical Physics, University of Michigan, 450 Church Street, Ann Arbor, Michigan 48109-1040, USA

[21]Department of Physics & Astronomy, University of Rochester, 206 Bausch and Lomb Hall, P.O. Box 270171, Rochester, NY 14627-0171, USA

[22]Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

[23]Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA

[24]Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

[25]Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México

[26]NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA

[27]Department of Physics & Astronomy and Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA

[28]Department of Astronomy, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

[29]Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Menlo Park, CA 94305, USA

[30]Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia

[31]Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia

[32]Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain

[33] Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain

[34] Departament de Física Quàntica i Astrofísica, Universitat de Barcelona, Martí i Franquès 1, E08028 Barcelona, Spain

[35] Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), c. Martí i Franquès, 1, 08028 Barcelona, Spain.

[36] Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA

[37] Department of Astrophysical Sciences, Princeton University, Princeton NJ 08544, USA

[38] Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

[39] Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), FR-75005 Paris, France

[40] Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

[41] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra Barcelona, Spain

[42] Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain

[43] Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA

[44] Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, U.K

[45] Department of Physics & Astronomy, University of Wyoming, 1000 E. University, Dept. 3905, Laramie, WY 82071, USA

[46] National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing, 100012, P.R. China

[47] Departamento de Física, Universidad de Guanajuato - DCI, C.P. 37150, Leon, Guanajuato, México

[48] Instituto Avanzado de Cosmología A. C., San Marcos 11 - Atenas 202. Magdalena Contreras, 10720. Ciudad de México, México

[49] Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

[50] Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

[51] Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada

[52] Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA

[53] Max Planck Institute for Extraterrestrial Physics, Gießenbachstraße 1, 85748 Garching, Germany

[54] Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea

[55] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, VIC 3122, Australia

[56] Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA