

# UC Irvine

## UC Irvine Previously Published Works

### Title

DESI DR2 reference mocks: clustering results from Uchuu-BGS and LRG

### Permalink

<https://escholarship.org/uc/item/888335bf>

### Journal

Journal of Cosmology and Astroparticle Physics, 2026(05)

### ISSN

1475-7516

### Authors

Fernández-García, E

Prada, F

Smith, A

et al.

### Publication Date

2026-05-01

### DOI

10.1088/1475-7516/2026/05/002

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

PAPER • OPEN ACCESS

# DESI DR2 reference mocks: clustering results from Uchuu-BGS and LRG

To cite this article: E. Fernández-García *et al* JCAP05(2026)002

View the [article online](#) for updates and enhancements.

## You may also like

- [ELG×LRG Distribution through Dark Matter Halo Dynamics](#)  
Ginevra Favole, Francisco-Shu Kitaura, Boryana Hadzhiyska et al.
- [HALO OCCUPATION DISTRIBUTION MODELING OF CLUSTERING OF LUMINOUS RED GALAXIES](#)  
Zheng Zheng, Idit Zehavi, Daniel J. Eisenstein et al.
- [Circumgalactic Medium at High Halo Masses—Signatures of Cold Gas Depletion in Luminous Red Galaxies](#)  
Marijana Smailagi, Jason Xavier Prochaska, Joseph Burchett et al.

## DESI DR2 reference mocks: clustering results from UCHUU-BGS and LRG

E. Fernández-García et al.

*Full author list at the end of the paper*

*E-mail:* [efdez@iaa.es](mailto:efdez@iaa.es)

**ABSTRACT:** The aim of this work is to construct mock galaxy catalogues that accurately reproduce the redshift evolution of galaxy number density, clustering statistics, and baryonic properties, such as stellar mass for luminous red galaxies (LRGs) and absolute magnitude in the  $r$ -band for the bright galaxy sample (BGS), based on the first three years of observations from the Dark Energy Spectroscopic Instrument (DESI). To achieve this, we applied the subhalo abundance matching (SHAM) technique to the UCHUU  $N$ -body simulation, which follows the evolution of 2.1 trillion particles within a volume of  $8 h^{-3} \text{Gpc}^3$ , assuming a Planck base- $\Lambda$ CDM cosmology. Using SHAM, we populated UCHUU subhalos with LRGs and BGS-BRIGHT ( $r < 19.5$ ) galaxies up to redshift  $z = 1.1$ , assigning stellar masses to LRGs and luminosities to BGS galaxies (up to  $M_r \leq 20$ ). Furthermore, we analyzed the clustering dependence on stellar mass and luminosity for each tracer. Our results show that the UCHUU BGS-BRIGHT and LRG mocks accurately reproduce the observed redshift evolution of clustering, with better than 5% agreement for separations of  $1 < r < 20 h^{-1} \text{Mpc}$  and below a 10% for  $0.1 < r < 1 h^{-1} \text{Mpc}$ . For the UCHUU-LRG mock, we successfully captured the stellar mass dependence of clustering, while for the UCHUU-BGS mock, we replicated the clustering for various volume-limited subsamples. We also find good agreement between the data and mocks in the dependence of large-scale bias on luminosity for BGS-BRIGHT galaxies and on stellar mass for LRGs. Altogether, these results equip DESI with robust tools for generating high-fidelity lightcones for the remainder of the survey, thereby enhancing our understanding of the galaxy-halo connection.

**KEYWORDS:** cosmological simulations, galaxy clustering, galaxy evolution, power spectrum

**ARXIV EPRINT:** [2507.01593](https://arxiv.org/abs/2507.01593)



---

**Contents**

|          |                                                      |           |
|----------|------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                  | <b>1</b>  |
| <b>2</b> | <b>DESI DR1 and DR2</b>                              | <b>3</b>  |
| 2.1      | DESI DR1                                             | 3         |
| 2.2      | DESI DR2                                             | 4         |
| <b>3</b> | <b>The UCHUU simulation</b>                          | <b>6</b>  |
| <b>4</b> | <b>SHAM method for BGS-BRIGHT and LRG</b>            | <b>7</b>  |
| 4.1      | UCHUU BGS-BRIGHT                                     | 8         |
| 4.2      | UCHUU LRG                                            | 12        |
| 4.3      | Constructing the UCHUU lightcones                    | 13        |
| <b>5</b> | <b>Redshift-space two-point correlation function</b> | <b>14</b> |
| 5.1      | Clustering dependence on luminosity and stellar mass | 19        |
| 5.2      | Redshift space distortions                           | 19        |
| <b>6</b> | <b>Redshift-space power spectrum</b>                 | <b>20</b> |
| <b>7</b> | <b>Large-scale bias</b>                              | <b>22</b> |
| <b>8</b> | <b>Summary</b>                                       | <b>24</b> |
| <b>A</b> | <b>Linear regime of the TPCF</b>                     | <b>27</b> |
| <b>B</b> | <b>Modelling of the colour distribution</b>          | <b>27</b> |
|          | <b>Author List</b>                                   | <b>32</b> |

---

**1 Introduction**

The use of simulated lightcones has become an essential tool in modern cosmology, particularly in the study of large-scale structure (LSS). These simulations aim to replicate the distribution of galaxies and matter in the universe, providing insights into its evolution and underlying physics [1–8]. By generating mock observations that closely resemble real survey data, researchers can validate theoretical models, refine observational strategies, and interpret cosmological measurements.

The concept of simulating cosmic structures traces back to the mid-20th century when early computational efforts sought to understand the dynamics of gravitational clustering. With the advent of large-scale galaxy surveys, such as the Sloan Digital Sky Survey (SDSS) [9] and the Two-Degree Field Galaxy Redshift Survey (2dFGRS) [10], the need for more realistic and comprehensive mock catalogues became apparent. These early surveys revealed the cosmic web — a vast network of galaxies, clusters, and voids — and posed challenges

in modelling the complex interplay of gravitational, hydrodynamical, and astrophysical processes shaping this structure.

As galaxy surveys expanded in scope and precision, the development of simulations advanced in parallel. Numerical cosmological simulations, such as the MILLENNIUM [11] and ILLUSTRIS [12] simulations, have been pivotal in predicting the clustering of dark matter and its connection to galaxy formation. Coupled with methods like the halo occupation distribution (HOD) and subhalo abundance matching (SHAM), these simulations enable the creation of lightcones that mimic observed properties of galaxies, such as luminosity, stellar mass, and clustering statistics.

Subhalo abundance matching (SHAM) [3, 13–19] has emerged as a powerful method for populating dark matter halos with galaxies. By assuming a monotonic relationship between a halo property, such as peak circular velocity or maximum mass (i.e. the peak value of the maximum circular velocity  $V_{\text{max}}$  over the history of the halo), and a galaxy property, such as stellar mass or luminosity, SHAM allows for a straightforward and physically motivated mapping of galaxies onto halos. This method requires minimal assumptions about the underlying physics of galaxy formation, making it particularly attractive for constructing mock catalogues that match observed galaxy statistics. Compared to HOD methods, which model the probability distribution of galaxies within halos as a function of halo mass, SHAM provides a direct galaxy-halo connection that naturally captures environmental dependencies. Moreover, SHAM-based mocks are particularly well suited for generating weak lensing, peculiar velocity, and next-generation survey mocks.

The UCHUU [20] reference mocks we present in this work exemplify the strengths of SHAM-based lightcones by providing high-fidelity galaxy distributions tailored for large-scale structure analyses. These mocks serve as a crucial resource for training and testing covariance mocks, modelling the evolution of clustering with redshift, and capturing the dependence of clustering on baryonic properties. In particular, each Luminous Red Galaxy (LRG) in the mock is assigned a stellar mass, while each Bright Galaxy Sample (BGS) galaxy is assigned an  $r$ -band absolute magnitude and  $g - r$  colour. Furthermore, they provide an accurate determination of halo occupation and bias, ensuring robust interpretations of observed galaxy clustering.

The Dark Energy Spectroscopic Instrument (DESI) [21, 22] represents a major milestone in galaxy surveys, offering an unprecedented view of the universe with millions of spectroscopically confirmed galaxies. To maximize the scientific return of the three years of measurements of DESI, it is crucial to have mock catalogues that replicate the survey’s selection functions and statistical properties. Simulated lightcones for specific galaxy populations, such as LRGs and BGS, play a key role in this endeavour. These mocks allow researchers to: 1) test the robustness of cosmological analyses by providing a controlled environment where the true underlying parameters are known; 2) quantify systematic effects, such as survey incompleteness and redshift errors, which can bias results; 3) optimize measurement techniques, including the estimation of the two-point correlation function and higher-order statistics; and 4) support the development of theoretical models by comparing predictions with observational data.

In this work, we present simulated lightcones designed to mimic the clustering statistics of DESI for Luminous Red Galaxies and the Bright Galaxy Survey from the first three years

of operations (DESI DR2, hereafter). These mocks are constructed using the SHAM method applied to the UCHUU halo simulation [20], one of the largest and most detailed  $N$ -body simulations available. By incorporating realistic number density evolution with redshift and accurately reproducing the two-point correlation function, among other key observables, our simulations bridge the gap between theory and observation, serving as a critical resource for advancing our understanding of the universe.

This article is structured as follows: in section 2, we provide a brief summary of DESI Data Release 1 (DR1) and of the three first years of measurements from DESI, and describe in detail the galaxy populations used in this work. In section 3, we introduce the high-resolution simulation used: UCHUU. In section 4, we present a detailed explanation of the SHAM methodology applied to construct BGS-BRIGHT and LRG mocks, as well as the procedure for generating the lightcone. Then, in section 5, we discuss the two-point correlation function of the mocks, comparing them with measurements from DESI DR2 and Y1 at small scales and also large scales, and also study the dependence of this function with absolute magnitude for different volume-limited samples for the case of the BGS and with different stellar mass for the case of the LRGs. We also study the redshift-space distortions. In section 6 we present the results from the power spectrum, and, finally, in section 7 we analyse the Mean Halo Occupancy and large-scale bias. To end this work, in section 8, we summarize our key findings.

## 2 DESI DR1 and DR2

The Dark Energy Spectroscopic Instrument (DESI) is a robotic, fiber-fed, highly multiplexed spectroscopic surveyor that operates on the Mayall 4-meter telescope at Kitt Peak National Observatory [23]. DESI, which can obtain simultaneous spectra of almost 5000 objects over a  $\sim 3^\circ$  field [22, 24, 25], is currently conducting a five-year survey of about a third of the sky. This campaign will obtain spectra for approximately 40 million galaxies and quasars [21].

The goal of DESI is to determine the nature of dark energy through the most precise measurement of the expansion history of the universe ever obtained [26]. DESI was designed to meet the definition of a Stage IV dark energy survey with only a 5-year observing campaign. Forecasts for DESI [22] predict a factor of approximately five to ten improvement on the size of the error ellipse of the dark energy equation of state parameters  $w_0$  and  $w_a$  relative to previous Stage-III experiments [27].

In this work, we use the DESI Year 1 (Y1) and Year 3 (Y3) catalogues for two main reasons. First, stellar mass estimates based on CIGALE are only available for the Y1 Luminous Red Galaxy (LRG) sample (see [28] for a detailed explanation of the CIGALE tool, and subsection 4.2 for a description of its application in this work). Second, the DESI Collaboration has not yet released pre-reconstruction measurements of the two-point correlation function (TPCF) at Baryon Acoustic Oscillation (BAO) scales for Y3. As a result, comparisons at these scales can only be performed using Y1 data. However, at smaller scales, we carry out the comparison using the Y3 TPCF.

### 2.1 DESI DR1

The DESI Data Release 1 (DR1, DESI DR1 hereafter) dataset [29] contains observations made with the DESI instrument during its first year of operations. These observations occurred

during the main survey operations, which began on May 14, 2021, following a period of survey validation [30], and continued until June 13, 2022. The DESI instrument captures spectra from 5,000 targets simultaneously by using robotic positioners to place optical fibres at the targets’ celestial coordinates [24]. These fibres are grouped into ten “petals”, which direct the light to ten separate climate-controlled spectrographs.

DESI data is gathered through observations of “tiles”, where each tile corresponds to a specific pointing on the sky with a field of view of approximately 8 square degrees. For each tile, targets are assigned to the 5,000 robotic fibres according to a prioritization scheme. DESI allocates its observation time between “bright” and “dark” time programs based on observing conditions. “Bright time” refers to nights with substantial moonlight or twilight periods, while “dark time” refers to nights with little or no moonlight. This distinction is important because the sky background significantly affects the detection of faint objects. DESI also employs a “backup” program during particularly poor observing conditions.

During DR1, 2,744 tiles were observed during “dark” time, and 2,775 tiles were observed during “bright” time. The bright galaxy sample [31, BGS] is observed during bright time, while the luminous red galaxies [32, LRG], quasars [33, QSO], and emission line galaxies [34, ELG] are observed during dark time. This division optimizes the use of telescope time by targeting brighter objects when sky conditions are less favourable and reserving the best observing conditions for fainter, higher-redshift targets. The data were first processed by the DESI spectroscopic pipeline [35] the morning after the observations for immediate quality checks and later underwent a uniform processing run (internally referred to as “iron”) to generate the redshift catalogues used in this paper.

The BGS sample consists of two types of targets: BGS-BRIGHT and BGS-FAINT. The BGS-BRIGHT sample includes galaxies with magnitudes  $r < 19.5$ , while the BGS-FAINT sample covers the range  $19.5 < r < 20.175$  and has also a selection in color to achieve high redshift efficiency [31]. The majority of BGS targets fall into the BGS-BRIGHT category, which is the sample used in this work.

In this work we will focus on the creation of the UCHUU-LRG and UCHUU-BGS lightcones. The methodology used in order to create the UCHUU-ELG and UCHUU-QSO lightcones is presented in [36], as well as the main results from the clustering statistics.

## 2.2 DESI DR2

On the other hand, we will utilize the galaxy catalogues containing data from the first three years of operations, spanning from May 14, 2021, to April 2024, which includes DR1 [37, 38]. This catalogue, referred to as DESI DR2, includes 6,671 tiles observed during ‘dark’ time and 5,171 tiles observed during ‘bright’ time. These numbers represent 2.4 and 2.3 times the number of tiles released in DR1, respectively.

In table 1 we list the basic properties of the LRG and BGS-BRIGHT samples used in this work. This table includes information such as the redshift ranges, sky area, total weighted number of galaxies and effective volume as a measure of constraining power. The effective area is calculated generating  $N_{\text{random}}=1 \times 10^7$  random points and checking how many of them are inside Y1 or Y3 bright (dark) footprint for the case of the BGS-BRIGHT (LRG),

| Sample        | Redshift range   | $z_{\text{med}}$ | $A_{\text{eff}}$ (deg <sup>2</sup> ) | $N_{\text{eff}}$ | $V_{\text{eff}}$ ( $h^{-3}$ Gpc <sup>3</sup> ) |
|---------------|------------------|------------------|--------------------------------------|------------------|------------------------------------------------|
| Y1 BGS-BRIGHT | $0.02 < z < 0.5$ | 0.203            | 7473                                 | 6189576          | 1.73                                           |
| Y3 BGS-BRIGHT | $0.02 < z < 0.5$ | 0.203            | 12416                                | 10069613         | 2.88                                           |
| Y1 LRG        | $0.4 < z < 1.1$  | 0.753            | 5740                                 | 3013218          | 7.98                                           |
| Y3 LRG        | $0.4 < z < 1.1$  | 0.753            | 10077                                | 5172900          | 14.01                                          |

**Table 1.** Basic properties of the DESI BGS-BRIGHT Y1 and Y3 (first and second row, respectively) and LRG Y1 and Y3 (third and fourth row, respectively): the redshift interval, median redshift ( $z_{\text{med}}$ ), effective area of the sky footprint weighted by completeness ( $A_{\text{eff}}$ ), number of galaxies weighted by completeness ( $N_{\text{eff}}$ ), and effective volume ( $V_{\text{eff}}$ ).

| $M_r$ | $z_{\text{min}}$ | $z_{\text{max}}$ | $N_{\text{eff}}$ | $V_{\text{eff}}$ ( $h^{-3}$ Gpc <sup>3</sup> ) |
|-------|------------------|------------------|------------------|------------------------------------------------|
| -22.0 | 0.02             | 0.50             | 311406           | 2.89                                           |
| -21.5 | 0.02             | 0.42             | 852877           | 1.83                                           |
| -21.0 | 0.02             | 0.34             | 1457497          | 1.03                                           |
| -20.5 | 0.02             | 0.28             | 1983958          | 0.60                                           |
| -20.0 | 0.02             | 0.23             | 2120063          | 0.35                                           |

**Table 2.** Minimum and maximum redshift for different volume-limited samples of the BGS survey, as well as the effective number of galaxies ( $N_{\text{eff}}$ ) and the effective volume of each sample.

$N_{\text{inside}}$ .  $A_{\text{eff}}$  is calculated then as

$$A_{\text{eff}} = \frac{N_{\text{inside}}}{N_{\text{random}}} \frac{360^2}{\pi}, \quad (2.1)$$

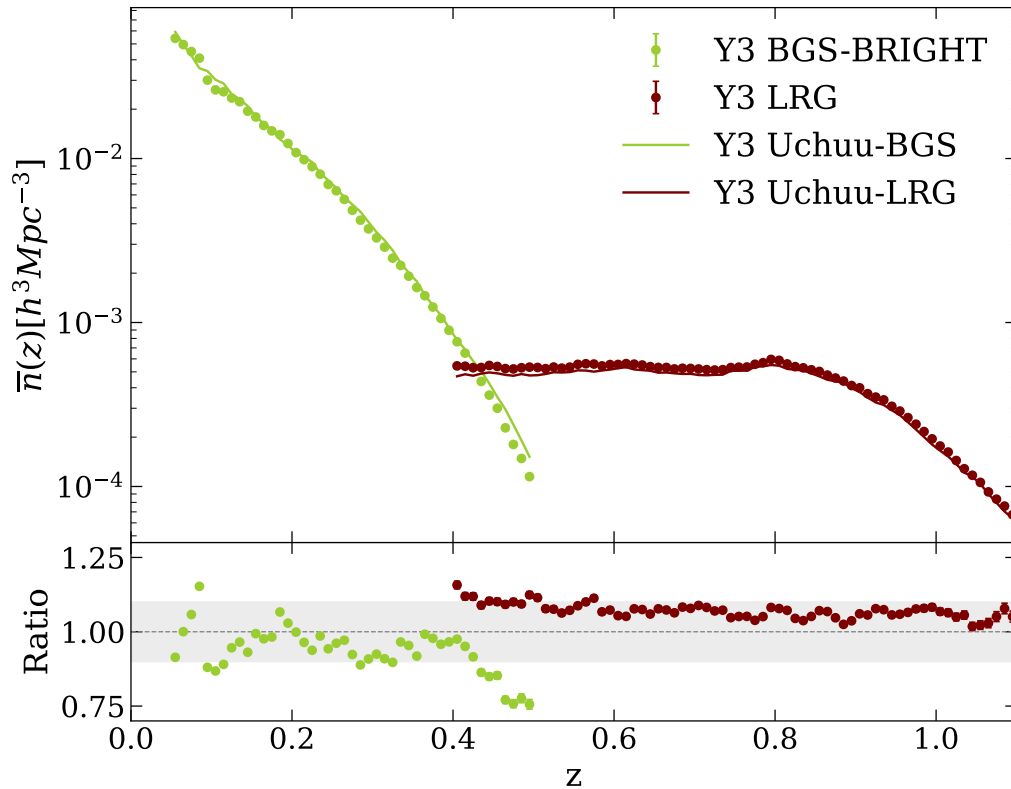
where  $\frac{360^2}{\pi}$  is the area of the full-sky lightcone in square degrees.  $V_{\text{eff}}$  is calculated then as

$$V_{\text{eff}} = \frac{4\pi}{3} \frac{N_{\text{inside}}}{N_{\text{random}}} (r_{\text{max}}^3 - r_{\text{min}}^3), \quad (2.2)$$

where  $r_{\text{min}}$  and  $r_{\text{max}}$  is the comoving distance that corresponds to the maximum and minimum redshift of each target survey (indicated in the second column of table 1). We consider a flat  $\Lambda$ CDM cosmology with Planck-15 parameters [39] when transforming redshift to comoving distances:  $h=0.6774$ ,  $\Omega_{\text{m}}=0.3089$ ,  $\Omega_{\text{b}}=0.0486$ ,  $n_{\text{s}}=0.9667$ ,  $\Omega_{\Lambda}=0.6911$ , and  $\sigma_8=0.8159$ .

The top panel of figure 1 shows the comoving number density for the two Y3 samples. The bottom panel presents the ratio between the data and the mock prediction. It is observed that for redshifts in the range of  $0.1 < z < 0.4$ , the mock prediction is quite accurate, with the ratio remaining below 10%. However, for higher redshifts ( $z > 0.4$ ), the deviation increases to approximately 25%. For the LRGs, the mock slightly underestimates the number density compared to the data, although the ratio remains below 10%, except near  $z \sim 0.4$ .

In figure 2 the BGS number density as a function of redshift can be seen for the different volume-limited samples constructed in this work (see table 2 for details). It can be proved that the mean ratio between the data and the mock is about a 5% for all the volume-limited samples if we take into account only the redshift range where  $n(z)$  is approximately constant.



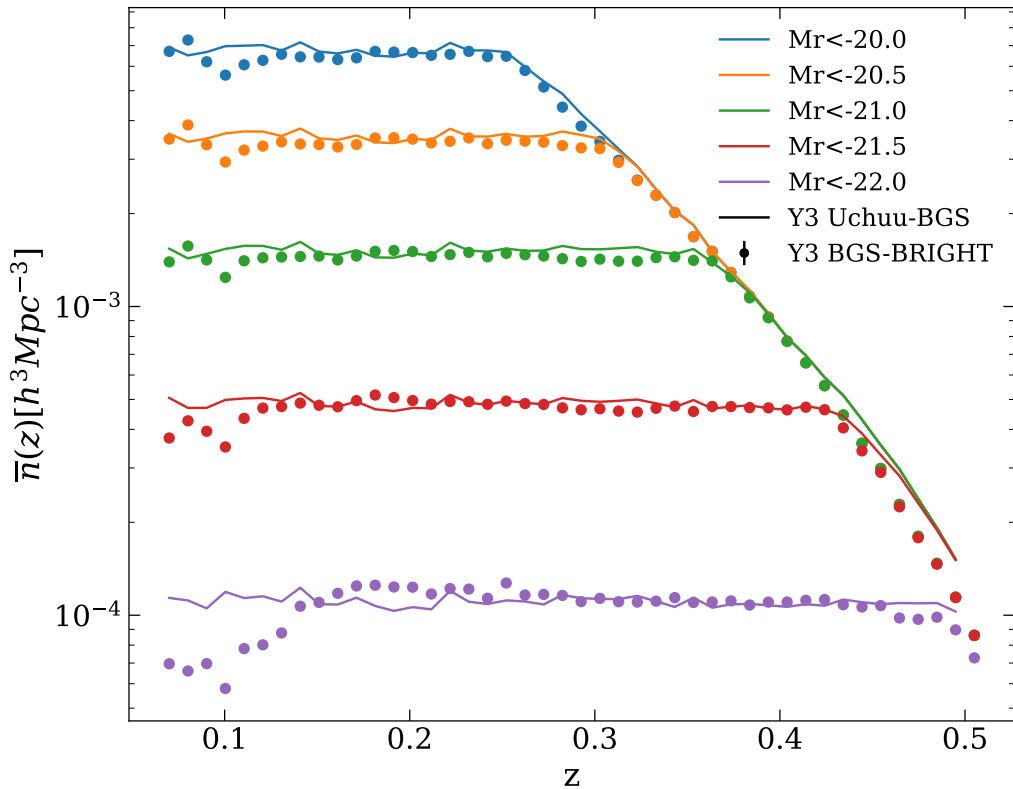
**Figure 1.** The comoving number density corrected by completeness of Y3 LRG and BGS-BRIGHT DESI (points) and Y3 UCHUU mock lightcones (solid lines) is shown in the top panel. The ratio between data and the prediction of the mock for each target is shown in the bottom panel. The shaded area represents the limits of the 10%.

### 3 The UCHUU simulation

In order to model the clustering signal of the DESI DR1 survey within the flat  $\Lambda$ CDM Planck cosmology [40], we utilized the UCHUU  $N$ -body simulation [20]. Designed specifically to model the DESI survey, UCHUU features high numerical resolution, which enables the resolution of low-mass dark matter haloes and subhaloes within its large volume. This capability allows us to apply the SHAM technique to populate UCHUU haloes and subhaloes with DESI galaxies, generating mock lightcones to reproduce the number density and predict the clustering of each DESI tracer.

The UCHUU simulation was run using the TreePM code GREEM [41]. The box has a comoving side length of  $2 \ h^{-1} \text{Gpc}$ , with  $12,800^3$  dark matter particles. The mass resolution and gravitational softening length are  $3.27 \times 10^8 \ h^{-1} \text{M}_\odot$  and  $4.27 \ h^{-1} \text{kpc}$ , respectively. The initial conditions were generated using the second-order Lagrangian Perturbation Theory (2LPT) approximation at  $z_{\text{init}}=127$ , and the simulation followed the growth of cosmic structures in the Planck 2015 flat  $\Lambda$ CDM cosmology [39]. We saved 50 snapshots of the particle distribution from  $z=14$  to  $z=0$ , and identified bound structures using the Rockstar<sup>1</sup>

<sup>1</sup>The minimum number of particles per halo identified by Rockstar is 10, corresponding to the mass of  $3.27e9 \ \text{Msun}/h$ . The force resolution used in Rockstar is the same with the softening length, and the other parameters were set to the default values.



**Figure 2.** Number density as a function of redshift for different volume-limited samples (detailed in table 2 calculated from the UCHUU-BGS lightcone (solid lines) and DESI BGS-BRIGHT Y3 (points).

phase-space halo\subhalo finder [42]. We constructed merger trees for these structures using a parallel version of the Consistent-Trees algorithm [43]. Additionally, we obtained the peak value of the maximum circular velocity,  $V_{\text{max}} = \max\left(\sqrt{\frac{GM(r)}{r}}\right)$ , over the history of each (sub)halo, denoted as  $V_{\text{peak}}$ . We measure the maximum circular velocity at each of the 50 redshift outputs, and take the maximum value as  $V_{\text{peak}}$ . We used this to implement the SHAM method for populating UCHUU haloes with DESI galaxies. For more information of the simulation methodology and performance, we refer the reader to [20]. All UCHUU data products are publicly available through Skies & Universes.<sup>2</sup>

#### 4 SHAM method for BGS-BRIGHT and LRG

In this section we explain the SHAM method used for BGS-BRIGHT and LRG tracers to assign absolute magnitudes and stellar masses, respectively. The choice of different target properties reflects the distinct nature of these galaxy populations: BGS-BRIGHT is a flux-limited sample where luminosity is the primary observational constraint, while LRG is a color-selected sample where stellar mass better characterizes the underlying galaxy population and its evolution.

<sup>2</sup><https://www.skiesanduniverses.org/Simulations/Uchuu>.

## 4.1 UCHUU BGS-BRIGHT

### 4.1.1 SHAM Procedure and target property choice

We follow the SHAM methodology from [8] to construct the UCHUU Y3 BGS-BRIGHT lightcones because:

1. BGS-BRIGHT is a magnitude-limited survey with a flux cut at  $r=19.5$ ;
2. luminosity directly relates to the observational selection function;
3. and the magnitude-limited nature makes absolute magnitude the most physically motivated property for abundance matching.

This approach has previously been applied to BOSS [17], SDSS [8] and the DESI Early Data Release [6, 44]. SHAM assumes that the most massive (sub)halos host the most luminous galaxies. To generate our simulated flux-limited BGS-BRIGHT sample, we use the luminosity function derived from Y1 DESI BGS-BRIGHT galaxies [see 45, for more details]. However, we only consider the luminosity function derived from the North Galactic Cap (NGC), as the one derived from the South Galactic Cap (SGC) systematically underpredicts the observed number density of BGS-BRIGHT galaxies in that region. This discrepancy likely originates from uncertainties in the luminosity computation, such as differences in photometric calibration or  $k + E$  corrections between the two regions.

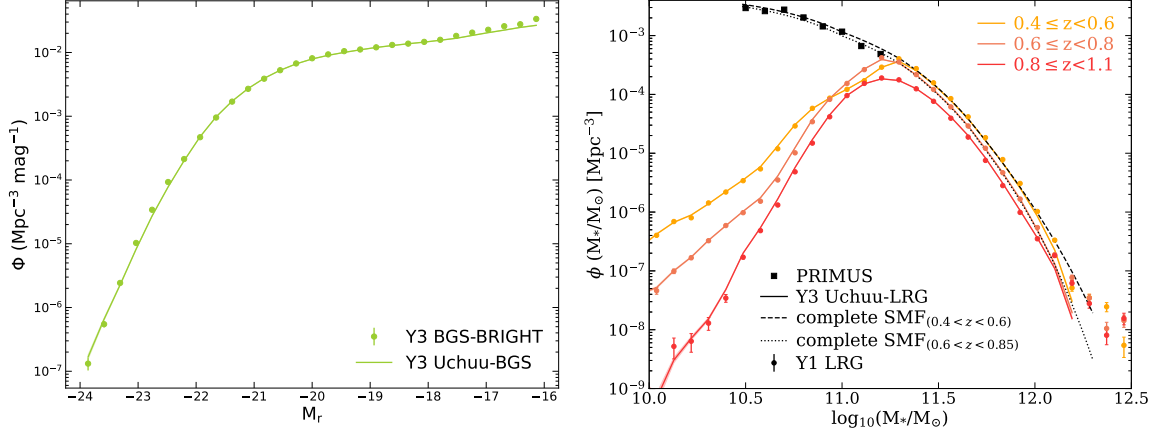
This luminosity function is  $k + E$  corrected, allowing us to apply the same function consistently across all UCHUU simulation boxes, each corresponding to a different redshift. The  $k$ -corrections are computed using polynomial fits based on the FastSpecFit  $k$ -correction catalogue from DESI [46], with a reference redshift of  $z_{\text{ref}} = 0.1$ . These corrections transform absolute magnitudes into the rest-frame SDSS  $r$ -band system,  $z = 0.1$ , ensuring an homogeneous comparison of galaxies across different redshifts. Additionally, we correct for passive evolution using the model of [47], where the evolution correction  $E(z)$  is parametrized as  $E(z) = -Q(z - z_{\text{ref}})$ , with  $Q$  determined by minimizing the  $V/V_{\text{max}}$  distribution to ensure an optimal fit to the observed luminosity evolution, where  $V_{\text{max}}$  is the maximum volume over which the galaxy could be seen inside the volume of the survey up to the observed redshift of the galaxy.

The luminosity functions from Y3 BGS-BRIGHT and UCHUU-BGS are shown in the left panel of figure 3. Uchuu-BGS luminosity function is compatible within  $1\sigma$  with Y3 BGS-BRIGHT luminosity function.

### 4.1.2 Step-by-step SHAM implementation

A summary of the SHAM method for BGS is the following:

1. Sort the haloes by  $V_{\text{peak}}$  in decreasing order, to compute the cumulative number density function  $n_h(> V_{\text{peak}})$ .
2. Assign an ‘unscattered’ galaxy  $r$ -band magnitude,  $^{0.1}M_h^r$ , to each halo by matching the above  $V_{\text{peak}}$  cumulative halo number density function to the target galaxy luminosity



**Figure 3.** Left: Luminosity function from Y3 NGC BGS-BRIGHT (green points) and UCHUU-BGS (green line). The shaded area and the error bars indicates the  $1\sigma$  uncertainty. Right: stellar mass function of Y1 LRG DESI (points) and Y1 and Y3 UCHUU-LRG lightcones (dotted and solid green lines) are shown for several redshift bins within the range  $0.4 < z < 1.1$ . The dashed and dotted black lines represent the complete SMF adopted in each redshift range, indicated in the legend.

function, preserving the ranking such that the large  $V_{\text{peak}}$  haloes host high-luminosity galaxies:

$$n_g(< {}^{0.1}M_h^r) = n_h(> V_{\text{peak}}). \quad (4.1)$$

3. Define a new ‘scattered’ value of the absolute magnitude,  ${}^{0.1}M_h^{r,\text{scat}}$ :

$${}^{0.1}M_h^{r,\text{scat}} = \mathcal{N}(0, \sigma^2) + {}^{0.1}M_h^r, \quad (4.2)$$

where  $\mathcal{N}(0, \sigma^2)$  is a normal distribution with median equal to zero and standard deviation  $\sigma$ , which is the only free parameter of this model.

4. Sort the haloes by  ${}^{0.1}M_h^{r,\text{scat}}$ , and compute the  ${}^{0.1}M_h^{r,\text{scat}}$  cumulative distribution,  $n_h(< {}^{0.1}M_h^{r,\text{scat}})$ .
5. Assign the final  $r$ -band magnitude by matching the cumulative distribution of  ${}^{0.1}M_h^{r,\text{scat}}$  to the target cumulative distribution function,

$$n_g(< {}^{0.1}M_h^r) = n_h(> {}^{0.1}M_h^{r,\text{scat}}). \quad (4.3)$$

The scatter implementation preserves the target luminosity function by design. After introducing scatter in step 3, we perform a re-ranking and re-assignment in step 4 that explicitly matches the scattered halo distribution back to the original target luminosity function. This ensures that the final galaxy population exactly reproduces the input luminosity function, while the scatter affects only the specific halo-galaxy assignments, not the marginal luminosity distribution.

We have verified this preservation by comparing the resulting luminosity function from UCHUU-BGS with the input Y3 BGS-BRIGHT luminosity function (left panel of figure 3), showing excellent agreement within  $1\sigma$  uncertainties.

Also, it is important to mention that fitting the scatter so that we match the monopole of the 2PCF at small scales<sup>3</sup> may mask the different cosmology between the data and the mocks. If the real cosmology of the data were approximately the cosmology from Planck15 or slightly different, then the scatter will absorb this difference. However, if the real cosmology were too different from Planck18, then one would expect that the shape of the 2PCF would be different (the peak of the BAO would be shifted, for example, and the scatter has no effect at these scales), and then the scatter wouldn't be enough in order to reproduce the 2PCF from the data.

#### 4.1.3 Luminosity-dependent scatter

In this work we have found that  $\sigma$  is a function of the luminosity. The value of the scatter is 0.05 for galaxies with  $^{0.1}M_h^r > -20.5$ , and 0.4 for  $^{0.1}M_h^r \leq -20.5$ . This dependence of the scatter with the absolute magnitude has been calibrated matching UCHUU-BGS TPCF with the one from DESI BGS-BRIGHT. However, we have tested it only for galaxies brighter than  $^{0.1}M_h^r < -20.0$ . For future works, we will improve the dependence of the scatter with the absolute magnitude so that we can find a functional form, and also reach the faintest galaxies. However, the dependence of the scatter with luminosity obtained in this work is the one expected from previous works, such as [48].

For clarity and readability, we will omit the superscript 0.1 and the subscript  $h$  in  $^{0.1}M_h^r$  from this point onward. The superscript 0.1 indicates that the  $k$ -corrections were computed using a reference redshift of  $z_{\text{ref}} = 0.1$ , while the subscript  $h$  denotes that the absolute magnitudes have been corrected for evolution.

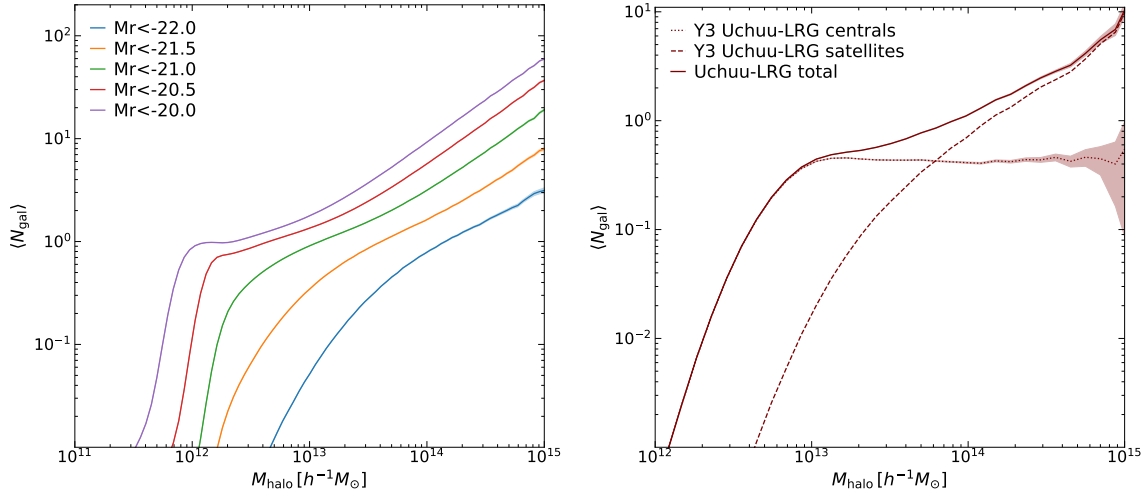
#### 4.1.4 Flux-limited sample construction

The last important step is to remove all those galaxies with apparent  $r$ -band magnitude fainter than  $r = 19.5$ . We have to do this because, as mentioned above, DESI BGS-BRIGHT is magnitude-limited. In order to do this, we need to transform absolute magnitudes to apparent magnitudes, and for that we need to assign a colour for each galaxy (in order to remove the  $k + E$  corrections). Following the methodology described in [49], the rest-frame  $g - r$  colour distribution is modelled as a double-Gaussian function, with parameters for the mean and dispersion of both the red and blue sequences, as well as the fraction of blue galaxies. Unlike previous implementations based on SDSS data, this approach employs a smoothly broken linear function to more accurately represent the colour distribution across different magnitudes. These colour distributions are evolved with redshift using the convention of [47] to ensure consistency with DESI BGS-BRIGHT Y1 observations. A more detailed explanation of how the evolution correction is calculated can be seen in [45]. Additionally, the modelling of colour differences between central and satellite galaxies is refined, improving the overall agreement with observed clustering properties. This enhancement leads to a more realistic colour-dependent clustering behaviour, better aligning the mock catalogue with empirical measurements from DESI BGS-BRIGHT Y1.

However, caution must be taken when using these colours. In appendix B, we present a comparison of the probability density functions for the colour distribution in the mock data, the observational data, and the adopted model. This comparison, shown in figure 14, is

---

<sup>3</sup>It can be checked that changing the value of the scatter only affects the 2PCF of to scales of  $20\text{--}30\text{ Mpc}h^{-1}$ .



**Figure 4.** Mean halo occupancy of BGS volume-limited samples (left panel) and LRG samples with stellar mass cuts (right panel) as determined from our SHAM UCHUU boxes. The redshift of the UCHUU-BGS box is  $z = 0.19$ , while the redshift of the UCHUU-LRG box is  $z = 0.78$ . The mean number of galaxies of a halo with a given mass  $M_{\text{halo}}$  is denoted by  $\langle N_{\text{gal}} \rangle$ . The solid lines represent the combined centrals and satellite occupation. The shaded area indicates the  $1\sigma$  uncertainty of the occupation measured from the UCHUU lightcone.

performed across different bins of magnitude and redshift. By construction, the agreement between the mock data and the model is excellent. However, discrepancies emerge when comparing the model to observational data, especially for the faintest galaxies and in the tails of the distribution (i.e., the reddest and bluest galaxies). These mismatches arise from the model’s assumption of a double Gaussian fit, which does not always provide a good approximation in all magnitude-redshift bins. Moreover, while the current methodology allows some flexibility in assigning colours differently to central and satellite galaxies, it does not account for the possibility of central galaxy assembly bias. Investigating this limitation will be an important step toward improving the realism of future mock catalogues.

The sub(halo) boxes we use to construct the UCHUU-BGS lightcone are those with redshift 0.49, 0.43, 0.36, 0.19, 0.093 and 0. As already mentioned above, the luminosity function considered in this work is  $k + E$  corrected. This allows us to apply the same function consistently across all the snapshots.

In the left panel from figure 4 the halo occupancy distribution (HOD) from the UCHUU-BGS for different volume-limited samples constructed from the galaxy box with  $z = 0.19$  is shown. We compare the DR2 UCHUU-BGS HOD with that presented in [6]. For the brightest volume-limited sample, the agreement between the two HODs is excellent, while larger discrepancies are observed for the fainter volume-limited samples. These differences are likely driven by a combination of effects. In particular, the relatively small volume of the One-Percent Survey makes its HOD measurements more susceptible to cosmic variance, especially for satellite galaxies. In addition, methodological differences may also play a role: while [6] applies a single scatter in  $V_{\text{peak}}$  to the full sample, in this work we allow the scatter to vary with magnitude, which can naturally impact the inferred HODs.

| $> \log M^*$ | $0.4 < z < 0.6$ | $0.6 < z < 0.8$ | $0.8 < z < 1.1$ |
|--------------|-----------------|-----------------|-----------------|
| 11.6         | 0.62            | 0.58            | 0.30            |
| 11.4         | 0.66            | 0.37            | 0.70            |
| 11.3         | 0.29            | 0.33            | 0.55            |
| 11.2         | 0.14            | 0.1             | 0.16            |
| 11.1         | 0.14            | 0.1             | 0.16            |
| 10.8         | 0.14            | 0.1             | 0.16            |

**Table 3.** Value of the scatter found for each cumulative cut in the stellar mass (first column) for each of the three redshift ranges:  $0.4 < z < 0.6$  (second column),  $0.6 < z < 0.8$  (third column) and  $0.8 < z < 1.1$  (fourth column).

## 4.2 UCHUU LRG

### 4.2.1 SHAM procedure and target property choice

For LRG, we use stellar mass as the target property because the LRG sample is designed to be complete only for massive, evolved galaxies. Using the luminosity function would result in a severely incomplete distribution across the full luminosity range. In contrast, while the stellar mass function is incomplete for low-mass galaxies, it remains complete for the high-mass regime that defines the LRG population. This completeness at high masses allows us to reliably model the incompleteness at lower masses, making stellar mass a more practical and robust choice for abundance matching LRGs.

To construct the UCHUU-LRG lightcones, we follow the SHAM approach introduced in section 4.1 of [3], which has been previously applied to the BOSS survey and the DESI Early Data Release [6]. This method assumes that the most massive galaxies are hosted by the most massive (sub)haloes. The main equation of the SHAM method is

$$n_g(> M^*) = n_h(> V_{\text{peak}}^{\text{scatter}}), \quad (4.4)$$

with

$$V_{\text{peak}}^{\text{scatter}} = (1 + \mathcal{N}(0, \sigma^2))V_{\text{peak}}, \quad (4.5)$$

where  $V_{\text{peak}}$  is the maximum circular velocity over the halo history,  $\mathcal{N}(0, \sigma^2)$  is a normal distribution with mean equal to zero and standard deviation  $\sigma$ , which is the only free parameter of the model.  $n_g$  is the theoretical (complete) stellar mass function from the LRG population.

In this work, we found that the optimal value of  $\sigma$  depends on the redshift and on the stellar mass. This dependence was found matching the mock two-point correlation function (TPCF) with that from the data split in the redshift ranges  $0.4 < z < 0.6$ ,  $0.6 < z < 0.8$  and  $0.8 < z < 1.1$  and also doing cuts in stellar mass ( $\log M^* > 10.8, 11.2, 11.3, 11.4$  and  $11.6$ ). The optimal values of the scatter with the stellar mass and redshift can be seen in table 3.

### 4.2.2 Stellar mass function and completeness modelling

In the right panel from figure 3 we present the stellar mass function of LRG obtained from the Y1 DESI survey using the CIGALE tool [50] to estimate individual stellar masses [see 28, for more details]. We had to use the SMF from Y1 DESI LRG as there are no estimations

for the stellar masses of Y3 LRG so far. However, we don't expect a big change in the SMF from Y1 LRG and Y3 LRG.

With this estimations of the stellar masses, we are able to calculate the complete LRG stellar mass function (SMF) at high masses, but we lack information on the shape of the SMF at low masses due to the selection function. To supplement our analysis, we incorporate the SMF measurements obtained from PRIMUS presented by [51], which is shown by the square symbols in the right panel from figure 3. It is worth noting that we do not consider the redshift evolution of the PRIMUS SMF in our analysis, as it has been shown to have a negligible impact on our results and is consistent with the findings of the PRIMUS survey.

To account for the observed evolution in the shape of the SMF with redshift (as shown in the right panel from figure 3), we have employed two different complete SMF fits in our SHAM method: one that characterizes a complete galaxy population at  $0.4 \leq z \leq 0.6$  (low- $z$ , represented by the dashed line), and another that characterizes a complete population at  $0.6 \leq z \leq 0.85$  (high- $z$ , represented by the dotted line).

We apply SHAM to the (sub)halo snapshots from UCHUU boxes at redshifts 0.49, 0.56, 0.63, 0.70, 0.78 and 0.86, 0.94 and 1.03 to cover interval  $0.4 \leq z \leq 1.1$ . As we only have the complete stellar mass function calculated at  $0.4 < z < 0.6$  and  $0.6 < z < 0.85$ , and we have halo boxes with redshift beyond  $z = 0.85$ , we apply the complete stellar mass function that corresponds to the range  $0.6 < z < 0.85$  to the two halo boxes with redshift  $z > 0.85$ .

#### 4.2.3 Incompleteness implementation: detailed procedure

It is important to remark that one additional step is needed in order to reproduce the observed stellar mass function (points) from figure 3: we need to randomly remove some galaxies, specially low-mass galaxies, so that we can reproduce the incompleteness that comes from observations. This random downsampling is done for each mass bin in the observed stellar mass function and in small redshift ranges (from 0.4 to 1.1 in steps of 0.01). This probabilistic treatment preserves the clustering properties of the high-mass population while realistically representing the observational limitations. Also, matching the observed stellar mass function we guarantee that the observed number density is also matched.

In the right panel from figure 4 the halo occupancy distribution (HOD) from UCHUU-LRG galaxy box at redshift  $z = 0.78$  is shown. Again, we can compare the HOD from this work with that presented in [6]. It can be checked that the HOD for central galaxies agrees perfectly well within  $1\sigma$  with that from [6] for the most massive haloes, although there is a big discrepancy for the least massive part of the HOD and for the satellite HOD. Therefore, the full HOD doesn't agree with the HOD from [6]. As in the case for the BGS, the main difference between the methodology in this work and in [6] is that in the latter work a single scatter in  $V_{\text{peak}}$  is applied, whereas in this work we apply different scatters for different galaxies depending on the mass. This could explain the difference observed in the HOD.

#### 4.3 Constructing the UCHUU lightcones

After applying the SHAM method to populate the simulation catalogues with galaxies, we generate UCHUU-DESI lightcones for each tracer by assembling cubic simulation boxes into spherical shells [see 6, 49, for a detailed explanation of this method].

This approach was applied to both DESI tracers considered in this work to construct their respective full-sky lightcones. The lightcones were built using shells with a redshift width of 0.05, selecting the simulation box with a redshift closest to the mean redshift of each bin.

The shell assignment process means that the same simulation snapshot may be used for multiple adjacent shells when it represents the closest available redshift. For instance, a simulation box at  $z=0.49$  might be assigned to shells covering both  $0.45 < z < 0.50$  and  $0.50 < z < 0.55$ . This results in the same galaxies appearing in the two shells, which is a necessary consequence of using finite box sizes to construct continuous lightcones. However, this overlap preserves large-scale structure continuity and does not significantly bias our clustering statistics, as the measurements average over large volumes and many galaxy pairs.

The lightcones were then cut to match the northern and southern areas of the Y1 and Y3 DESI Survey footprints. In this study, we retained all objects within the survey footprint, regardless of completeness of the footprint, for both tracers.

Figure 5 illustrates the incredible scope and detail of the data included in Uchuu-DESI DR2. In this plot a two-dimensional projection of redshift and right ascension for a narrow wedge of the DESI DR2 footprint ( $\pm 5$  degrees in declination) is shown, unveiling the large-scale structure traced by the BGS, LRG, ELG, and QSO targets out to  $z \approx 2$ .

## 5 Redshift-space two-point correlation function

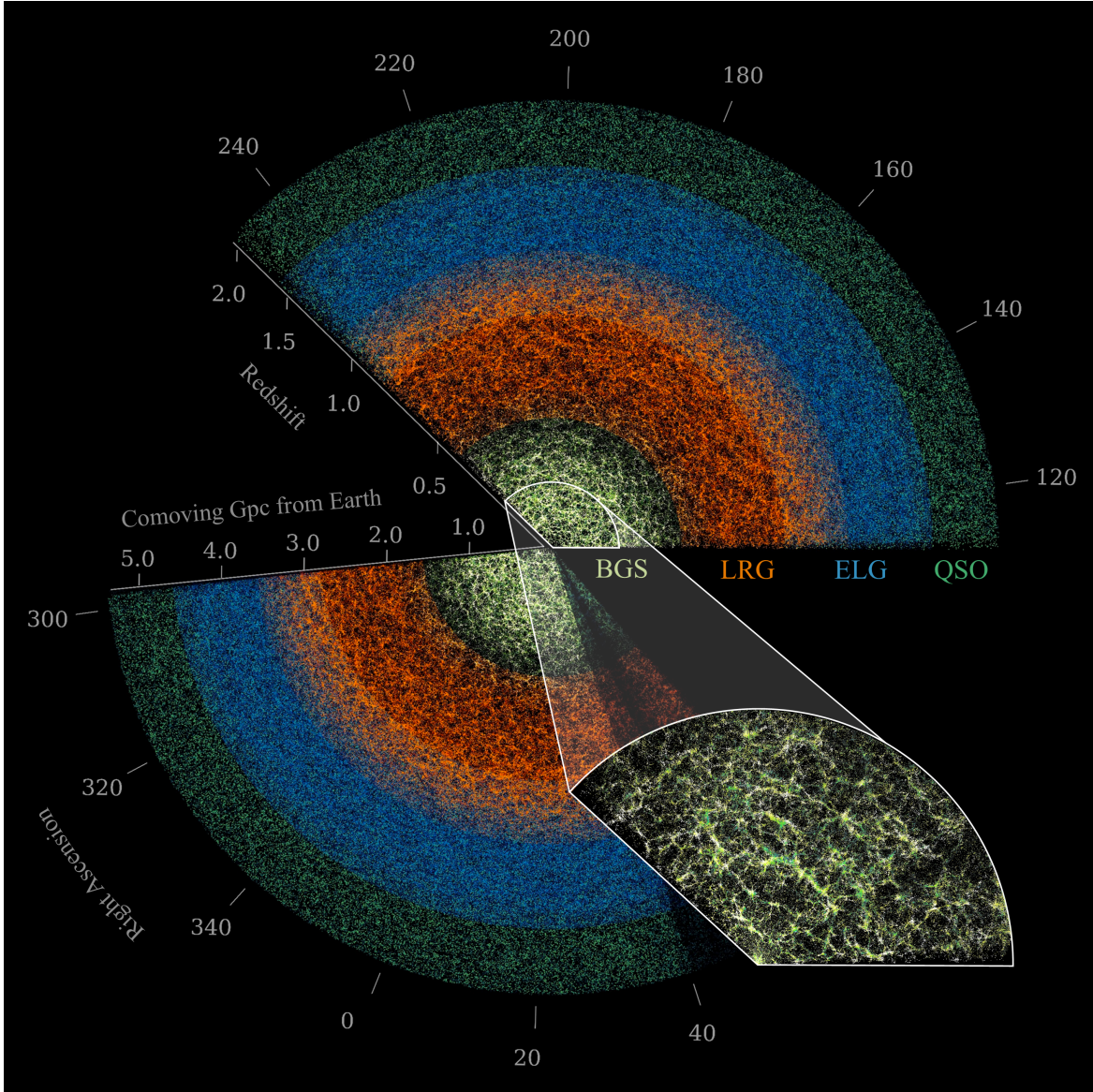
In this section, we compare the two-point correlation function (TPCF) measured for each of the galaxy samples in the DESI DR2 Survey [52] with that predicted by the Planck cosmology using our UCHUU Y3 mock lightcones from LRG and BGS-BRIGHT, as described in the previous section. Additionally, we explore the dependence of the galaxy clustering on luminosity and stellar mass for BGS and LRG galaxies, respectively.

The LRG data is primary weighted using the Pairwise Inverse Probability (PIP) weights [53], which reflect whether pairs of galaxies would have been observed in 128 alternate realizations of DESI, in addition to which pairs are observed in the actual survey [54]. The PIP weights are combined with the FKP weights [55],  $w_{\text{FKP}}$ , to account for the inhomogeneous sampling density of the data sample, which are defined as

$$w_{\text{FKP}} = \frac{1}{1 + P_0 n(z, n_{\text{tile}})} \quad (5.1)$$

where  $n(z, n_{\text{tile}})$  is the weighted number density,  $n_{\text{tile}}$  is the number of overlapping tiles, and  $P_0$  is the power spectrum value at a scale of  $k = 0.15 \, h\text{Mpc}^{-1}$  ( $7000 \, h^{-3} \text{Mpc}^3$  for BGS and  $10000 \, h^{-3} \text{Mpc}^3$  for LRG [44]).

FKP weights are designed to optimize the signal-to-noise ratio by down-weighting high-density regions and up-weighting low-density ones, effectively flattening the redshift distribution  $n(z)$ . In the case of the BGS-BRIGHT sample (with  $M_r < -20.0$ ), which is flux-limited,  $n(z)$  varies significantly with redshift. While FKP weights can still be applied in such cases to mitigate this variation, doing so would alter the effective  $n(z)$  of the data relative to that of the mocks, where no FKP weights have been applied. To maintain consistency between data and simulations and avoid introducing artificial differences in clustering — especially



**Figure 5.** A slice of the universe mapped by Uchuu DR2 drawn from a small wedge of the DESI footprint between  $\pm 5$  degrees in declination out to  $z \approx 4$ . We render the four major extragalactic samples — bright galaxy survey (BGS) galaxies, luminous red galaxies (LRG), emission-line galaxies (ELG), and QSOs — using yellow, orange, blue, and green points, respectively. Within each target class, the shade of the colour maps to declination (lighter colours correspond to higher declination). The inset shows a subset of the BGS survey extending out to redshift  $z = 0.2$ , highlighting the large-scale structure traced by galaxies in the densest survey region. For reference, this small wedge of the BGS survey represents less than 0.1% of the comoving cosmological volume in DR2.

at high redshift where  $n(z)$  is lower — we do not apply FKP weights when analysing the full, flux-limited sample.

We use the Landy-Szalay [56] estimator to measure the two-dimensional correlation function,  $\xi(s, \mu)$ , where  $s$  is the pair separation in units of  $h^{-1}$  Mpc, and  $\mu$  is the cosine of the angle between the separation vector and the line of sight.

For large-scale clustering analysis, we use linearly spaced separation bins from 50 to 200  $h^{-1}$  Mpc with a bin width of 4  $h^{-1}$  Mpc. For small-scale studies, we adopt 49 logarithmically spaced bins between 0.01 and 100  $h^{-1}$  Mpc, although our plots primarily focus on the smaller scales. However, the full measurement range will be publicly available. In both cases, we use 30 bins for  $\mu$ , spanning from -1 to 1.

Additionally, when studying the Baryonic Acoustic Oscillation (BAO) scales we will compare the mock predictions with the measurements from DESI DR1, and when studying the smaller scales we will compare the mock prediction with the measurements from DESI DR2.

We measure the monopole and quadrupole, which are the first non-zero Legendre multipoles of the redshift-space TPCF. To account for the selection function, we generate our random samples that are 16 times larger than the data and use them to estimate the data-random and random-random pair counts for each tracer. We estimate the TPCF using the Python package PYCORR<sup>4</sup> [57, 58], which is a wrapper for correlation function estimation wrapping a modified version of CORRFUNC [57]. For the UCHUU lightcones, we create a different set of uniform randoms for each tracer, which match the same footprint as the mocks.

The left panel of figure 6 shows the monopole (top) and quadrupole (bottom) of the two-point correlation function (TPCF) for Y3 UCHUU-BGS (lines) and Y3 BGS-BRIGHT (points).

For the monopole, the mean deviation between the mock and the data is approximately 10% for scales in the ranges  $1 < s < 5 h^{-1}$  Mpc and  $5 < s < 20 h^{-1}$  Mpc, decreasing to 5% for  $0.1 < s < 1 h^{-1}$  Mpc.

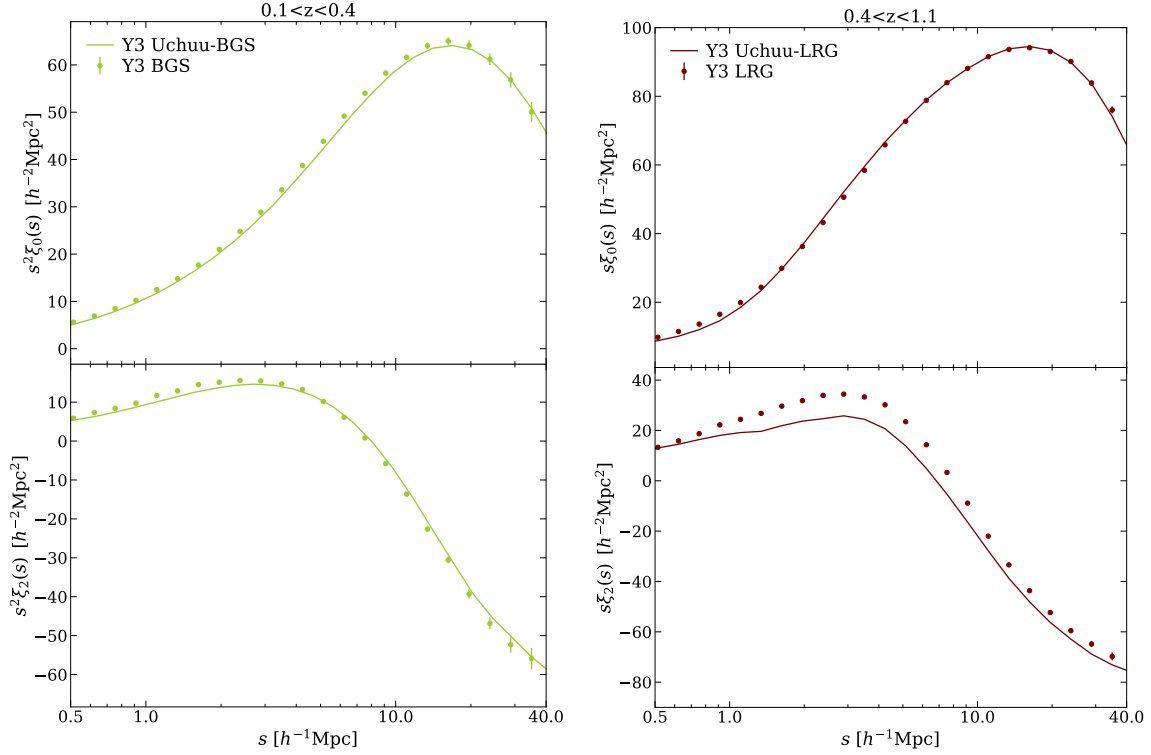
For the quadrupole, the mean deviation is around 4% for  $1 < s < 5 h^{-1}$  Mpc, decreasing to 2% for  $5 < s < 20 h^{-1}$  Mpc, but increasing to about 5% for the smallest scales,  $0.1 < s < 1 h^{-1}$  Mpc.

In the right panel of figure 6, we show the two-point correlation function (TPCF) for the full sample of luminous red galaxies (LRGs) with redshifts in the range  $0.4 < z < 1.1$ . The monopole (upper panel) displays excellent agreement between the mocks and the data, with a mean deviation of only 0.2% across the scales  $1 < s < 5 h^{-1}$  Mpc and  $5 < s < 20 h^{-1}$  Mpc. At smaller scales,  $0.1 < s < 1 h^{-1}$  Mpc, the deviation increases to around 10%.

In the lower panel, which shows the quadrupole, a noticeable discrepancy appears at scales larger than  $s = 1 h^{-1}$  Mpc. This mismatch is very likely driven by stellar-mass incompleteness in the sample, which affects in particular the less massive galaxies and has a direct impact on redshift-space clustering. This interpretation is supported by the fact that the agreement in the quadrupole improves significantly for more massive samples, as shown below.

In addition, the wide redshift bins used in the analysis — especially at high redshift, where the redshift distribution  $n(z)$  departs significantly from being constant — can further amplify the effect of stellar-mass incompleteness on the quadrupole. It is worth noting that the agreement in the lower redshift bin,  $0.4 < z < 0.6$ , is very good, and that the

<sup>4</sup><https://github.com/cosmodesi/pycorr>.



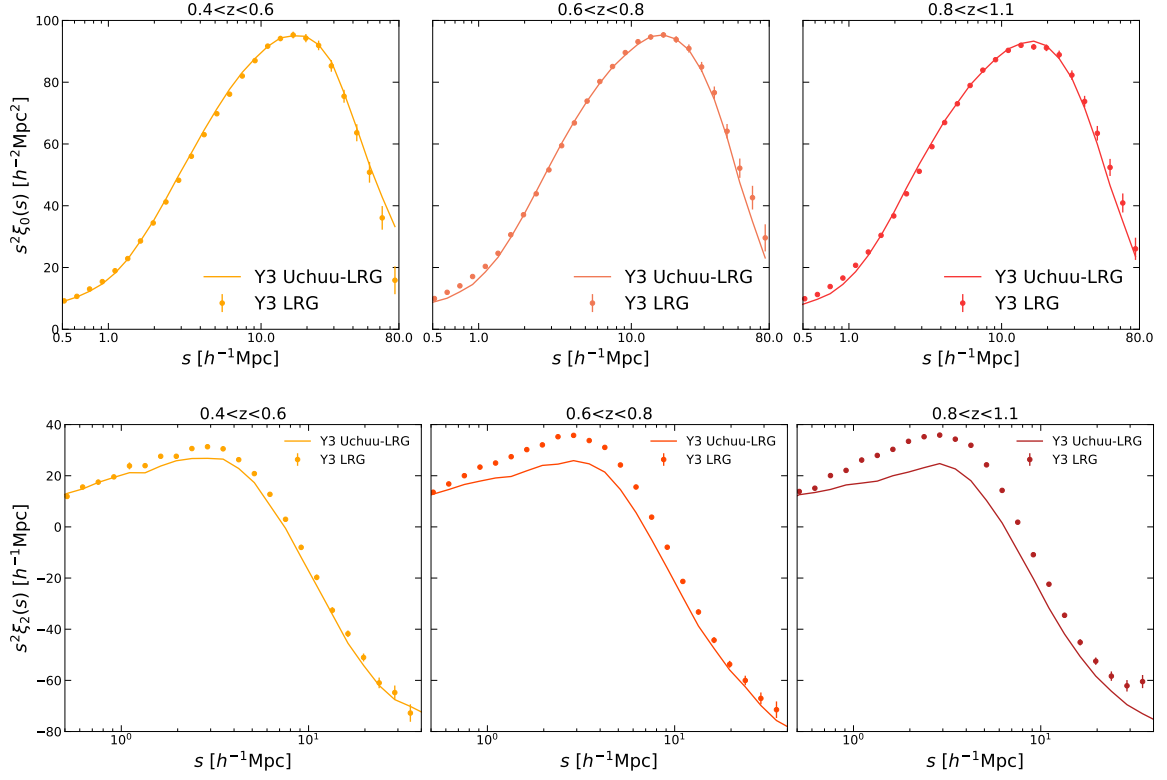
**Figure 6.** The monopole of the two point correlation function obtained for UCHUU-BGS (UCHUU-LRG) and BGS-BRIGHT (LRG) from Y1 in are shown with solid line and points, respectively, in the left (right) plot. In the bottom panels of both plots the quadrupole can be seen.

quadrupole measurements for the various volume-limited samples considered in this work for BGS-BRIGHT also show good agreement.

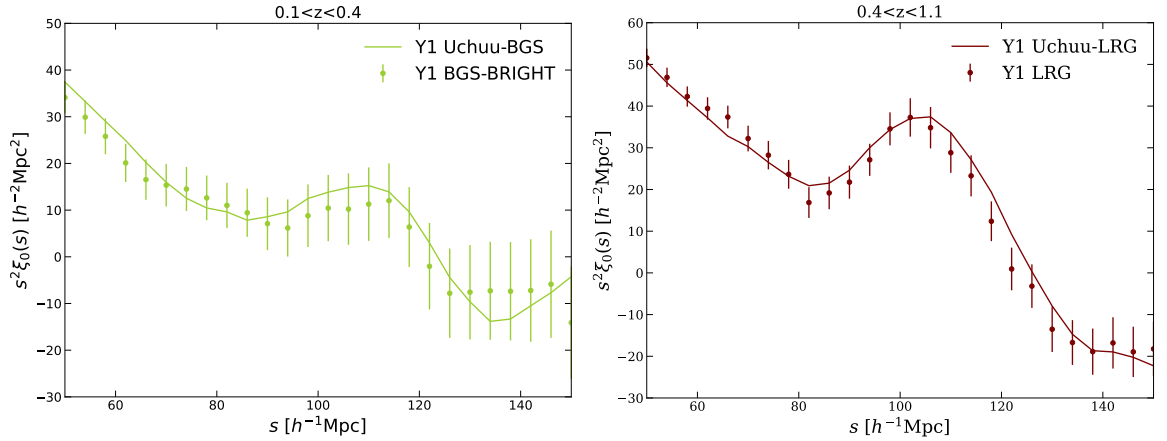
We can also study the dependence of the TPCF with redshift. This can be seen in figure 7, where the monopole of the TPCF from Y3 UCHUU-LRG (lines) and Y3 LRG (points) for the redshift ranges  $0.4 < z < 0.6$  (first row),  $0.6 < z < 0.8$  (second row) and  $0.8 < z < 1.1$  (third row) is shown. From these figures, one can see that the difference in the monopole between the measurement from the data and the prediction from the mock is below a 2% for the three redshift ranges for  $0.1 < s < 5 \ h^{-1} \text{ Mpc}$ , below a 1% for  $5 < s < 20 \ h^{-1} \text{ Mpc}$  and about a 15% for  $0.1 < s < 1 \ h^{-1} \text{ Mpc}$ .

In the bottom panels from the same figure, the quadrupole in each of the three redshift ranges is shown. It can be seen that the discrepancy between the mock and the data seen in figure 6 comes from the two higher redshift ranges.

Next, we focused on large scales to study the Baryonic Acoustic Oscillation (BAO) feature in the monopole of the TPCF. A comparison of the TPCF for large scales is shown in the left (right) panel of figure 8 for the BGS-BRIGHT (LRG) tracers. For both tracers, we observe that the agreement between the data and the mock prediction is quite good, and they are compatible within  $1\sigma$ .



**Figure 7.** The monopole (quadrupole) of the two point correlation function obtained for UCHUU-LRG (solid lines) and LRG (points) Y1 is shown in the top (bottom) panel for three different redshift ranges ( $0.4 < z < 0.6$ ,  $0.6 < z < 0.8$ , and  $0.8 < z < 1.1$ ).



**Figure 8.** The monopole of the two point correlation function obtained for UCHUU-BGS (solid lines) and BGS-BRIGHT (points) Y1 is shown for the redshift range  $0.1 < z < 0.4$  in the left panel. In the right panel the same can be seen for UCHUU-LRG lightcone and LRG Y1 in the redshift range  $0.4 < z < 1.1$ .

### 5.1 Clustering dependence on luminosity and stellar mass

Now, we can study the dependence of the TPCF with cuts in absolute magnitude (BGS-BRIGHT) and stellar mass (LRG). As already mentioned, matching the dependence of the TPCF from the mocks to the measurements from data is accomplished by introducing a dependence of the scatter (see eqs. (4.2) and (4.5)) with the stellar mass and redshift.

In the top-left panel from figure 9, one can see the monopole of the TPCF for the different volume-limited samples indicated in table 2. The mean deviation between the data and the prediction from the mock for scales between  $2 < s < 30 \ h^{-1} \text{Mpc}$  is about a 2% for  $M_r < -21.5$ , below a 1% for  $M_r < -21.0$ , and  $-20.5$ , and about a 2% for  $M_r < -20.0$ . We can see that there is an increasing deviation between the mock and the data for the faintest volume-limited samples, and this can't be fixed decreasing the scatter value (the monopole of the TPCF increases when decreasing the scatter, as demonstrated in [8]), because the value of the scatter for the faintest galaxies considered in this work is quite close to zero. Therefore, we suspect that the  $k + E$  corrections applied to both the data and the mock have to be improved.

In the bottom-left panel from the same figure, one can see the quadrupole of the TPCF for the same volume-limited samples. Similar tendencies as in the monopole are observed — the agreement between the data and the mock is better for the brightest volume-limited samples.

Next, we studied the dependence of the monopole of the TPCF for different cuts in stellar mass for the LRG tracer. We show the results for the middle redshift range only,  $0.6 < z < 0.8$ , but the reader can check using the public catalogues that similar agreement between the data and the mocks is obtained for the other two redshift ranges. The results from the UCHUU-LRG lightcone and DESI LRG Y1<sup>5</sup> can be seen in the top-right panel from figure 9. We have calculated that the average deviation of the mocks from the data in the range  $5 < s < 20 \ h^{-1} \text{Mpc}$  is below a 5% for all the stellar mass cuts, and between a 10 and a 15% for scales in the range  $0.1 < s < 1 \ h^{-1} \text{Mpc}$ .

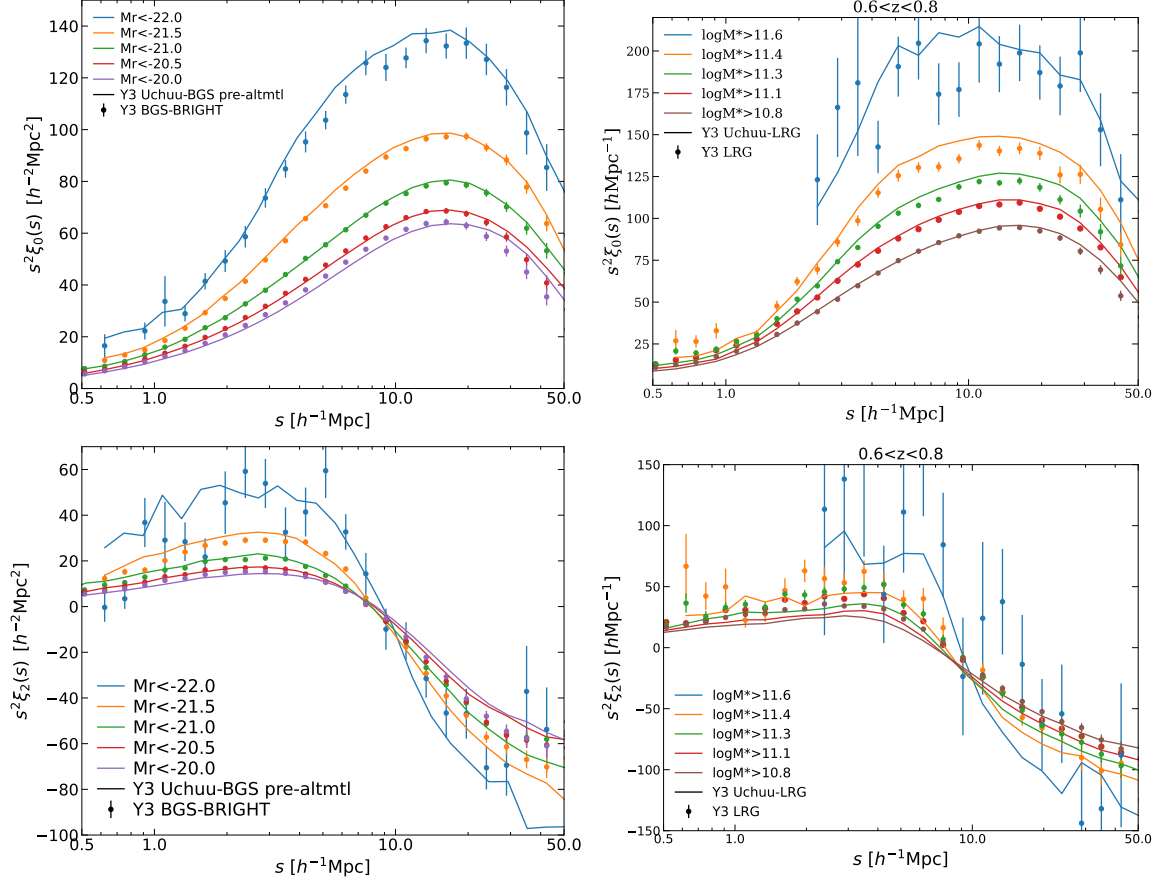
Finally, the dependence of the quadrupole with different cuts in the stellar mass for the LRGs can be seen in the bottom-right panel from figure 9. As expected from figures 6 and 7, the agreement between the data and the mock is not good at small scales, specially for the subsamples that include almost all galaxies (or, in other words, subsamples with a low-massive stellar mass cut).

### 5.2 Redshift space distortions

Top panels from figure 10 represent the TPCF from UCHUU-BGS and UCHUU-LRG lightcone and BGS-BRIGHT and LRG Y3 plotted as a function of pair separation perpendicular ( $r_p$ ) and parallel ( $\pi$ ) to the line of sight. From this figure the presence of redshift space distortions (RSD) is clear by the elongation along the  $\pi$  axis at scales within  $5 \ h^{-1} \text{Mpc}$ . The remarkable agreement between our mock and the data is noteworthy.

To facilitate a more precise comparison between the data and the mock on small scales, we have integrated out  $\pi$  and projected the measurements onto the  $r_p$  axis. This is illustrated in the bottom panels of figure 10. For the case of the BGS (LRG), the discrepancy between the mock and the data is approximately 1 (0.7) % at  $r_p = 1.5h^{-1} \text{Mpc}$  and decreases (increases)

<sup>5</sup>Note that we do the comparison with DESI LRG Y1 because the estimations for the stellar mass using the CIGALE tool have been done only for Y1 so far.



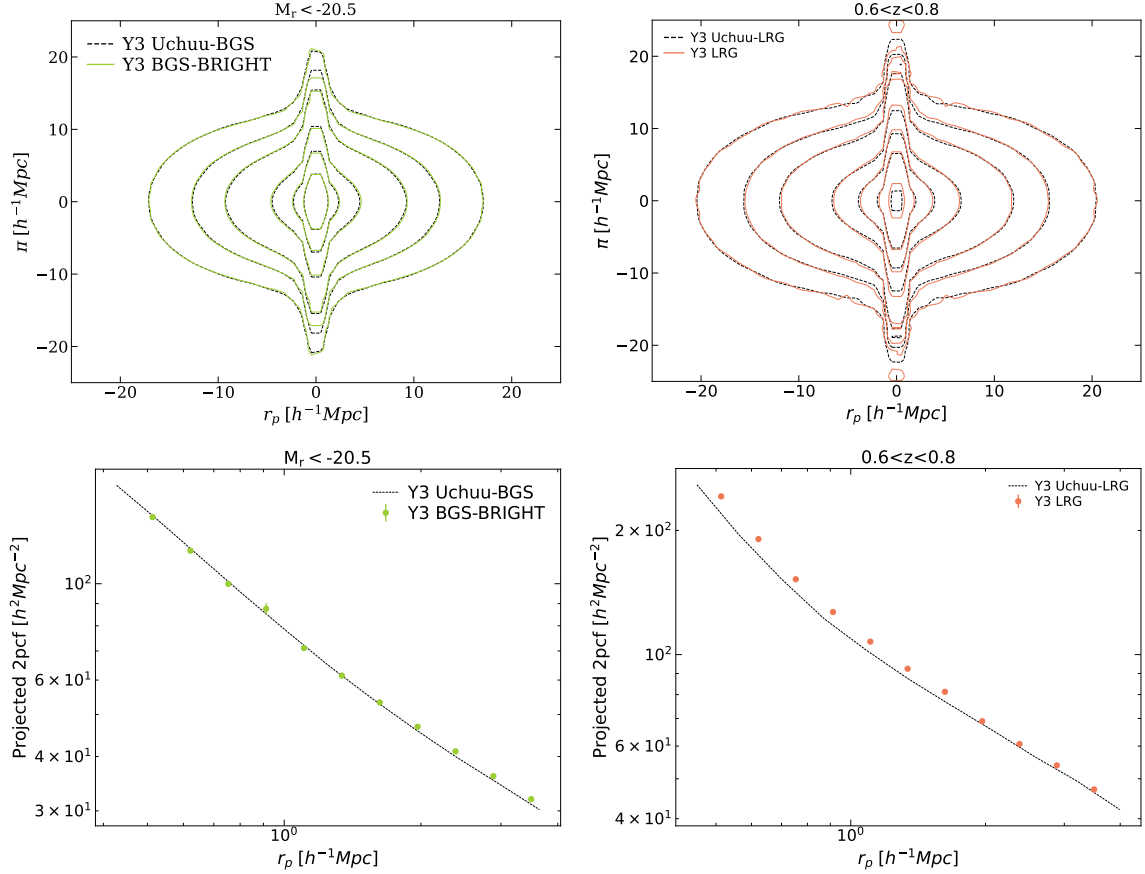
**Figure 9.** The monopole of the two point correlation function obtained for the different volume-limited samples detailed in table 2 from UCHUU-BGS (solid lines) and BGS-BRIGHT (points) is shown in the top-left panel. The same function but for different cuts in stellar mass and for  $0.6 < z < 0.8$  for the UCHUU-LRG lightcone (solid lines) and LRG Y1 data (points) is shown in the top-right panel. In the bottom panels the quadrupole is shown.

to about 0.8 (2) % at  $r_p = 0.5h^{-1} \text{ Mpc}$ . The larger deviation at the smallest scales for the case of the LRGs may be attributed to the absence of a fibre collision correction in our analysis. Additionally, since LRGs are more distant, the angular scales corresponding to the fibre collisions would map to larger physical scales. This may contribute to the observed discrepancies at the smallest scales, as the fibre collision effects are more pronounced for the more distant objects.

## 6 Redshift-space power spectrum

In this section, we also compare the power spectrum monopole  $P_0(k)$  and quadrupole  $P_2(k)$  for the galaxy samples considered in the previous section, measured over the wavenumber range  $0 < k < 0.2 h \text{ Mpc}^{-1}$  with a bin width  $\Delta k = 0.02 h \text{ Mpc}^{-1}$ . The power spectrum estimator is based on the fast Fourier transform (FFT) algorithm, as developed by Hand et al. [59] and implemented in the Python package PYPOWER.<sup>6</sup> Incompleteness, redshift

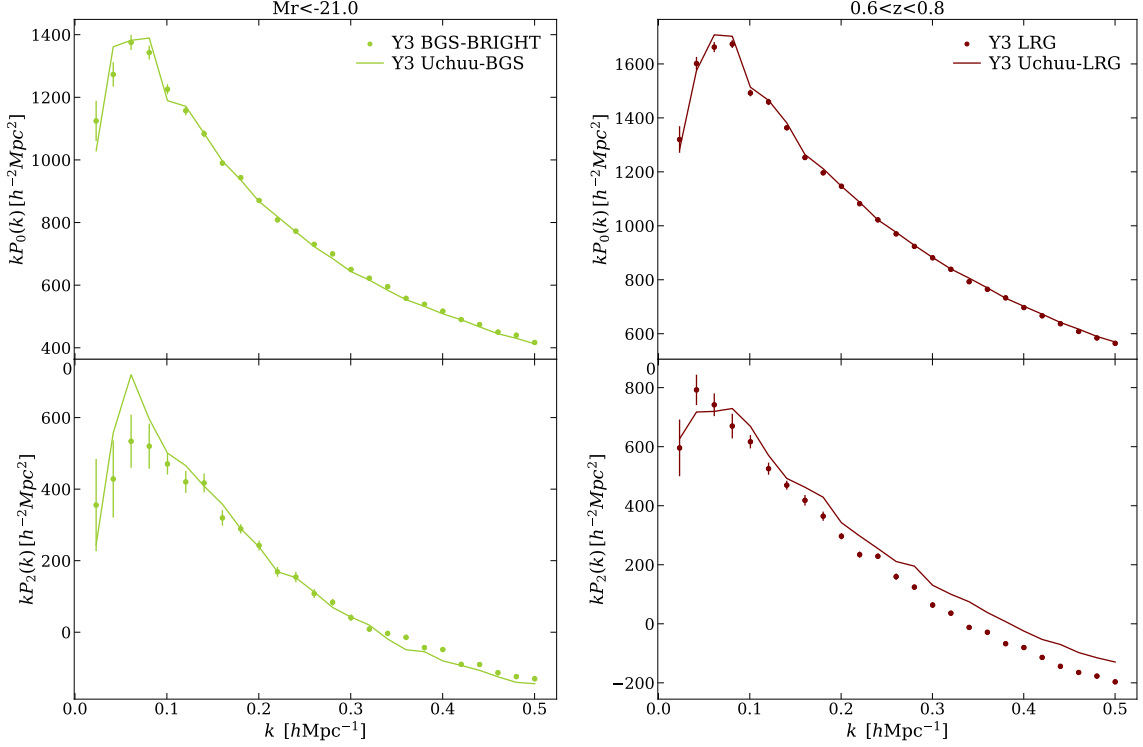
<sup>6</sup><https://github.com/cosmodesi/pypower>.



**Figure 10.** The TPC from UCHUU-LRG (solid line) and Y3 LRG (dashed line) plotted as a function of pair separation perpendicular ( $r_p$ ) and parallel ( $\pi$ ) to the line of sight can be seen in the top panel. To illustrate deviations from circular symmetry, the first quadrant is replicated with reflections in both axes. The contours levels shown are [0.5, 0.8, 2, 4, 8, 20]. In the bottom panels we show the same measurements of the projected TPCF integrating  $\pi$  out.

failure and imaging systematic weights are applied to the survey data, and FKP weights are calculated and applied to both the survey data and the UCHUU lightcones. For each catalogue, the particles are assigned to a mesh grid using the piecewise cubic spline (PCS) scheme, with a minimum grid number in each dimension chosen such that the Nyquist wavenumber is above  $0.4 h \text{Mpc}^{-1}$  to avoid aliasing effects. Finally, the power spectrum normalisation factor is calculated as a sum over both the galaxy and random mesh grids of grid cell size  $10 h^{-1} \text{Mpc}$ , such that it approximates the window function amplitude and thus the power spectrum amplitude can be approximately compared across different survey geometries [52].

In figure 11, we show the power spectrum monopole and quadrupole measured for BGS volume-limited sample with  $M_r < -21.0$  and for LRG in the redshift bin  $0.6 < z < 0.8$ . We observe excellent agreement between the Y3 survey data and the Uchuu lightcones, especially for LRG. However, one caveat is that the slightly different window functions (of which the various weights are an input) complicate the interpretation as they change the shape of the observed power spectrum [55]; in addition, the power spectrum normalisation



**Figure 11.** Power spectrum monopole (top panels) and quadrupole measurements (bottom panels) of the BGS-BRIGHT volume-limited sample with  $M_r < -21.0$  (left) and LRG with  $0.6 < z < 0.8$  (right) tracers. The solid lines and the points show the measurements from UCHUU and the Y3 survey data, respectively. The error bars represent the standard deviation calculated from the 25 AbacusSummit mocks.

discussed above only approximately accounts for the amplitude modulation induced by the window function. By contrast, the Landy-Szalay type of TPCF estimators eliminate the effect of the window function.

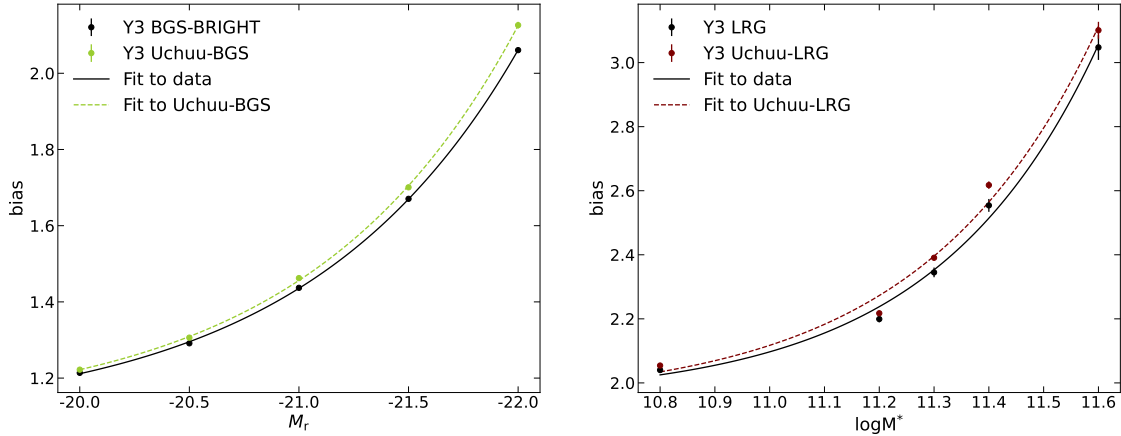
## 7 Large-scale bias

The large-scale bias,  $b$ , for both DESI tracers was measured from the DESI DR2 Survey and compared to their prediction obtained from the UCHUU lightcones in the Planck cosmology. The results are presented in figure 12. We measure the linear bias by fitting

$$\xi_0(s) = b^2 \left( 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 \right) \xi_{\text{lin}}(s) \quad (7.1)$$

to the correlation function monopole measurements,  $\xi_0(s)$ , over a given range of separations.  $\xi_{\text{lin}}(s)$  is from the linear power spectrum at the redshift of our galaxy sample, and  $\beta = \Omega_m^{0.6}/b$  [see 60, 61].

The linear bias for both BGS and LRG samples is measured over the separation range  $10 < s < 40 h^{-1} \text{Mpc}$  at the median redshift of each volume-limited sample. For the LRGs, we consider only galaxies in the redshift range  $0.6 < z < 0.8$ . As expected, the brightest



**Figure 12.** Bias of the BGS (LRG) magnitude threshold volume limited-samples (samples with stellar mass cut), at the median redshift of each sample is shown in the left (right) panel for the data (points) and UCHUU mocks (solid lines).

and most massive galaxies exhibit the highest bias, as they reside in the most massive halos, which are more strongly clustered.

Comparing the bias from Uchuu-BGS to that from BGS-BRIGHT, we find that they are consistent within  $1\sigma$  for the faintest volume-limited samples. For the brightest samples with  $M_r < -21$ , the biases show a larger discrepancy. The uncertainties shown for the bias are derived from the fitting procedure and do not include cosmic variance, which can be significant given the finite volume probed by the survey and the simulation. Including cosmic variance would increase the total error budget and could reduce the apparent discrepancy. Despite this, the agreement between mock and data is very good for the 2PCF monopole and quadrupole, the projected 2PCF, and the power spectrum, suggesting that the small difference in bias is not driven by limitations in the SHAM modelling or redshift-space clustering.

A similar trend is observed for the LRGs. In both cases, the bias measured from the mock is higher than that from the data. This is consistent with figure 9, which shows that the monopole of the TPCF at scales  $10 < s < 40 h^{-1} \text{ Mpc}$  is larger in the mock than in the data for the brightest and most massive galaxies. Furthermore, at larger scales, the mock predictions systematically exceed the measurements from the data.

The final step we took to study the dependence of the bias with the absolute magnitude and stellar mass is finding a functional dependence between these variables. To do this, we used the same function as that proposed in [62] for the BGS:

$$b(< M_r) = B_0 + B_1 \times 10^{B_2(M_r^* - M_r)/2.5} \quad (7.2)$$

where  $M_r = -20.44$ .

For the LRGs, we use the same equation but change  $M_r$  by  $\log(M^*)$  and fix the value of  $M_r^*$  to 12.3, which is the typical mass for a LRG.

The values we obtain for  $B_0$ ,  $B_1$  and  $B_2$  can be seen in the table 4. The results from these two fits can be seen in the left panel from figure 12 with continuous lines for Y3 BGS-BRIGHT and Y3 LRG, and dashed lines for the mock. For both tracers, the results

| Survey        | $B_0$             | $B_1$             | $B_2$            |
|---------------|-------------------|-------------------|------------------|
| Y3 BGS-BRIGHT | $1.087 \pm 0.008$ | $0.196 \pm 0.008$ | $1.12 \pm 0.02$  |
| Y3 Uchuu-BGS  | $1.10 \pm 0.013$  | $0.200 \pm 0.012$ | $1.14 \pm 0.03$  |
| Y3 LRG        | $1.95 \pm 0.07$   | $0.40 \pm 0.08$   | $-3.65 \pm 0.55$ |
| Y3 Uchuu-LRG  | $1.94 \pm 0.10$   | $0.46 \pm 0.12$   | $-3.42 \pm 0.69$ |
| DESI-BGS 1%   | $0.59 \pm 0.43$   | $0.62 \pm 0.45$   | $0.43 \pm 0.29$  |
| DESI-LRG 1%   | $1.90 \pm 0.018$  | $0.26 \pm 0.02$   | $-5.69 \pm 0.69$ |

**Table 4.** The table presents the best-fit values of the parameters  $B_0$ ,  $B_1$ , and  $B_2$  in the functional equation of the galaxy bias as a function of luminosity and stellar mass. The first four rows correspond to the Y3 BGS-BRIGHT and Y3 LRG surveys, along with their respective Uchuu-based mock counterparts. The results for both real and mock data are consistent within  $1\sigma$ , as shown in figure 12. For LRGs, the bias is modelled as a function of  $\log(M^*)$ , with the characteristic mass fixed at  $M_r^* = 12.3$ . The last two rows contain the corresponding values for the DESI One-Percent Survey from [6]. Despite the larger uncertainties in DESI 1% due to its limited survey volume, the obtained parameters are compatible with the results found in this work.

from the mock and the data are compatible within  $1\sigma$ . We have also added in the last two rows of the table the results obtained for the DESI One-Percent Survey in [6]. We can see that the results obtained in that work are compatible with the results obtained in this work, although the uncertainties obtained for the One-Percent survey were much larger because of the tiny volume of that survey.

To conclude this section, we have computed the linear Two-Point Correlation Function (TPCF) using the bias measurements derived earlier for each BGS volume-limited sample and LRG sample with different stellar mass cuts. The results are shown in figure 13 from appendix A, where the left panel displays the linear TPCF at redshift  $z = 0$  for the data (represented by points) and mocks (depicted by lines) for BGS-BRIGHT targets, while the right panel shows the corresponding data for LRGs. Additionally, the linear TPCF calculated from Uchuu’s matter power spectrum is shown as a dark dashed line. The bottom panels of both plots present the ratio between the data and mock linear TPCF. The shaded gray area represents the 5% limits. From these panels, it is evident that the agreement between the data and mock linear TPCF with that from Uchuu is excellent at scales between  $10 < s < 20 \ h^{-1} \text{Mpc}$  (the ratio is below a 5%), and remains strong even at larger scales. For the calculation of the large-scale bias, we assumed that the TPCF was linear within the range  $10 < s < 20 \ h^{-1} \text{Mpc}$ , and this result confirms the validity of that assumption.

## 8 Summary

This study details the methodology used to generate simulated lightcones for the LRG and BGS-BRIGHT DESI DR2 tracers. These lightcones are built within the flat- $\Lambda$ CDM Planck cosmology framework using UCHUU, a 2.1 trillion-particle  $N$ -body simulation tailored for the DESI survey. UCHUU provides high-resolution modelling of dark matter haloes and subhaloes across a vast volume, capturing structures from massive galaxy clusters to dwarf galaxies.

To populate the UCHUU haloes with DESI galaxies, we apply the Subhalo Abundance Matching (SHAM) technique, using peak maximum circular velocity as a proxy for (sub)halo mass. The construction of BGS and LRG lightcones follows the standard SHAM approach, incorporating the redshift evolution of tracers and the dependence of clustering on key properties such as luminosity and stellar mass.

We compare the predicted clustering signals from UCHUU with measurements from DESI DR2 for each galaxy sample. Additionally, we analyse the halo occupation distribution and large-scale bias factors for the DESI targets.

Our main results are summarized as follows:

- We measured the redshift-space two-point correlation function monopole and quadrupole over the scales from  $0.01 h^{-1} \text{ Mpc}$  to  $200 h^{-1} \text{ Mpc}$ . Overall, we find consistency between the DESI DR2 (and Y1) Survey measurements and the theoretical predictions based on Planck cosmology using the UCHUU lightcones. The agreement between the data and the prediction from the mock for both tracers is excellent for the monopole, although we find some discrepancy in the quadrupole for the LRGs, which may be caused to the fact that the SHAM method does not accurately model galaxy velocities. If we focus on BAO scales, we find that the result predicted by the mocks and those measured in DESI DR1 are statistically compatible within  $1\sigma$  for the full sample for both tracers.
- We have also shown that we can reproduce quite accurately the monopole of the TPCF for different volume-limited samples for the BGS and different stellar mass cuts for the LRGs. For the BGS, we've been able to reach the limit in absolute magnitude of  $M_r < -20.0$ . We've found that the disagreement between the monopole from the data and that predicted by UCHUU-BGS mocks becomes bigger the faintest the volume-limited samples is, even though we decrease the scatter value to values close to zero.
- We have also checked there is generally good agreement in the power spectrum monopole and quadrupole between the Uchuu lightcones and the DESI DR2 sample for BGS with  $M_r < -20.0$  and LRG in the redshift bin  $0.6 < z < 0.8$ , although the interpretation is complicated by the slightly different window functions.
- The linear bias factors were measured for both tracers from the DESI DR2 sample and compared to predictions based on the UCHUU lightcones in the Planck cosmology. We measured the bias as a function of absolute magnitude threshold for the various BGS volume-limited samples and as a function of the stellar mass for LRGs. For both tracers, we obtain that UCHUU bias is slightly larger than that obtained for DESI LRG Y3.

The results of this study are instrumental in refining key aspects of cosmological modelling, particularly in the generation of simulated lightcones and the development of more robust galaxy-halo connection techniques. These advancements will play a crucial role in the future of the DESI survey, enabling more accurate full-shape analyses of the Baryon Acoustic Oscillations (BAO) to constrain cosmological parameters. In particular, the mocks generated, especially the BGS catalog, will be invaluable for studies of large-scale structure, including the identification of voids. By enhancing our understanding of clustering signals, these findings will allow for a more precise interpretation of DESI's large-scale structure observations,

leading to better constraints on cosmological models and advancing our knowledge of the universe’s expansion and matter distribution.

A thorough analysis of the DESI DR2 Survey data allows us to assess and enhance the accuracy of methodologies used in cosmological studies, ultimately increasing the reliability of the final survey’s results. Furthermore, insights gained from the current DESI dataset provide a foundation for optimizing future observational strategies and theoretical models. These advancements are essential for maximizing the scientific impact of the full DESI survey, ensuring that it delivers the most accurate constraints on the fundamental properties of the universe.

## Acknowledgments

EFG acknowledges financial support from the Severo Ochoa grant CEX2021-001131-S funded by MCIN/AEI/ 10.13039/501100011033.

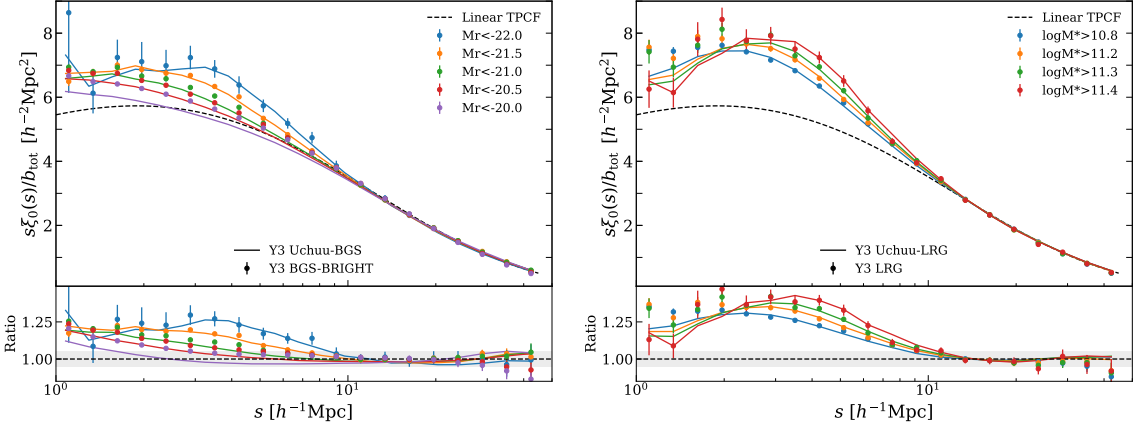
EFG and FP thank Instituto de Astrofísica de Andalucía (IAA-CSIC), Centro de Supercomputación de Galicia (CESGA) and the Spanish academic and research network (RedIRIS) in Spain for hosting Uchuu DR1, DR2 and DR3 in the Skies & Universes site for cosmological simulations. The Uchuu simulations were carried out on Aterui II supercomputer at Center for Computational Astrophysics, CfCA, of National Astronomical Observatory of Japan, and the K computer at the RIKEN Advanced Institute for Computational Science. The Uchuu Data Releases efforts have made use of the skun@IAA\_RedIRIS and skun6@IAA computer facilities managed by the IAA-CSIC in Spain (MICINN EU-Feder grant EQC2018-004366-P).

EFG and FP acknowledge support from the Spanish MICINN funding grant PGC2018-101931-B-I00.

TI has been supported by IAAR Research Support Program in Chiba University Japan, MEXT/JSPS KAKENHI (Grant Number JP19KK0344 and JP25H00662), MEXT as “Program for Promoting Researches on the Supercomputer Fugaku” (JPMXP1020230406), and JICFuS.

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the U.S. National Science Foundation (NSF), Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF’s National Optical-Infrared Astronomy Research Laboratory; the Science and Technology Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Humanities, Science and Technology of Mexico (CONAHCYT); the Ministry of Science, Innovation and Universities of Spain (MICIU/AEI/10.13039/501100011033), and by the DESI Member Institutions: <https://www.desi.lbl.gov/collaborating-institutions>. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation, the U.S. Department of Energy, or any of the listed funding agencies.

The authors are honored to be permitted to conduct scientific research on I’oligam Du’ag (Kitt Peak), a mountain with particular significance to the Tohono O’odham Nation.



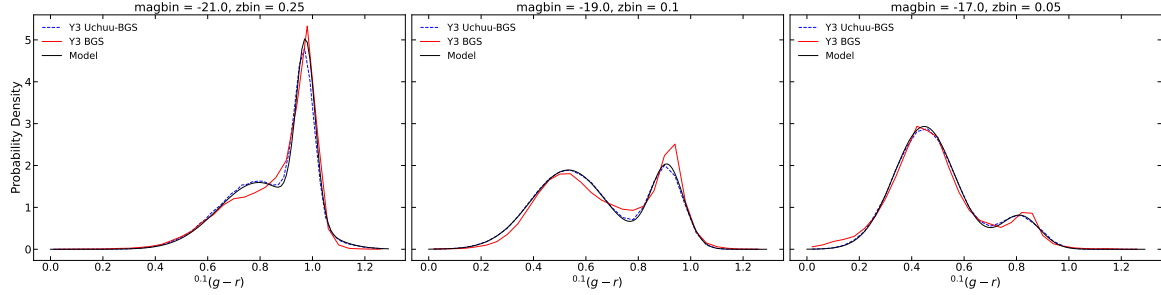
**Figure 13.** The left panel shows the comparison of the monopole of the linear Two-Point Correlation Function (TPCF) with the bias,  $b$ , measured at redshift  $z = 0$ , for different BGS volume-limited samples. The linear TPCF from the Uchuu simulation (dashed black line), also measured at redshift  $z = 0$ , is shown for reference. The right panel displays the same comparison, but for different LRG samples with various stellar mass cuts. The bottom panels show the ratio of the data points to the lines, relative to the linear TPCF. The shaded area represents the 5% limits.  $b_{\text{tot}}$  is equal to  $b^2 \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)$ .

## A Linear regime of the TPCF

In this section, we show that the TPCF is linear at separation scales between  $10 < s < 40$   $h^{-1}$  Mpc, which is the separation range considered in order to measure the large-scale bias in section 7. This can be seen in figure 13, where the linear Two-Point Correlation Function (TPCF) using the bias measurements derived in the aforementioned section for each BGS volume-limited sample and LRG sample with different stellar mass cuts. The left panel displays the linear TPCF at redshift  $z = 0$  for the data (represented by points) and mocks (depicted by lines) for BGS-BRIGHT targets, while the right panel shows the corresponding data for LRGs. Additionally, the linear TPCF calculated from Uchuu’s matter power spectrum is shown as a dark dashed line. The bottom panels of both plots present the ratio between the data and mock linear TPCF. The shaded gray area represents the 5% limits. From these panels, it is evident that the agreement between the data and mock linear TPCF with that from Uchuu is excellent at scales between  $10 < s < 20$   $h^{-1}$  Mpc (the ratio is below a 5%), and remains strong even at larger scales. For the calculation of the large-scale bias, we assumed that the TPCF was linear within the range  $10 < s < 20$   $h^{-1}$  Mpc, and this result confirms the validity of that assumption.

## B Modelling of the colour distribution

In this appendix, we present a comparison of the probability density functions for the colour distribution of the mock data, observational data, and the adopted model. This comparison is illustrated in figure 14, where each panel corresponds to different bins of magnitude and redshift. By construction, the agreement between the mock data and the model is excellent. However, the model does not match the observational data as well, particularly for the faintest



**Figure 14.** Comparison of the probability density of the colour distribution for Y3 BGS-BRIGHT (red line), Y3 Uchuu-BGS mock (blue line) and the model adopted for assigning a colour to each galaxy (black line). Each panel shows the result for different absolute magnitude bins (magbin) and redshift bins (zbin). The bin width of magbin and zbin is 0.2 and 0.02, respectively.

samples in the tails of the distribution (i.e., the reddest and bluest galaxies). This discrepancy arises because the model assumes a double Gaussian fit to the data, which is not always a suitable approximation for certain magnitude and redshift bins. Addressing this limitation will be crucial for improving the next generation of mock catalogues.

**Data Availability Statement.** The DR1 UCHUU-LRG and UCHUU-BGS lightcones will be made public at <https://data.desi.lbl.gov/doc/vac/>. DR2 lightcones will be made public with the DESI DR2. The data points corresponding to the figures from this paper will be available in a Zenodo repository.

## References

- [1] S. de la Torre and J.A. Peacock, *Reconstructing the distribution of haloes and mock galaxies below the resolution limit in cosmological simulations*, *Mon. Not. Roy. Astron. Soc.* **435** (2013) 743 [[arXiv:1212.3615](#)] [[INSPIRE](#)].
- [2] M. White, J.L. Tinker and C.K. McBride, *Mock galaxy catalogues using the quick particle mesh method*, *Mon. Not. Roy. Astron. Soc.* **437** (2014) 2594 [[arXiv:1309.5532](#)] [[INSPIRE](#)].
- [3] S.A. Rodríguez-Torres et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: modelling the clustering and halo occupation distribution of BOSS CMASS galaxies in the Final Data Release*, *Mon. Not. Roy. Astron. Soc.* **460** (2016) 1173 [[arXiv:1509.06404](#)] [[INSPIRE](#)].
- [4] A. Smith et al., *A Lightcone Catalogue from the Millennium-XXL Simulation*, *Mon. Not. Roy. Astron. Soc.* **470** (2017) 4646 [[arXiv:1701.06581](#)] [[INSPIRE](#)].
- [5] DES collaboration, *The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs*, [[arXiv:1901.02401](#)] [[INSPIRE](#)].
- [6] F. Prada et al., *The DESI One-Percent Survey: Modelling the clustering and halo occupation of all four DESI tracers with UCHUU*, *Astron. Astrophys.* **698** (2025) A170 [[arXiv:2306.06315](#)] [[INSPIRE](#)].
- [7] J. Ereza et al., *The Uchuu-glam BOSS and eBOSS LRG lightcones: exploring clustering and covariance errors*, *Mon. Not. Roy. Astron. Soc.* **532** (2024) 1659 [[arXiv:2311.14456](#)] [[INSPIRE](#)].

- [8] C.A. Dong-Páez et al., *The Uchuu-SDSS galaxy light-cones: a clustering, redshift space distortion and baryonic acoustic oscillation study*, *Mon. Not. Roy. Astron. Soc.* **528** (2024) 7236 [[arXiv:2208.00540](#)] [[INSPIRE](#)].
- [9] EBOSS collaboration, *Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory*, *Phys. Rev. D* **103** (2021) 083533 [[arXiv:2007.08991](#)] [[INSPIRE](#)].
- [10] 2DFGRS collaboration, *The 2dF Galaxy Redshift Survey: Spectra and redshifts*, *Mon. Not. Roy. Astron. Soc.* **328** (2001) 1039 [[astro-ph/0106498](#)] [[INSPIRE](#)].
- [11] V. Springel et al., *Simulating the joint evolution of quasars, galaxies and their large-scale distribution*, *Nature* **435** (2005) 629 [[astro-ph/0504097](#)] [[INSPIRE](#)].
- [12] D. Nelson et al., *The Illustris Simulation: Public Data Release*, *Astron. Comput.* **13** (2015) 12 [[arXiv:1504.00362](#)] [[INSPIRE](#)].
- [13] C. Marinoni and M.J. Hudson, *The mass-to-light function of virialized systems and the relationship between their optical and x-ray properties*, *Astrophys. J.* **569** (2002) 101 [[astro-ph/0109134](#)] [[INSPIRE](#)].
- [14] A.V. Kravtsov et al., *The dark side of the halo occupation distribution*, *Astrophys. J.* **609** (2004) 35 [[astro-ph/0308519](#)] [[INSPIRE](#)].
- [15] A. Vale and J.P. Ostriker, *Linking halo mass to galaxy luminosity*, *Mon. Not. Roy. Astron. Soc.* **353** (2004) 189 [[astro-ph/0402500](#)] [[INSPIRE](#)].
- [16] C. Conroy, R.H. Wechsler and A.V. Kravtsov, *Modeling luminosity-dependent galaxy clustering through cosmic time*, *Astrophys. J.* **647** (2006) 201 [[astro-ph/0512234](#)] [[INSPIRE](#)].
- [17] S. Saito et al., *Connecting massive galaxies to dark matter haloes in BOSS — I. Is galaxy colour a stochastic process in high-mass haloes?*, *Mon. Not. Roy. Astron. Soc.* **460** (2016) 1457 [[arXiv:1509.00482](#)] [[INSPIRE](#)].
- [18] S. Safonova, P. Norberg and S. Cole, *Rosella: a mock catalogue from the P-Millennium simulation*, *Mon. Not. Roy. Astron. Soc.* **505** (2021) 325 [[arXiv:2009.00005](#)] [[INSPIRE](#)].
- [19] J. DeRose, M.R. Becker and R.H. Wechsler, *Modeling Redshift-space Clustering with Abundance Matching*, *Astrophys. J.* **940** (2022) 13 [[arXiv:2105.12104](#)] [[INSPIRE](#)].
- [20] T. Ishiyama et al., *The Uchuu simulations: Data Release 1 and dark matter halo concentrations*, *Mon. Not. Roy. Astron. Soc.* **506** (2021) 4210 [[arXiv:2007.14720](#)] [[INSPIRE](#)].
- [21] DESI collaboration, *The DESI Experiment Part I: Science, Targeting, and Survey Design*, [arXiv:1611.00036](#) [[INSPIRE](#)].
- [22] DESI collaboration, *The DESI Experiment Part II: Instrument Design*, [arXiv:1611.00037](#) [[INSPIRE](#)].
- [23] DESI collaboration, *Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **164** (2022) 207 [[arXiv:2205.10939](#)] [[INSPIRE](#)].
- [24] DESI collaboration, *The Robotic Multiobject Focal Plane System of the Dark Energy Spectroscopic Instrument (DESI)*, *Astron. J.* **165** (2023) 9 [[arXiv:2205.09014](#)] [[INSPIRE](#)].
- [25] DESI collaboration, *The Optical Corrector for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **168** (2024) 95 [[arXiv:2306.06310](#)] [[INSPIRE](#)].
- [26] DESI collaboration, *The DESI Experiment, a whitepaper for Snowmass 2013*, [arXiv:1308.0847](#) [[INSPIRE](#)].

- [27] A. Albrecht, *Report from the Dark Energy Task Force*, in *APS April Meeting Abstracts*, APS Meeting Abstracts, p. G1.002 (2006) [<https://ui.adsabs.harvard.edu/abs/2006APS..APR.G1002A>].
- [28] M. Siudek et al., *Value-added catalog of physical properties for more than 1.3 million galaxies from the DESI survey*, *Astron. Astrophys.* **691** (2024) A308 [[arXiv:2409.19066](#)] [[INSPIRE](#)].
- [29] DESI collaboration, *Data Release 1 of the Dark Energy Spectroscopic Instrument*, *Astron. J.* **171** (2026) 285 [[arXiv:2503.14745](#)] [[INSPIRE](#)].
- [30] DESI collaboration, *Validation of the Scientific Program for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **167** (2024) 62 [[arXiv:2306.06307](#)] [[INSPIRE](#)].
- [31] C.H. Hahn et al., *The DESI Bright Galaxy Survey: Final Target Selection, Design, and Validation*, *Astron. J.* **165** (2023) 253 [[arXiv:2208.08512](#)] [[INSPIRE](#)].
- [32] DESI collaboration, *Target Selection and Validation of DESI Luminous Red Galaxies*, *Astron. J.* **165** (2023) 58 [[arXiv:2208.08515](#)] [[INSPIRE](#)].
- [33] E. Chaussidon et al., *Target Selection and Validation of DESI Quasars*, *Astrophys. J.* **944** (2023) 107 [[arXiv:2208.08511](#)] [[INSPIRE](#)].
- [34] A. Raichoor et al., *Target Selection and Validation of DESI Emission Line Galaxies*, *Astron. J.* **165** (2023) 126 [[arXiv:2208.08513](#)] [[INSPIRE](#)].
- [35] DESI collaboration, *The Spectroscopic Data Processing Pipeline for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **165** (2023) 144 [[arXiv:2209.14482](#)] [[INSPIRE](#)].
- [36] R. Vaisakh et al., *DESI Y3 reference mocks: clustering results from Uchuu QSO and ELG lightcones*, in preparation (2026).
- [37] DESI collaboration, *DESI DR2 results. II. Measurements of baryon acoustic oscillations and cosmological constraints*, *Phys. Rev. D* **112** (2025) 083515 [[arXiv:2503.14738](#)] [[INSPIRE](#)].
- [38] DESI collaboration, *Validation of the DESI DR2 measurements of baryon acoustic oscillations from galaxies and quasars*, *Phys. Rev. D* **112** (2025) 083512 [[arXiv:2503.14742](#)] [[INSPIRE](#)].
- [39] PLANCK collaboration, *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13 [[arXiv:1502.01589](#)] [[INSPIRE](#)].
- [40] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [*Erratum ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [41] T. Ishiyama, K. Nitadori and J. Makino, *4.45 Pflops Astrophysical N-Body Simulation on K computer — The Gravitational Trillion-Body Problem*, [arXiv:1211.4406](#) [[INSPIRE](#)].
- [42] P.S. Behroozi, R.H. Wechsler and H.-Y. Wu, *The Rockstar Phase-Space Temporal Halo Finder and the Velocity Offsets of Cluster Cores*, *Astrophys. J.* **762** (2013) 109 [[arXiv:1110.4372](#)] [[INSPIRE](#)].
- [43] P.S. Behroozi et al., *Gravitationally Consistent Halo Catalogs and Merger Trees for Precision Cosmology*, *Astrophys. J.* **763** (2013) 18 [[arXiv:1110.4370](#)] [[INSPIRE](#)].
- [44] DESI collaboration, *The Early Data Release of the Dark Energy Spectroscopic Instrument*, *Astron. J.* **168** (2024) 58 [[arXiv:2306.06308](#)] [[INSPIRE](#)].
- [45] S.M. Moore et al., *DESI Y1 Luminosity Functions*, in preparation (2026).
- [46] J. Moustakas et al., *FastSpecFit: Fast spectral synthesis and emission-line fitting of DESI spectra*, Astrophysics Source Code Library, record [ascl:2308.005](#) (2023).

- [47] T. McNaught-Roberts et al., *Galaxy And Mass Assembly (GAMA): the dependence of the galaxy luminosity function on environment, redshift and colour*, *Mon. Not. Roy. Astron. Soc.* **445** (2014) 2125 [[arXiv:1409.4681](#)].
- [48] P. Behroozi, R.H. Wechsler, A.P. Hearin and C. Conroy, *UniverseMachine: The correlation between galaxy growth and dark matter halo assembly from  $z = 0$ –10*, *Mon. Not. Roy. Astron. Soc.* **488** (2019) 3143 [[arXiv:1806.07893](#)] [[INSPIRE](#)].
- [49] A. Smith et al., *Solving small-scale clustering problems in approximate light-cone mocks*, *Mon. Not. Roy. Astron. Soc.* **516** (2022) 1062 [[arXiv:2206.08763](#)] [[INSPIRE](#)].
- [50] M. Boquien et al., *CIGALE: a python Code Investigating GALaxy Emission*, *Astron. Astrophys.* **622** (2019) A103 [[arXiv:1811.03094](#)].
- [51] J. Moustakas et al., *PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function From  $z = 0$ –1*, *Astrophys. J.* **767** (2013) 50 [[arXiv:1301.1688](#)] [[INSPIRE](#)].
- [52] DESI collaboration, *DESI 2024 II: sample definitions, characteristics, and two-point clustering statistics*, *JCAP* **07** (2025) 017 [[arXiv:2411.12020](#)] [[INSPIRE](#)].
- [53] D. Bianchi and W.J. Percival, *Unbiased clustering estimation in the presence of missing observations*, *Mon. Not. Roy. Astron. Soc.* **472** (2017) 1106 [[arXiv:1703.02070](#)] [[INSPIRE](#)].
- [54] DESI collaboration, *Production of alternate realizations of DESI fiber assignment for unbiased clustering measurement in data and simulations*, *JCAP* **01** (2025) 127 [[arXiv:2404.03006](#)] [[INSPIRE](#)].
- [55] H.A. Feldman, N. Kaiser and J.A. Peacock, *Power spectrum analysis of three-dimensional redshift surveys*, *Astrophys. J.* **426** (1994) 23 [[astro-ph/9304022](#)] [[INSPIRE](#)].
- [56] S.D. Landy and A.S. Szalay, *Bias and variance of angular correlation functions*, *Astrophys. J.* **412** (1993) 64 [[INSPIRE](#)].
- [57] M. Sinha and L.H. Garrison, *corrfunc — a suite of blazing fast correlation functions on the CPU*, *Mon. Not. Roy. Astron. Soc.* **491** (2020) 3022 [[arXiv:1911.03545](#)] [[INSPIRE](#)].
- [58] M. Sinha and L.H. Garrison, *Corrfunc: Blazing fast correlation functions with AVX512F SIMD Intrinsics*, [arXiv:1911.08275](#) [[DOI:10.1007/978-981-13-7729-7\\_1](#)] [[INSPIRE](#)].
- [59] N. Hand, Y. Li, Z. Slepian and U. Seljak, *An optimal FFT-based anisotropic power spectrum estimator*, *JCAP* **07** (2017) 002 [[arXiv:1704.02357](#)] [[INSPIRE](#)].
- [60] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. Roy. Astron. Soc.* **227** (1987) 1 [[INSPIRE](#)].
- [61] A.J.S. Hamilton, *Linear redshift distortions: A Review*, in the proceedings of the *Ringberg Workshop on Large Scale Structure*, Ringberg, Germany, 23–28 September 1996 [[DOI:10.1007/978-94-011-4960-0\\_17](#)] [[astro-ph/9708102](#)] [[INSPIRE](#)].
- [62] SDSS collaboration, *Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity*, *Astrophys. J.* **736** (2011) 59 [[arXiv:1005.2413](#)] [[INSPIRE](#)].

## Author List

E. Fernández-García<sup>1,\*</sup>, F. Prada<sup>1</sup>, A. Smith<sup>2</sup>, J. DeRose<sup>3</sup>, A.J. Ross<sup>4,5,6</sup>, S. Bailey<sup>7</sup>, M.S. Wang<sup>8</sup>, Z. Ding<sup>9</sup>, C. Guandalin<sup>8</sup>, C. Lamman<sup>10</sup>, R. Vaisakh<sup>11</sup>, R. Kehoe<sup>11</sup>, J. Lasker<sup>12</sup>, T. Ishiyama<sup>13</sup>, S.M. Moore<sup>12</sup>, S. Cole<sup>12</sup>, M. Siudek<sup>14</sup>, A. Amalbert<sup>11</sup>, A. Salcedo<sup>15</sup>, A. Hearin<sup>16</sup>, B. Joachimi<sup>17</sup>, A. Rocher<sup>18</sup>, S. Saito<sup>19</sup>, A. Krolewski<sup>20,21,22</sup>, Z. Slepian<sup>23</sup>, Q. Li<sup>24</sup>, K.S. Dawson<sup>24</sup>, E. Jullo<sup>25</sup>, J. Aguilar<sup>7</sup>, S. Ahlen<sup>26</sup>, D. Bianchi<sup>27,28</sup>, D. Brooks<sup>17</sup>, T. Claybaugh<sup>7</sup>, A. de la Macorra<sup>29</sup>, P. Doel<sup>17</sup>, S. Ferraro<sup>30,7</sup>, A. Font-Ribera<sup>31</sup>, J.E. Forero-Romero<sup>32,33</sup>, S. Gontcho A Gontcho<sup>7</sup>, G. Gutierrez<sup>34</sup>, K. Honscheid<sup>4,35,6</sup>, M. Ishak<sup>36</sup>, R. Joyce<sup>37</sup>, S. Juneau<sup>37</sup>, D. Kirkby<sup>38</sup>, T. Kisner<sup>7</sup>, A. Kremin<sup>7</sup>, O. Lahav<sup>17</sup>, A. Lambert<sup>7</sup>, M. Landriau<sup>7</sup>, M.E. Levi<sup>7</sup>, M. Manera<sup>39,40</sup>, R. Miquel<sup>40,41</sup>, J. Moustakas<sup>42</sup>, S. Nadathur<sup>43</sup>, W.J. Percival<sup>44,45</sup>, I. Pérez-Ràfols<sup>46</sup>, G. Rossi<sup>47</sup>, E. Sanchez<sup>48</sup>, D. Schlegel<sup>7</sup>, H. Seo<sup>49</sup>, J. Silber<sup>7</sup>, D. Sprayberry<sup>37</sup>, G. Tarlé<sup>50</sup>, B.A. Weaver<sup>37</sup>, P. Zarrouk<sup>51</sup>, R. Zhou<sup>7</sup>

<sup>1</sup> Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain

<sup>2</sup> Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, U.K.

<sup>3</sup> Physics Department, Brookhaven National Laboratory, Upton, NY 11973, U.S.A.

<sup>4</sup> Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.

<sup>5</sup> Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, U.S.A.

<sup>6</sup> The Ohio State University, Columbus, 43210 OH, U.S.A.

<sup>7</sup> Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.

<sup>8</sup> Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, U.K.

<sup>9</sup> University of Chinese Academy of Sciences, Nanjing 211135, People's Republic of China

<sup>10</sup> Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, U.S.A.

<sup>11</sup> Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, U.S.A.

<sup>12</sup> Astrophysics & Space Institute, Schmidt Sciences, New York, NY 10011, U.S.A.

<sup>13</sup> Digital Transformation Enhancement Council, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba, 263-8522, Japan

<sup>14</sup> Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain

<sup>15</sup> Steward Observatory, University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85721, U.S.A.

<sup>16</sup> Argonne National Laboratory, High-Energy Physics Division, 9700 S. Cass Avenue, Argonne, IL 60439, U.S.A.

<sup>17</sup> Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, U.K.

<sup>18</sup> Institute of Physics, Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, Chemin Pegasi 51, CH-1290 Versoix, Switzerland

<sup>19</sup> Institute for Multi-messenger Astrophysics and Cosmology, Department of Physics, Missouri University of Science and Technology, 1315 N Pine St, Rolla, MO 65409, U.S.A.

<sup>20</sup> Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

<sup>21</sup> Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada

<sup>22</sup> Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

<sup>23</sup> Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611, U.S.A.

- <sup>24</sup> *Department of Physics and Astronomy, The University of Utah,  
115 South 1400 East, Salt Lake City, UT 84112, U.S.A.*
- <sup>25</sup> *Aix-Marseille Univ., CNRS, CNES, LAM, Marseille, France*
- <sup>26</sup> *Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, U.S.A.*
- <sup>27</sup> *Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,  
Via Celoria 16, I-20133 Milano, Italy*
- <sup>28</sup> *INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy*
- <sup>29</sup> *Instituto de Física, Universidad Nacional Autónoma de México, Circuito de la Investigación Científica,  
Ciudad Universitaria, Cd. de México C. P. 04510, México*
- <sup>30</sup> *University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, U.S.A.*
- <sup>31</sup> *Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology,  
Edifici Cn, Campus UAB, 08193, Bellaterra (Barcelona), Spain*
- <sup>32</sup> *Departamento de Física, Universidad de los Andes,  
Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia*
- <sup>33</sup> *Observatorio Astronómico, Universidad de los Andes,  
Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia*
- <sup>34</sup> *Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, U.S.A.*
- <sup>35</sup> *Department of Physics, The Ohio State University,  
191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.*
- <sup>36</sup> *Department of Physics, The University of Texas at Dallas,  
800 W. Campbell Rd., Richardson, TX 75080, U.S.A.*
- <sup>37</sup> *NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, U.S.A.*
- <sup>38</sup> *Department of Physics and Astronomy, University of California, Irvine, 92697, U.S.A.*
- <sup>39</sup> *Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona,  
08193 Bellaterra (Barcelona), Spain*
- <sup>40</sup> *Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology,  
Edifici Cn, Campus UAB, 08193, Bellaterra (Barcelona), Spain*
- <sup>41</sup> *Institució Catalana de Recerca i Estudis Avançats,  
Passeig de Lluís Companys, 23, 08010 Barcelona, Spain*
- <sup>42</sup> *Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, U.S.A.*
- <sup>43</sup> *Institute of Cosmology and Gravitation, University of Portsmouth,  
Dennis Sciamia Building, Portsmouth, PO1 3FX, U.K.*
- <sup>44</sup> *Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*
- <sup>45</sup> *Waterloo Centre for Astrophysics, University of Waterloo,  
200 University Ave W, Waterloo, ON N2L 3G1, Canada*
- <sup>46</sup> *Departament de Física, EEBE, Universitat Politècnica de Catalunya,  
c/Eduard Maristany 10, 08930 Barcelona, Spain*
- <sup>47</sup> *Department of Physics and Astronomy, Sejong University,  
209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*
- <sup>48</sup> *CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*
- <sup>49</sup> *Department of Physics & Astronomy, Ohio University, 139 University Terrace, Athens, OH 45701, U.S.A.*
- <sup>50</sup> *Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, U.S.A.*
- <sup>51</sup> *Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE),  
FR-75005 Paris, France*

\* Corresponding author