

## The Aemulus Project V: Cosmological constraint from small-scale clustering of BOSS galaxies

ZHONGXU ZHAI,<sup>1, 2, 3, 4</sup> JEREMY L. TINKER,<sup>5</sup> ARKA BANERJEE,<sup>6, 7</sup> JOSEPH DEROSA,<sup>8, 9</sup> HONG GUO,<sup>10</sup> YAO-YUAN MAO,<sup>11, \*</sup>  
 SEAN MC LAUGHLIN,<sup>12, 13</sup> KATE STOREY-FISHER,<sup>5</sup> AND RISA H. WECHSLER<sup>12, 13</sup>

<sup>1</sup>*Department of Astronomy, School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*

<sup>2</sup>*Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai 200240, China*

<sup>3</sup>*Waterloo Center for Astrophysics, University of Waterloo, Waterloo, ON N2L 3G1, Canada*

<sup>4</sup>*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*

<sup>5</sup>*Center for Cosmology and Particle Physics, Department of Physics, New York University, 726 Broadway, New York, NY 10003, USA*

<sup>6</sup>*Department of Physics, Indian Institute of Science Education and Research, Homi Bhabha Road, Pashan, Pune 411008, India*

<sup>7</sup>*Fermi National Accelerator Laboratory, Cosmic Physics Center, Batavia, IL 60510, USA*

<sup>8</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

<sup>9</sup>*Berkeley Center for Cosmological Physics, Department of Physics, UC Berkeley, CA 94720, USA*

<sup>10</sup>*Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Shanghai 200030, China*

<sup>11</sup>*Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA*

<sup>12</sup>*Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA*

<sup>13</sup>*Department of Particle Physics and Astrophysics, SLAC National Accelerator Laboratory, Stanford, CA 94305, USA*

## ABSTRACT

We analyze clustering measurements of BOSS galaxies using a simulation-based emulator of two-point statistics. We focus on the monopole and quadrupole of the redshift-space correlation function, and the projected correlation function, at scales of  $0.1 \sim 60 h^{-1} \text{Mpc}$ . Although our simulations are based on  $w\text{CDM}$  with general relativity (GR), we include a scaling parameter of the halo velocity field,  $\gamma_f$ , defined as the amplitude of the halo velocity field relative to the GR prediction. We divide the BOSS data into three redshift bins. After marginalizing over other cosmological parameters, galaxy bias parameters, and the velocity scaling parameter, we find  $f\sigma_8(z=0.25) = 0.413 \pm 0.031$ ,  $f\sigma_8(z=0.4) = 0.470 \pm 0.026$  and  $f\sigma_8(z=0.55) = 0.396 \pm 0.022$ . Compared with Planck observations using a flat  $\Lambda\text{CDM}$  model, our results are lower by  $1.9\sigma$ ,  $0.3\sigma$  and  $3.4\sigma$  respectively. These results are consistent with other recent simulation-based results at non-linear scales, including weak lensing measurements of BOSS LOWZ galaxies, two-point clustering of eBOSS LRGs, and an independent clustering analysis of BOSS LOWZ. All these results are generally consistent with a combination of  $\gamma_f^{1/2}\sigma_8 \approx 0.75$ . We note, however, that the BOSS data is well fit assuming GR, i.e.  $\gamma_f = 1$ . We cannot rule out an unknown systematic error in the galaxy bias model at non-linear scales, but near-future data and modeling will enhance our understanding of the galaxy–halo connection, and provide a strong test of new physics beyond the standard model.

**Keywords:** large-scale structure of universe — methods: numerical — methods: statistical

## 1. INTRODUCTION

Clustering analysis of galaxies provides important information for us to understand the spatial distribution and evolution of the underlying dark matter. The relationship between luminous galaxies and dark matter can serve as a constraint on galaxy formation physics, which is necessary for an unbiased determination of cosmological parameters. The observations from large-scale cosmological surveys, e.g., the Sloan Digital Sky Survey (SDSS-I/II, York et al. 2000;

Abazajian et al. 2009), the Two Degree Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001; Cole et al. 2005), WiggleZ (Drinkwater et al. 2010), BOSS (Dawson et al. 2013), and eBOSS (Dawson et al. 2016) have provided spatial information for millions of galaxies and produced significant results that inform our understanding of the universe. Ongoing and future spectroscopic surveys such as the Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al. 2016), 4MOST (de Jong et al. 2016), PFS (Takada et al. 2014), Euclid (Laureijs et al. 2011, 2012), and NASA’s Nancy Grace Roman Space Telescope (Roman, Green et al. 2012; Dressler et al. 2012; Spergel et al. 2015; Wang et al. 2021) will continue to map the structure of the Universe with unprecedented volumes and precision. The compilation and analysis of these current and future data will enhance our

Corresponding author: Zhongxu Zhai  
 zhongxuzhai@sjtu.edu.cn

\* NASA Einstein Fellow

ability to measure the structure and evolution of the universe in the past billions of years. Using a method developed previously in this series, in this paper we analyze BOSS data to extract cosmological information from clustering on small scales and put constraints on cosmological parameters.

Retrieval of cosmological information on small scales is challenging due to the lack of an accurate and convenient analytic model to describe the non-linear dynamics of dark matter, as well as our incomplete understanding of the baryonic processes that impact the spatial distribution of galaxies. In Zhai et al. (2019), we proposed an emulator approach for galaxy clustering based on  $N$ -body simulations to investigate this problem, targeting massive BOSS galaxies at  $z = 0.55$ . The idea is to run a limited number of high-resolution cosmological simulations that can efficiently sample cosmological parameter space. We then augment these dark matter only (DMO) simulations with a galaxy bias model to compute galaxy correlation functions. Our galaxy bias model is based on the Halo Occupation Distribution (HOD; see Wechsler & Tinker 2018 and references therein). This theoretical template is combined with Gaussian Processes (GP) to model the dependence of galaxy statistics on both cosmology and galaxy bias. The resultant product is able to make predictions for an arbitrary set of parameters, both cosmological as well as those controlling the galaxy–halo connection, in an accurate and efficient manner. Therefore, it is possible to sample this enlarged parameter space to achieve the posterior distribution of unknown cosmological parameters and correctly marginalize over nuisance parameters. We demonstrate that there is more constraining power at non-linear scales than linear scales to constrain the growth rate of structure. In this paper, we apply this method to BOSS galaxies distributed within  $0.18 < z < 0.62$  to constrain various parameters of our cosmological model and measure the growth rate of structure. Since the pioneering works of Heitmann et al. (2009, 2010), Lawrence et al. (2010), and Heitmann et al. (2014), the emulator approach has been widely applied in literature to model the statistics of galaxy and dark matter on non-linear scales. Using the  $N$ -body simulation from the Aemulus Project (DeRose et al. 2019), we have constructed emulators for the halo mass function (McClintock et al. 2019a), the halo bias function (McClintock et al. 2019b), the non-linear power spectra of biased tracers (Kokron et al. 2021) and for investigating galaxy assembly bias (McLaughlin et al. in preparation). This approach has also been used in studies such as the EuclidEmulator (Euclid Collaboration et al. 2019) to model the non-linear corrections to the dark matter power spectrum in preparation for the Euclid mission, the DARK EMULATOR (Nishimichi et al. 2019; Kobayashi et al. 2020; Miyatake et al. 2020) for dark matter halo statistics over a wide redshift range, extensions to  $\Lambda$ CDM cosmologies (Giblin et al. 2019; Ramachandra et al. 2020), baryonic effects in the matter power spectrum (Aricò et al. 2020), the modeling of the Lyman- $\alpha$  forest flux power spectrum (Bird et al. 2019; Rogers et al. 2019; Walther et al. 2020; Pedersen et al. 2020), galaxy lensing and clustering for BOSS-LOWZ

galaxies (Wibking et al. 2017, 2020), weak lensing signal of galaxy clusters (Salcedo et al. 2020), and so on.

Based on the Aemulus suite, Lange et al. (2021) measured the linear growth rate for the BOSS-LOWZ galaxies using clustering measurement on small scales, while Chapman et al. (2021) applied the emulator method of Zhai et al. (2019) and performed a cosmological analysis with eBOSS LRG data using a similar range of scales. This work extends the previous works in three different ways. First, we use a different, and larger, redshift range. Second, we incorporate a model for the effects of galaxy assembly bias. Third, through a scaling parameter, we decouple the velocity field of dark matter halos from that predicted by general relativity (GR), enabling us to constrain departures from GR.

This paper is a direct application of Zhai et al. (2019) to data from BOSS DR12. The method to construct the emulator for the galaxy correlation function remains the same but with two main differences in the analysis: (1) galaxy selection from BOSS DR12 and (2) incorporating assembly bias in the HOD model. Zhai et al. (2019) modeled BOSS-CMASS galaxies at effective redshift  $z \sim 0.55$  with the number density of  $n = 4.2 \times 10^{-3} [h^{-1}\text{Mpc}]^{-3}$ , corresponding to the peak value over the redshift range. However, the flux limit of the selection can introduce incompleteness in terms of luminosity or stellar mass. Thus, in this paper, we apply additional selection criteria to the BOSS galaxies to reduce the effect of sample incompleteness, although this effect is shown to be minor for the clustering analysis (Tinker et al. 2017; Zhai et al. 2017).

In Zhai et al. (2019), the connection between galaxies and dark matter halos is modeled by populating halos in the DMO simulations with a parameterized HOD. In particular, we assumed that the relationship can be simply modeled by  $P(N|M)$ , the probability distribution that a halo of mass  $M$  contains  $N$  galaxies of a given class, combined with parameters for spatial and velocity bias of galaxies *within* halos. This basic form assumes that the galaxy population, and therefore the clustering, is determined by halo mass only. However, result based on  $N$ -body simulations reveals that the halo clustering may depend on properties other than halo mass, which has been referred to variously as assembly bias or secondary bias (Wechsler et al. 2002; Sheth & Tormen 2004; Gao et al. 2005; Harker et al. 2006; Wechsler et al. 2006; Wechsler & Tinker 2018). The investigation of this assembly bias effect has been performed extensively with  $N$ -body simulations and semi-analytical models for galaxy formation and evolution. Depending on which secondary property is studied, the clustering of dark matter halos may have different levels of correlation. This effect can, theoretically, propagate into the distributions of galaxies that live in these halos and thus add additional complexity to the galaxy clustering analysis. Current studies have shown small but non-negligible correlation of galaxy and halo clustering with internal properties such as halo age, concentration, or spin, or external properties such as large scale environment (see e.g., Mao et al. 2015; Hearn et al. 2016; Mao et al. 2018; Villarreal et al. 2017; Lehmann et al. 2017; Shi & Sheth 2018;

Salcedo et al. 2018; Zentner et al. 2019; Contreras et al. 2020 and references therein). Although the precise details and physical reasons for the assembly bias are not fully understood, it is now clear that it is necessary to incorporate this effect into the standard HOD approach to create a robust model for the analysis of BOSS galaxy clustering, and to correctly marginalize over any such effect when obtaining constraints on cosmological parameters. Therefore, in this work, we extend the basic HOD model by adding more parameters to describe the clustering dependence on halo environment.

In addition to yielding unbiased cosmological constraints, searching for the galaxy assembly bias itself is also an active topic of research. However, recent attempts based on SDSS and BOSS galaxies present contradictory results on the significance of galaxy assembly bias (Vakili & Hahn 2019; Walsh & Tinker 2019; Salcedo et al. 2020; Yuan et al. 2020). Although the construction of the assembly bias models in these studies are different, it implies that our understanding of the galaxy assembly bias is not complete, necessitating a general and flexible parameterization within the galaxy bias model. Using the emulator approach and the extended HOD model, we will investigate this problem and examine any bias induced in the cosmological constraint in this paper.

The emulator approach applied in this paper enables constraints on fundamental cosmological parameters using small scale galaxy clustering, with an emphasis on the growth of the dark matter structure. Our pilot study (Zhai et al. 2019) demonstrated that small scale clustering has more constraining power than large scales using perturbation theory to measure the parameter combination  $f\sigma_8$ . The result of this analysis can give accurate measurement over the entire redshift range of BOSS galaxies. In addition, our modeling of the dark matter halo velocity field also marginalizes over the modification of underlying gravity. We employ a phenomenological model by scaling the velocity field of dark matter halos with a free parameter. This extra degree of freedom can mimic, in a simplified manner, the effect of modified gravity. Therefore, clustering analyses in redshift space are able to probe deviations from GR. The result can help us better understand the families of cosmological models proposed to explain the cosmic acceleration.

Our paper is organized as follows: In Section 2, we introduce the BOSS galaxies and the sample selection. Section 3 describes the simulation suites used in the analysis. Section 4 lays out the galaxy–halo connection model. Section 5 introduces galaxy statistics for the clustering measurement. Section 6 describes the construction of the covariance matrix and prior for the likelihood analysis. Section 7 presents our cosmological measurements and systematics analysis. We discuss and list our conclusions in Section 8.

## 2. OBSERVATIONAL DATA

In this paper, the analysis uses the large-scale structure catalog created from the BOSS observations as described in Reid et al. (2016), including the survey footprint, veto masks and observational systematics. We use both the galaxy and random catalogs created for clustering measurements. BOSS

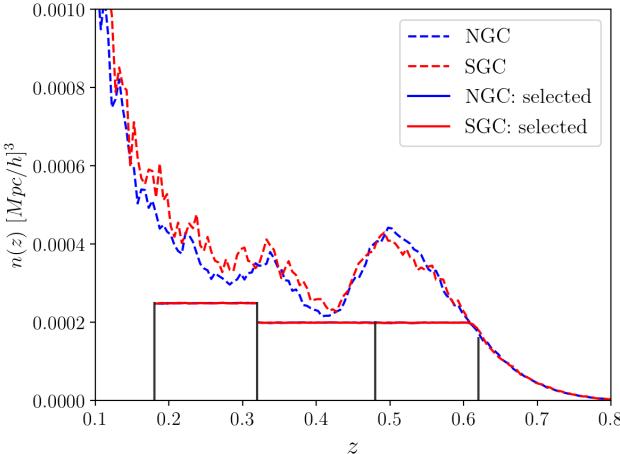
**Table 1.** Number of galaxies used in our analysis.

| Redshift          | Type     | NGC    | SGC    |
|-------------------|----------|--------|--------|
| $0.18 < z < 0.32$ | DR12     | 115187 | 54399  |
|                   | Selected | 82103  | 34249  |
| $0.32 < z < 0.48$ | DR12     | 230093 | 103861 |
|                   | Selected | 158610 | 66509  |
| $0.48 < z < 0.62$ | DR12     | 342844 | 143024 |
|                   | Selected | 209697 | 89950  |

targets galaxies with two selection algorithms: the LOWZ sample at  $z \sim 0.3$  and the CMASS sample at  $z \sim 0.55$ . We follow the strategy of Alam et al. (2017) to use a combined LOWZ+CMASS sample covering the entire redshift range of  $0.2 < z < 0.7$ . We note that the color cuts and flux limits used in target selection introduce incompleteness in the BOSS galaxies, i.e., the BOSS galaxy sample is not a volume-limited sample. A typical halo occupation analysis assumes that the galaxy sample being modeled is volume limited. Although the incompleteness of BOSS galaxies has been quantified in Leauthaud et al. (2016) and Tinker et al. (2017), and the results show that the incompleteness does not have a significant impact on the clustering measurement or HOD analysis of the data (Zhai et al. 2017), this incompleteness should be minimized in order to construct galaxy samples that are best suited for HOD analysis. Therefore we prepare the galaxy selection as follows.

We first compute the galaxy number density,  $n(z)$ , of the BOSS sample assuming a spatially flat  $\Lambda$ CDM with  $\Omega_m = 0.31$  as a fiducial model to compute the cosmic volume. The result is shown in Figure 1 for the North Galactic Cap (NGC) and South Galactic Cap (SGC) respectively. The local minimum at  $z \sim 0.4$  represents the transition between CMASS and LOWZ. We then split the galaxy sample into three redshift slices:  $0.18 < z < 0.32$  (low- $z$ ),  $0.32 < z < 0.48$  (med- $z$ ) and  $0.48 < z < 0.62$  (high- $z$ ) and analyze the clustering measurements separately for each redshift slice. For each subsample, we apply a thin redshift binning with e.g.,  $\Delta z = 0.005$ . In each of these fine slices, we convert the  $i$ -band apparent magnitude of the galaxies to absolute magnitude and rank-order by the luminosity. Then we select the bright end of this galaxy subsample to have a number density that is constant across the redshift slice. For low- $z$ , we use  $2.5 \times 10^{-4} [h^{-1}\text{Mpc}]^{-3}$ . For both med- $z$  and high- $z$ , we use  $2.0 \times 10^{-4} [h^{-1}\text{Mpc}]^{-3}$ . This results a galaxy sample with a constant number density across each of the redshift range of interest, as shown in the solid lines in Figure 1. We apply this selection separately for NGC and SGC and ensure they have the same number density by definition. Table 1 summarizes the number of galaxies in the resultant samples. In Figure 2, we display the distribution of  $i$ -band absolute magnitude for each galaxy sample in NGC and SGC respectively. Although the luminosity threshold varies some across each bin in order to preserve the number density, this is significantly closer to

a volume-limited sample than the original flux-limited target selection. Thus, these new galaxy samples are more appropriate for halo occupation analysis.



**Figure 1.** The comoving number density of the BOSS DR12 galaxies as a function of redshift. The sample used in our analysis is defined by the galaxy brightness such that both NGC and SGC can reach a constant number density, as indicated by the horizontal lines. The grey vertical lines split the galaxies into low- $z$ , med- $z$  and high- $z$  subsamples that will be analyzed individually.

### 3. SIMULATIONS

The non-linear dynamics of dark matter can be well captured by  $N$ -body simulations (see, e.g., Klypin et al. 2011, 2016). This also provides the theoretical framework for us to extract cosmological information from the BOSS galaxies. In this section we introduce the simulation suites employed in our work, as well as examinations of systematics. Table 2 summarizes the key information of these simulation suites and their functions in our analysis. Briefly, we use three different types of simulations: (1) The AEMULUS simulation suite (DeRose et al. 2019), which is used to build the emulator, (2) high-resolution simulations that resolve substructure, which are used to test the emulator, and (3) lower-resolution PM simulations run with the GLAM code (Klypin & Prada 2018) to construct a covariance matrix for the BOSS clustering measurements.

#### 3.1. The Aemulus suite

We use the AEMULUS<sup>1</sup> simulation presented in DeRose et al. (2019) to build and test the emulator for galaxy clustering. The AEMULUS suite comprises 75 boxes of dark-matter-only  $N$ -body simulations, with 40 simulations at different cosmologies for training the emulator, and 35 addi-

tional simulations at 7 cosmologies for testing. The test set has 5 realizations at each cosmology for better statistics. The cosmologies of the training set are chosen based on  $w$ CDM model in an optimized Latin hyper-cube designed parameter space (Heitmann et al. 2009). The cosmological parameters for the AEMULUS suite are the matter density  $\Omega_m$ , the baryon energy density  $\Omega_b$ , the amplitude of matter fluctuations  $\sigma_8$ , the dimensionless Hubble parameter  $h$ , the spectral index of the primordial power spectrum  $n_s$ , the equation of state of dark energy  $w$  and the number of relativistic species  $N_{\text{eff}}$ . All the simulations have a volume of  $1.05h^{-1}\text{Gpc}$  with  $1400^3$  dark matter particles, yielding a mass resolution appropriate for resolving halos that host massive galaxy populations like the BOSS samples. For  $\Omega_m = 0.3$ , the particle mass is  $\sim 3.5 \times 10^{10}h^{-1}\text{M}_\odot$ . We use the training set to build the emulator for the different galaxy clustering statistics, and we use the test set to evaluate the emulator performance and quantify the uncertainty, as done in Zhai et al. (2019). In Table 3, the first section summarizes the cosmological parameters and the range relevant for the emulator construction in the following sections. We refer the readers to DeRose et al. (2019) for more details on the AEMULUS simulations.

#### 3.2. Uchuu and UNIT

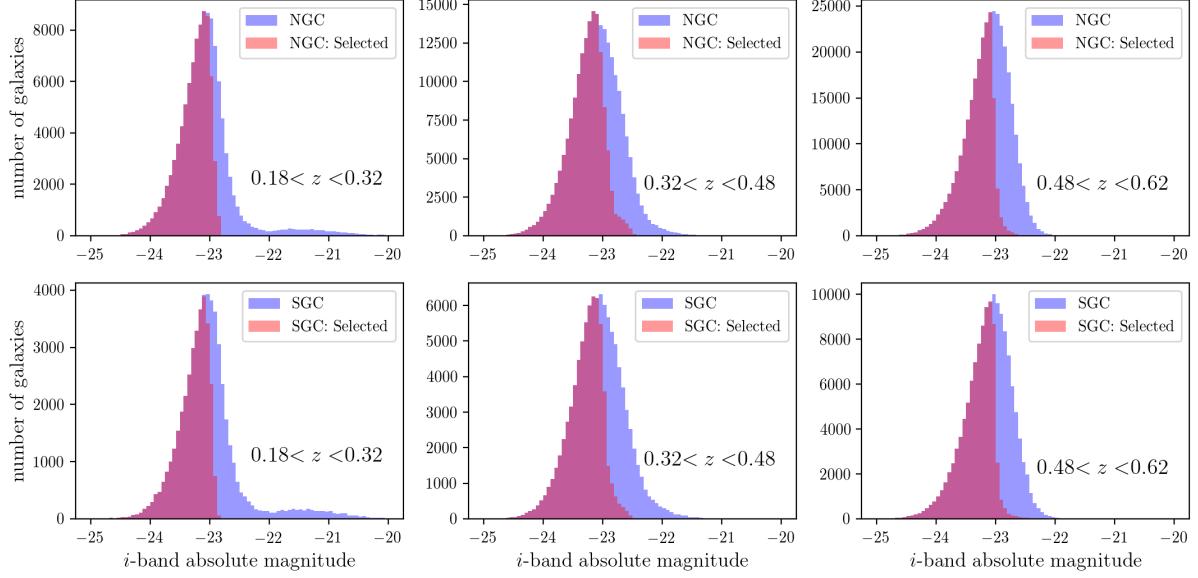
We use the Uchuu<sup>2</sup> (Ishiyama et al. 2020) and UNIT<sup>3</sup> (Chuang et al. 2019) simulations to build galaxy mocks and do recovery tests of our emulator. The galaxy–halo connection in our emulator is based on an HOD in which the parameters of the mean occupation function vary with halo environment, thus mimicking the effect of galaxy assembly bias. In order to validate the robustness of this approach in modeling the galaxy clustering at non-linear scale and the inferred cosmological measurement, we test this approach against mock galaxy catalogs produced via the SubHalo Abundance Matching method (SHAM; i.e., Kravtsov et al. 2004; Vale & Ostriker 2004; Conroy et al. 2006; see Wechsler & Tinker 2018 for a review). This model assigns galaxies to dark matter halos based on the assumption that the stellar mass or luminosity of a galaxy is correlated with the properties of dark matter halo or subhalo hosting this galaxy. The SHAM model has far fewer parameters than HOD but it can well match observed galaxy statistics (e.g., Lehmann et al. 2017). The standard mapping process between galaxies and dark matter halos in the SHAM approach yields some amount of galaxy assembly bias. Therefore testing our HOD-based model with this SHAM model is able to tell us whether different models of assembly bias can bias the cosmological constraints.

We use two simulation suites to create SHAM mocks: Uchuu and UNIT; both adopt the Planck 2015 cosmology (Planck Collaboration et al. 2015). The Uchuu simulation has 2.1 trillion dark matter particles in a  $8[h^{-1}\text{Gpc}]^3$  box with a particle mass of  $3.27 \times 10^8 h^{-1}\text{M}_\odot$ . This volume is  $\sim 8$  times larger than the test boxes from AEMULUS, enabling a

<sup>1</sup> <https://aemulusproject.github.io>

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

<sup>3</sup> <https://unitsims.ft.uam.es>



**Figure 2.** Distribution of the  $i$ -band absolute magnitude of the BOSS galaxies. Our sample corresponds to the brighter end with a nearly hard cut at  $-23$ . This redshift-dependent selection can provide a galaxy sample with constant number density across redshift, which is close to a volume-limited sample.

| Simulation | Cosmology   | Box size                 | Number of simulations | What it is used for   | Reference              |
|------------|-------------|--------------------------|-----------------------|-----------------------|------------------------|
| Aemulus    | multiple    | $1.05 h^{-1} \text{Gpc}$ | 75                    | Building the emulator | DeRose et al. (2019)   |
| GLAM       | Planck-like | $1.0 h^{-1} \text{Gpc}$  | 986                   | Covariance matrix     | Klypin & Prada (2018)  |
| Uchuu      | Planck      | $2.0 h^{-1} \text{Gpc}$  | 1                     | External test of SHAM | Ishiyama et al. (2020) |
| UNIT       | Planck      | $1.0 h^{-1} \text{Gpc}$  | 2 pairs               | External test of SHAM | Chuang et al. (2019)   |

**Table 2.** Simulations used for the analysis in this paper.

more precise measurement of clustering statistics. For the UNIT simulations, we use the  $1h^{-1}\text{Gpc}$  boxes with a particle number of  $4096^3$ . This simulation adopts the inverse phase technique (Angulo & Pontzen 2016) to reduce cosmic variance. We use all the 4 boxes (2 pairs) in our analysis for galaxy clustering, implying that the effective volume is higher than  $4[h^{-1}\text{Gpc}]^3$ . For both simulations, we use the method of Lehmann et al. (2017) to assign galaxies to dark matter halos and subhalos. In this method, the property used to rank halos is a combination of the maximum circular velocity within the halo,  $v_{\max}$ , and the virial velocity,  $v_{\text{vir}}$ . This combination allows the user to vary the amount of assembly bias the galaxies exhibit. More details on the parameters used and the results of the tests are in Appendix E.

We use the GLAM<sup>4</sup> simulations to construct the covariance matrix for the likelihood analysis. The GLAM simulations are run with the new Parallel Particle-Mesh (PM)  $N$ -body code (PPM-GLAM) with a box size of  $1h^{-1}\text{Gpc}$  and  $2000^3$  particles. The high speed of this code has enabled nearly 1000 independent realizations at the Planck cosmology, a set of simulations large enough to construct a robust covariance matrix for all our clustering measurements. In our work, we use 986 boxes with redshift outputs equal to the mean redshifts of our BOSS samples, to estimate the covariance matrix for our BOSS correlation function measurements. The details of the GLAM simulation can be found in Klypin & Prada (2018).

#### 4. GALAXY-HALO CONNECTION MODEL

In this paper, we adopt the HOD approach to model the galaxy–halo connection. The HOD describes the galaxy pop-

#### 3.3. GLAM simulations

<sup>4</sup> <http://www.skiesanduniverses.org/Simulations/GLAM/>

ulation within dark matter halos in a statistical manner. We start with the model of Zhai et al. (2019) to define the mean occupation function for central and satellite galaxies, which is in turn based on the Zheng et al. (2005) HOD model. Our implementation includes three additional parameters that control the concentration of the radial distribution of satellite galaxies and velocity biases of both centrals and satellites. The parameters are summarized in the second section of Table 3. Note that the ranges of certain parameters are enlarged compared to Zhai et al. (2019) to better fit the BOSS measurements. This is due to the fact that the galaxy number density in this paper is lower than that of the Zhai et al. (2019) emulator, thus we are modeling the clustering of an intrinsically brighter galaxy population with a higher clustering amplitude. For reference, in Appendix A, we show the clustering measurements for the BOSS galaxies with and without our selection based on luminosity. Due to the correlation between galaxy luminosity and their host halo mass, these galaxies are likely living in more massive halos and are thus more clustered.

Compared with the previous model in Zhai et al. (2019), the critical change in this analysis is the addition of a galaxy assembly bias model, in which galaxy occupation depends on halo properties other than mass. With the lack of consensus on observational constraints on galaxy assembly bias, there is flexibility to choose secondary halo properties to investigate the assembly bias, including both internal and external properties. In this work, we focus on the bias introduced by the environment, i.e., an external property. In particular, we define the halo environment as the dark matter overdensity of dark matter halo. We measure the relative density  $\delta$  of dark matter halos using a top-hat window function with radius of  $10h^{-1}\text{Mpc}$ . Then we modify the HOD model by scaling the parameter  $M_{\min}$  through

$$\bar{M}_{\min} = M_{\min} \left[ 1 + f_{\text{env}} \text{erf} \left( \frac{\delta - \delta_{\text{env}}}{\sigma_{\text{env}}} \right) \right]. \quad (1)$$

This functional form modulates the dependence of  $M_{\min}$  as a function of local density and its value can be determined by other HOD parameters when galaxy number density is fixed (Zhai et al. 2019). The amplitude parameter  $f_{\text{env}}$  controls the overall strength of the dependence and the resultant level of assembly bias, the position parameter  $\delta_{\text{env}}$  determines the threshold to split halos living in over- and under-dense regions, and the width parameter  $\sigma_{\text{env}}$  controls the smoothness of the transition from under-density to over-density. This assembly bias model allows the minimum mass scale for dark matter halos to host a central galaxy to depend on halo environment, which has a direct impact on the occupancy of centrals and satellites and therefore can change the clustering signal compared with basic HOD model, see e.g., Figure 4 of Walsh & Tinker (2019). The degrees of freedom introduced through this parameterization can also enable the investigation of galaxy formation physics by looking at galaxies formed in halos of same mass but different environment. For the following analysis, setting  $f_{\text{env}} = 0$  can simply turn off the modeling of assembly bias and returns the basic HOD

model. The parameterization chosen allows for significant flexibility in assembly bias; the impact on clustering can be negligible or very large. The change induced by the assembly bias can be to either increase or decrease the clustering amplitude, and the effect can occur at any density.

The HOD model assumes the central occupancy  $N_{\text{cen}}$  approaches 1 at the very massive end, i.e., the most massive halos must host a LRG at the center. However, this is not necessarily true given the target selection of BOSS galaxies and the mass incompleteness (Leauthaud et al. 2016). In order to investigate its impact on our measurement of linear growth rate from small scales, we introduce parameter  $f_{\max}$  to scale the amplitude of  $N_{\text{cen}}$  within the range [0.1, 1.0]. This parameter is equivalent to  $f_f$  in Lange et al. (2021) for a similar discussion. Other earlier works for a similar modeling can be found in Hoshino et al. (2015) and Chapman et al. (2021).

The addition of these new parameters makes our previous HOD model more flexible. The construction of the emulator in this extended parameter space for galaxy correlation function follows the method developed in Zhai et al. (2019). More details, and the performance of the emulator, are described in Appendix C. We note that the choice of the assembly bias model in this paper is not the only option. It is possible to incorporate galaxy assembly bias that depends on other properties of dark matter halos. However, recent studies based on hydrodynamic simulation and semi-analytic model show that the local environment of the halo at the present day is an excellent predictor of assembly bias (Han et al. 2019; Yuan et al. 2020; Xu et al. 2020), as well as results from  $N$ -Body simulation (Yuan et al. 2021). Therefore we apply this particular model of assembly bias throughout this work, but extensions to other model are also possible. In addition, this assembly bias model only applies to the host halos and not to subhalos. Correlations caused by satellites, for example introduced by relationships with concentration, halo formation time etc can be modeled using a more flexible method; but we do not address those here. We note that there are alternatives to extend the basic HOD model with assembly bias, e.g. the decorated HOD algorithm (Hearin et al. 2016; see also McLaughlin et. al. (in preparation)). This model redistributes galaxies within the same halo mass bin by secondary halo property, while preserving the original HOD after marginalization. Our implementation is less physically motivated, but has more flexibility to describe the types of galaxy assembly bias signals induced by various forms the assembly bias may take.

## 5. MEASURING GALAXY CLUSTERING

### 5.1. Two-point Correlation Function (2PCF)

We quantify the clustering for both BOSS and simulated data sets using the two-point correlation function  $\xi(r)$ , which measures the excess probability of finding two galaxies separated by a vector distance  $\mathbf{r}$ , relative to a random distribution, for all  $|\mathbf{r}| = r$ . In practical applications, the distance to galaxies is determined by redshift, which can be distorted by peculiar velocities, also known as the redshift-space dis-

|                       | Parameter             | Meaning   | Range          |
|-----------------------|-----------------------|---|----------------|
| Cosmology             | $\Omega_m$            | The matter energy density   | [0.255, 0.353] |
|                       | $\Omega_b$            | The baryon energy density   | [0.039, 0.062] |
|                       | $\sigma_8$            | The amplitude of matter fluctuations on 8 $h^{-1}$ Mpc scales.          | [0.575, 0.964] |
|                       | $h$                   | The dimensionless Hubble constant                                       | [0.612, 0.748] |
|                       | $n_s$                 | The spectral index of the primordial power spectrum                     | [0.928, 0.997] |
|                       | $w$                   | The dark energy equation of state                                       | [-1.40, -0.57] |
|                       | $N_{\text{eff}}$      | The number of relativistic species                                      | [2.62, 4.28]   |
| HOD                   | $\gamma_f$            | The amplitude of halo velocity field relative to $w\text{CDM+GR}$       | [0.5, 1.5]     |
|                       | $\log M_{\text{sat}}$ | The typical mass scale for halos to host one satellite                  | [14.0, 15.5]   |
|                       | $\alpha$              | The power-law index for the mass dependence of the number of satellites | [0.2, 2.0]     |
|                       | $\log M_{\text{cut}}$ | The mass cut-off scale for the satellite occupation function            | [10.0, 13.7]   |
|                       | $\sigma_{\log M}$     | The scatter of halo mass at fixed galaxy luminosity                     | [0.05, 1.0]    |
|                       | $\eta_{\text{con}}$   | The concentration of satellites relative the dark matter halo           | [0.2, 2.0]     |
|                       | $\eta_{\text{vc}}$    | The velocity bias for central galaxies                                  | [0.0, 0.7]     |
| Assembly Bias         | $\eta_{\text{vs}}$    | The velocity bias for satellite galaxies                                | [0.2, 2.0]     |
|                       | $f_{\text{max}}$      | The incompleteness parameter for central occupancy                      | [0.1, 1.0]     |
|                       | $f_{\text{env}}$      | Amplitude parameter for assembly bias                                   | [-0.3, 0.3]    |
| $\delta_{\text{env}}$ | $\delta_{\text{env}}$ | Position parameter for assembly bias                                    | [0.5, 2.0]     |
|                       | $\sigma_{\text{env}}$ | Width parameter for assembly bias                                       | [0.1, 1.0]     |

**Table 3.** Parameters used in our emulator, their physical meaning, and the parameter space range for each parameter. For the HOD parameters, assembly bias parameters, and for  $\gamma_f$ , the range is used as a flat prior in analysis. For the cosmological parameters, we use a prior defined by the training cosmologies themselves. See text for further details.

tortion (RSD) effect. Therefore the measured galaxy distribution in redshift space is different than in real space, but these differences are driven by the amplitude of the peculiar velocity field and thus contain information about the growth rate of large-scale structure. We measure  $\xi_Z(r_p, \pi)$  on a two-dimensional grid of separations perpendicular  $r_p$  and parallel ( $\pi$ ) to the line of sight through

$$\pi = \frac{\mathbf{s} \cdot \mathbf{l}}{|\mathbf{l}|}, \quad r_p = \mathbf{s} \cdot \mathbf{s} - \pi^2, \quad (2)$$

with  $\mathbf{l} = (\mathbf{s}_1 + \mathbf{s}_2)/2$  (Davis & Peebles 1983; Fisher et al. 1994), the subscript "Z" denotes redshift space. In order to reduce the effect of redshift-space distortions and extract information in real space (with subscript "R" in the following equation), we compute the projected correlation function (Davis & Peebles 1983)

$$w_p(r_p) = 2 \int_0^\infty d\pi \xi_Z(r_p, \pi) = 2 \int_0^\infty d\pi \xi_R(r = \sqrt{r_p^2 + \pi^2}). \quad (3)$$

This integral needs to be truncated at some scale in the measurement from the observational sample or mock catalog. We choose  $\pi_{\text{max}} = 80 h^{-1}\text{Mpc}$ , which is large enough to give stable results. To encapsulate the clustering in redshift space, we measure the multipoles of the correlation function. With  $\mu = r_p/s$ , we use the standard decomposition with Legendre polynomial to obtain

$$\xi_\ell(s) = \frac{2\ell+1}{2} \int_{-1}^1 L_\ell(\mu) \xi_Z(s, \mu) d\mu, \quad (4)$$

where  $L_\ell$  is the Legendre polynomial of order  $\ell$ . Most of the information in redshift space is contained in the first few even multipoles and in this work we use  $\xi_0$  and  $\xi_2$ .

## 5.2. Measurement from BOSS

Using the galaxy samples defined in Section 2, we measure the 2PCF through the estimator (Landy & Szalay 1993)

$$\xi(r_p, \pi) = \frac{DD - 2DR + RR}{RR}, \quad (5)$$

where  $DD$ ,  $DR$ , and  $RR$  are suitably normalized numbers of (weighted) data–data, data–random, and random–random pairs in each separation bin. The positions of the BOSS galaxies are converted from RA, DEC and redshift to Cartesian coordinates assuming a Planck 2015  $\Lambda\text{CDM}$  model with  $\Omega_m = 0.307$ . The choice of cosmology for this transformation has a negligible impact on the final result (Chapman et al. 2021; Lange et al. 2021).

In BOSS clustering measurements, there is an important systematic due to “fiber collisions”. The spectroscopic redshift of BOSS galaxies is obtained by fibers, which have a physical scale of 62 arcsec on a given tile; any two fibers can not be placed closer than this scale. This leads to a fraction of galaxies without redshift determination. The loss of these galaxies has an impact on the clustering measurement at all scales, and is more significant on small scales. This fiber collision effect has been corrected using multiple methods in the clustering measurements, such as nearest neighbor or angular

up-weighting methods. In our analysis, we adopt the method developed in [Guo et al. \(2012\)](#), which is based on the fact that the fiber collision can be resolved in areas that are observed within more than one tile. This method can recover the projected and redshift space two-point correlation function on scales below and above the collision scale. Our final measurements of the BOSS galaxies are shown in Figure 3 for all three redshift bins. For  $w_p$  and  $\xi_\ell$ , we choose logarithmically-spaced bins for  $r_p$  or  $s$  from 0.1 to  $60.2 h^{-1} \text{Mpc}$ , resulting in 9 data points for each statistic. For  $\xi_\ell$ ,  $\mu$  is binned linearly with 40 bins from 0 to 1. Both the measurement from BOSS galaxies and from the mocks use the same binning scheme.

## 6. LIKELIHOOD ANALYSIS

### 6.1. Covariance matrix

The covariance matrix is of critical importance in the likelihood analysis. Here we pursue several methods of quantifying the covariance matrix of the BOSS clustering measurements in order to determine the sensitivity of our results to the details of the matrix construction. It can be estimated using a large number of simulations, or through the data themselves, and the matrix can employ a combination of these two approaches. Based on the observational data, we first measure the covariance matrix through jackknife resampling with 400 roughly equal sub-areas of the BOSS angular footprint. We note that this method may lead to a noise-dominated covariance matrix due to limited size of subsamples. Simulation-based methods have significantly less noise, but have inherent assumptions when constructing the mock galaxies used. To construct these mocks, we use the emulator of the 2PCF as constructed in the previous section to find an HOD model that can give consistent clustering measurements with the data. In particular, we fix the cosmological parameters to be those of the GLAM cosmology, and we also assume no assembly bias (i.e.,  $f_{\text{env}} = 0$ ). The resulting HOD model is then used to populate the halos within the GLAM simulations to produce 986 galaxy mocks. We repeat this process for all three redshifts. The correlation matrix of these two approaches give quite consistent behavior of the galaxy 2PCF, with the fact that the mock-based method is substantially smoother and thus reduces the effect of noise in the calculation of the likelihood.

With these two methods, we construct the following four covariance matrices:

- Jackknife: Full covariance matrix from Jackknife resampling.
- GLAM: Full covariance matrix from the GLAM mocks.
- Fiducial: Diagonal elements from jackknife resampling combined with the correlation matrix from GLAM mocks.
- Modified Fiducial: Same as the fiducial, but using a different HOD model for the galaxy mocks.

As listed above, our fiducial matrix uses the data to set the amplitude of the errors, but we use the simulations for the shape of the covariance matrix. The off-diagonal elements are scaled by the corresponding diagonal elements. In the modified fiducial matrix, we choose a different HOD model from one of the training sets that is close to the 2PCF measurement from BOSS and use this HOD model to regenerate 986 GLAM mocks. This tests any sensitivity to the details of the galaxy bias model. We apply this test to the high redshift  $z = 0.55$  subsample. We find minimal dependence on the details of the HOD model. Further details on this test and the construction of the various covariance matrices can be found in Appendices B and G.

The covariance matrices described above correspond to the contribution from sample variance in the data  $C_{\text{sam}}$ . When we perform the actual analysis, we also need to take into account the uncertainties from the emulator itself. Thus the final covariance matrix is

$$C = C_{\text{sam}} + C_{\text{emu}}, \quad (6)$$

where  $C_{\text{emu}}$  is the intrinsic error from the emulator prediction. To compute this, we adopt the same method as in [Zhai et al. \(2019\)](#). Simply put, the raw emulator performance show as the shaded area in Figure 15 has two contributions: intrinsic error of emulator and sample variance of the testing simulations. We assume these two terms are independent and therefore the intrinsic error of the emulator is the total error with the sample variance subtracted off in quadrature. In addition, we also assume the emulator error is independent among different  $r_p$  and  $s$  bins, thus  $C_{\text{emu}}$  is diagonal only. This differs from [Zhai et al. \(2019\)](#), in which we assumed  $C_{\text{emu}}$  had the same correlation matrix as  $C_{\text{sam}}$ , but our tests show this has a minor impact on the final constraints. A followup analysis is left for a future work (Storey-Fisher et al, in preparation).

### 6.2. Sampling algorithm and priors

We perform the our analysis using the likelihood function

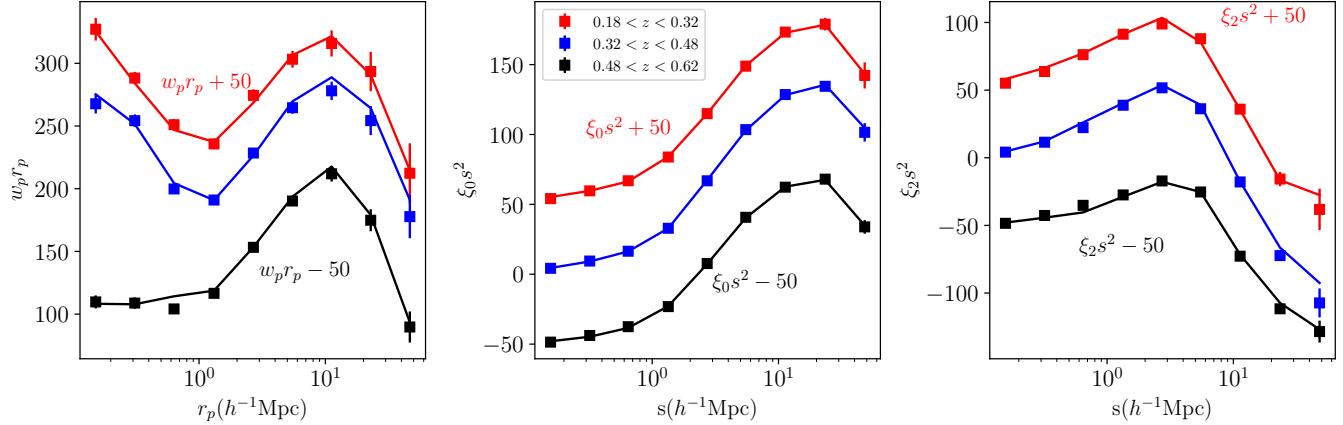
$$\ln \mathcal{L} = -\frac{1}{2}(\xi_{\text{emu}} - \xi_{\text{obs}})C^{-1}(\xi_{\text{emu}} - \xi_{\text{obs}}), \quad (7)$$

where  $\xi_{\text{emu}}$  and  $\xi_{\text{obs}}$  are the correlation function from the emulator and observational data respectively and  $C$  is the covariance matrix as defined above. Depending on tests,  $\xi_{\text{obs}}$  is from either BOSS measurements or galaxy mocks.

The likelihood analysis is done through Bayesian statistics. In particular, we explore the parameter space using the nested sampling algorithm ([Skilling 2004](#)) implemented in MULTINEST ([Feroz et al. 2009; Buchner et al. 2014](#)) package. This method can compute the Bayes Evidence through the integral

$$\mathcal{Z} = \int \mathcal{L}(\mathbf{p})p(\mathbf{p})d\mathbf{p}, \quad (8)$$

where  $p(\mathbf{p})$  is the prior given parameter vector  $\mathbf{p}$ . The Bayes Evidence has been widely used to evaluate the model selection in different scientific fields. The output from MULTINEST also gives posterior distributions of the parameters in



**Figure 3.** The two-point correlation function of BOSS galaxies, including  $w_p$ (left),  $\xi_0$ (middle), and  $\xi_2$  (right) for all three redshift bins. The lines are the prediction of the best-fit model with varying cosmology, HOD and assembly bias parameter, assuming the fiducial model for covariance matrix (see more details in Section 6.1). The best-fit model uses data from  $w_p + \xi_0 + \xi_2$ . The results for low- $z$  and high- $z$  are shifted slightly for plotting purpose.

the model. Compared with traditional methods to obtain the posterior, such as Markov Chain Monte Carlo (MCMC)-like method as we used in Zhai et al. (2019), the nested sampling needs fewer evaluations to reach convergence. In our analysis, we choose 1000 live-point to sample the high-dimensional parameter space. A typical run for likelihood analysis using emulators of all three statistics  $w_p + \xi_0 + \xi_2$  takes roughly five thousand CPU hours to get converged results.

Another critical ingredient is the prior  $p(\mathbf{p})$ . For our HOD parameters, we choose a flat and un-informative prior defined by the range of the parameters (Table 3). However, the cosmological parameter space in our analysis is restricted by the sampling of the training cosmologies, see for instance Figure 3 of DeRose et al. (2019). This means the that emulator is only guaranteed to produce reliable prediction of the 2PCF within this CMB+BAO+SNIa defined area. Therefore we define a prior space for the cosmological parameters based on the training cosmologies. In particular, this prior is defined by an ellipsoid in seven-dimensional space and we restrict the nested sampling to be within this prior range. We present more details of the training area in Appendix D.

### 6.3. Recovery tests

Before we apply this likelihood analysis to the BOSS galaxies, we first perform a recovery test using the SHAM galaxy mocks. The details and results are shown in Appendix E. It shows that with different details for the SHAM mock construction, our HOD-based emulator is able to recover the input cosmology successfully with the parameters constrained within  $1\sigma$  level, and thus validate the emulator construction. In Appendix G, we present the constraints using different setups for the covariance matrix. The consistency between different covariance matrices shows that the effect on the final cosmological measurement is not significant. Compared with the smoother correlation matrix from

GLAM mock, the noise in the jackknife resampling method does not bias the cosmological constraints.

## 7. RESULTS FROM BOSS GALAXIES

In this section, we present constraints using BOSS galaxies. We begin with a presentation of our results when implementing myriad priors and assumptions on the analysis. We then focus on the impact that galaxy assembly bias has on our results. Our results are summed up in our constraints on the growth rate of structure, both when assuming GR or when allowing  $\gamma_f$ , the scaling parameter of the velocity field, to be a free parameter. Last, we compare our results to others in the field that use simulation-based approaches to extract cosmological information from small scales.

### 7.1. Constraints on key cosmological parameters

The fiducial covariance matrix is constructed using the jackknife resampling for the uncertainty and GLAM mocks for the off-diagonal elements of the correlation matrix as explained in the previous section. All results in this section will use this covariance matrix. The observational data for galaxy statistics are  $w_p + \xi_0 + \xi_2$ . In order to have a comprehensive investigation, we perform several tests as follows:

- Fiducial: Varying cosmological parameters + HOD parameters + assembly bias parameters.
- Planck Prior: Gaussian prior on a subset of the cosmological parameters ( $\Omega_m, \Omega_b, \sigma_8, h, n_s$ ) using the latest Planck 2018 measurements. In particular we use the result from the chain `plikHM_TTTEEE_lowl_lowE_lensing`. We use this particular chain as Planck 2018 measurements throughout the paper, unless it is described explicitly. The other cosmological parameters  $w$ ,  $N_{\text{eff}}$  and  $\gamma_r$  have the same prior as the fiducial case.

- Fixing  $\gamma_f = 1.0$ : This forces the analysis to use General Relativity to describe gravity.
- Fixing  $w = -1$ : This forces the analysis assuming  $\Lambda\text{CDM}$ .
- No assembly bias: We set  $f_{\text{env}} = 0$  to turn off the assembly bias modeling.

In Figure 4, we present the constraints on the key parameters of interest ( $(\Omega_m, \sigma_8, \gamma_f)$ ) for the three redshift bins respectively. The constraints on the model parameters are also summarized in Table 4. Using our fiducial priors, all three redshift bins yield similar constraints on these parameters that influence the growth of structure. As expected, there is a degeneracy between  $\sigma_8$  and  $\gamma_f$ , with higher values of  $\sigma_8$  yielding lower values of  $\gamma_f$ . The amplitude of this degeneracy curve lies below the Planck+GR value of  $(\sigma_8, \gamma_f) = (0.82, 1.0)$ . This implies that the peculiar velocity field of BOSS galaxies is roughly 15% lower than the Planck+GR prediction. Although we focus on the cosmological parameters that control the growth of structure, the constraints on all cosmological parameters, for each redshift bin, are shown in Figure 5 and presented in Appendix F.

Although the MULTINEST algorithm adopted in our likelihood analysis is not designed for goodness-of-fit analysis, the posterior produced from the computation allows the search of a minimum for  $\chi^2$  that is close to the result from a minimization algorithm. In our fiducial analysis, the free parameters include 7 cosmological parameters, 1 parameter for dark matter halo velocity field  $\gamma_f$ , and 12 total HOD parameters. The data vector has 9 points for each of  $w_p$ ,  $\xi_0$  and  $\xi_2$ . Therefore the number of degrees of freedom is 8. We find  $\chi^2 = 6.88$  for the low- $z$  subsample,  $\chi^2 = 6.52$  for the med- $z$  subsample, and  $\chi^2 = 16.8$  for the high- $z$  subsample. We note that the high- $z$  subsample gives a large  $\chi^2$  relative to the degrees of freedom. However this result is dominated by the fit to  $\xi_0$  at the smallest radial bin. We recompute the  $\chi^2$  by excluding this single data point using the best-fit model. The resulting value is  $\chi^2 = 12.32$ , indicating a more reasonable result. Although these results imply a statistically good fit to the data, we note that  $\chi^2$  per degree of freedom is only a rough indicator. The data points are correlated, reducing the number of degrees of freedom. However, for the clustering analysis in this paper and our particular attention to the measurement of growth rate of structure, most of the constraining power is only from a subset of parameters, implying that the “effective” number of degrees of freedom is actually *higher* than the 8 listed above (i.e., see the discussion in Lange et al. 2021).

In addition to our fiducial constraints on the key cosmological parameters, Figure 4 also shows the results when adopting a Planck prior for all cosmological parameters. This then yields the value of the halo velocity field relative to GR,  $\gamma_f$ , that is required to match the data. As expected from the fiducial results, when enforcing this prior, the best-fit values of  $\gamma_f$  are all below unity, with values of 0.87, 0.97, 0.84, all with errors of  $\sim 0.05$ . More notable, however, is that the  $\chi^2$  values of the best-fit models are higher than in the fiducial

analysis, with  $\Delta\chi^2 = 5, 4$ , and 5, for low- $z$ , med- $z$ , and high- $z$  respectively, indicating the mild difficulty of the Planck cosmology to fit the data, even with the added freedom of scaling the halo velocity field and an 12-parameter halo occupation model that includes assembly bias.

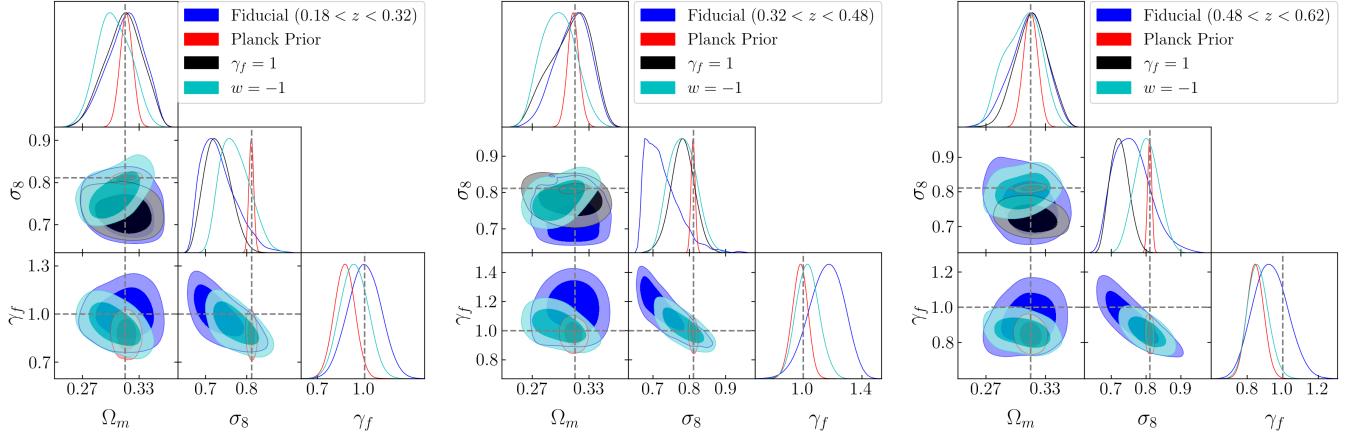
For comparison, these three figures also show results for an analysis in which the cosmological parameters adopt the fiducial prior, but with  $\gamma_f = 1$ . This analysis is thus a fully  $\Lambda\text{CDM+GR}$  fit to the data. These results are indicated with the black contours in the  $\Omega_m-\sigma_8$  panel. In all three redshift slices, the results are consistent with the fiducial analysis, and yield minimal changes to the best-fit  $\chi^2$  values, with  $\Delta\chi^2 \approx 0, 3$ , and 0 from low- $z$  to high- $z$ , respectively. The constraints on  $\Omega_m$  are consistent with the Planck results, but the constraints on  $\sigma_8$  are significantly lower, with best-fit values of  $\sim 0.75$ . This analysis demonstrates that a  $\Lambda\text{CDM+GR}$  model is sufficient to describe the data, but there is tension with the amplitude of clustering inferred from the CMB. The last test is to fix  $w = -1$ , i.e. force the model to be  $\Lambda\text{CDM}$ . Reducing this degree of freedom can change the best-fit  $\chi^2$  values by  $\sim 1-2$  for all three redshift bins. The final constraint on the key cosmological parameters is quite consistent with the fiducial test but we note that fixing  $w = -1$  can alter the degeneracies between some parameter pairs. The improvement on the measurement of linear growth rate is not surprising as in earlier work (Chapman et al. 2021).

To visually demonstrate the impact of these prior assumptions on the fits to the data, in Figure 6 we present the best-fit emulator predictions for the galaxy correlation functions in all redshift bins, with residuals of the fits relative to the data. For  $w_p(r_p)$ , the fits are relatively consistent regardless of assumption. However, the adoption of the Planck prior impacts the amplitude of the monopole, especially for the low- $z$  and high- $z$  redshift slices. It is this change that drives the  $\Delta\chi^2$  values.

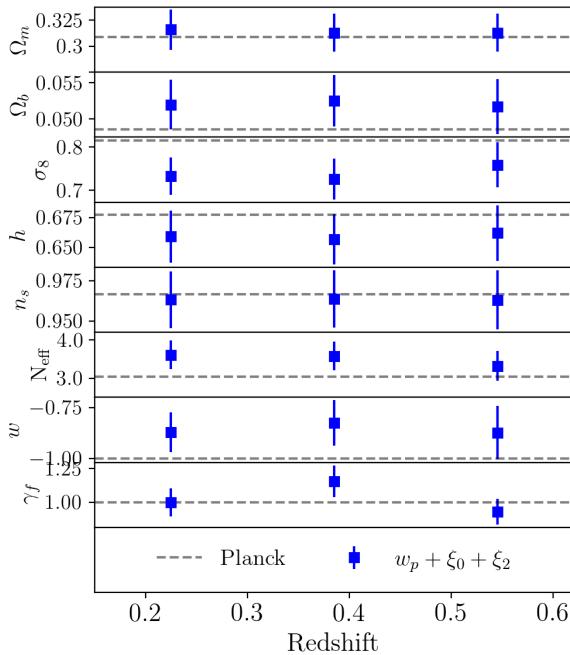
## 7.2. Halo occupation parameter constraints

We note again that we include two parameters that encapsulate velocity bias between galaxies and dark matter. First, the orbits of satellite galaxies within their host halos may move faster or slower than expected from the virial velocity dispersion ( $\eta_{vs}$ ). Second, central galaxies may have non-zero, random, velocities with respect to their host halos ( $\eta_{vc}$ ). The velocity bias for BOSS central galaxies shows deviation from zero with significance of a few  $\sigma$ , and the deviation increases as we go to higher redshift. This indicates that centrals are not at rest with respect to their host halos, consistent with earlier findings in Guo et al. (2015), Yuan et al. (2020), and Lange et al. (2021). Note that we use a slightly different definition of the velocity bias for centrals than these works, but the results are consistent.

The velocity bias for satellites is consistent with unity, indicating the velocity distribution of the satellites is well described by the virial dispersion of the dark matter halos. This is in agreement with Lange et al. (2021), who perform a similar clustering analysis using BOSS galaxies at  $z = 0.25$  and 0.4, as well as with the higher-redshift eBOSS LRG anal-



**Figure 4.** Constraint on some of the key parameters of the low- $z$  (left), med- $z$  (middle), and high- $z$  (right) subsamples, using  $w_p + \xi_0 + \xi_2$ . The contours show 1- and 2- $\sigma$  confidence levels. The result shows a comparison of three different tests: (1) Fiducial test (blue): varying cosmology+HOD+assembly bias parameters with an uninformative prior; (2) adopting a Gaussian prior on a subset of the cosmological parameters ( $\Omega_m, \Omega_b, \sigma_8, h, n_s$ ) using Planck 2018 observation (red); (3) fixing  $\gamma_f = 1$  with no deviation from GR (black); and (4) fixing  $w = -1$  to be  $\Lambda$ CDM (cyan). The dashed lines for cosmological parameters indicate the best-fit model of Planck with  $\gamma_f = 1$  and  $f_{\text{env}} = 0$ .



**Figure 5.** Constraints on cosmological parameters from our analysis. The dashed lines correspond to the best-fit measurement from Planck using the baseline  $\Lambda$ CDM model.

ysis of Chapman et al. (2021). This is in tension at some level with the result in Guo et al. (2015) where the authors report a  $2\sigma$  constraint of  $\alpha_s < 1$  using BOSS DR11 galaxies, but we note their analysis is at fixed cosmology. Using

CMASS galaxies, Yuan et al. (2020) finds that satellite galaxies slightly prefer higher velocity than the dark matter particles but this result is still consistent with ours. However, their result has a large variation depending on the details of the fit. This implies the necessity of more detailed investigation of the velocity field traced by satellites.

In our model, satellite galaxies follow an NFW density profile, with a concentration parameter that is proportional to that of the dark matter ( $\eta_{\text{con}}$ ). For the med- $z$  and high- $z$  bins, the satellite concentration parameter is roughly half that of the dark matter. For the low- $z$  bin, the best-fit value of  $\eta_{\text{con}}$  is higher, but the overall constraints on satellite concentrations are weak. Our constraints in the low- $z$  bin are consistent with those of Lange et al. (2021) at the same redshift. It is important to note that  $\eta_{\text{con}}$  is not degenerate with any of our cosmological parameters.

In all three redshift bins, the data prefer a slight non-zero value of the assembly bias parameter  $f_{\text{env}}$ , the parameter that governs the amplitude and sign of the bias. Positive  $f_{\text{env}}$  implies that the HOD halo mass scale increases at high densities by approximately 10%, lowering the number of galaxies at high densities and reducing the overall amplitude of clustering. For reference, this value of  $f_{\text{env}}$  lowers the large-scale galaxy bias by  $\sim 5\%$  when all other HOD parameters are fixed. However, we note that the statistical significance of the result is  $\lesssim 1\sigma$  in each redshift bin. This is in agreement with some earlier attempts, such as Walsh & Tinker (2019), Salcedo et al. (2020), and Lange et al. (2021), but in contrast to analyses by Zentner et al. (2019) and Yuan et al. (2020). We note that these works employ different models for assembly bias with different data sets and simulations, thus a direct comparison of each result is not straightforward. Galaxy assembly bias may depend on multiple halo properties, such as concentration, age, and environment (Han et al. 2019). Although the local density is an excellent indicator, it is possible

that assembly bias due to other halo properties or their combinations is not fully accounted for. A detailed comparison of the assembly bias models and an evaluation of their impacts on the clustering analysis is necessary but beyond the scope of this paper. In McLaughlin et. al. (in preparation), we apply the similar emulator approach to model projected galaxy correlation function and galaxy lensing signal, and explicitly investigate the impact of assembly bias on the cosmological inference.

For reference, we present the full 1D and 2D constraints on all cosmological and halo occupation parameters in our analysis in Appendix F.

### 7.3. Effect of assembly bias

Incorporation of assembly bias in our emulator enables the investigation of its impact on the cosmological inference. Using our BOSS clustering data, we repeat the analysis by turning off the assembly bias, i.e., assuming a prior of  $f_{\text{env}} = 0$  in the likelihood analysis. The constraints on  $\sigma_8$  and  $\gamma_f$  for three redshift bins are shown in Figure 7. Restricting the analysis in this way has a minimal impact on the key cosmological constraints. The largest shift on the contour plot is seen in the high- $z$  subsample, but it is still well within  $1\sigma$  level. This result is consistent with expectation from the fiducial analysis that the constraint on  $f_{\text{env}}$  shows only statistically weak deviations from 0. For comparison, the predictions of the emulator using the best-fit model without assembly bias is shown as the green dot-dashed curves in Figure 6.

Although this analysis demonstrates that the inclusion of galaxy assembly bias is not required to achieve unbiased cosmological constraints from the BOSS sample, we still must note potential caveats. Although our model is designed with flexibility in mind, it may not mimic all potential impacts of galaxy assembly bias. The assembly bias can be a combination of multiple secondary properties and the choice of the model in a particular analysis can be arbitrary. Our current model only applies to the host halos within the HOD formalism, and not to satellite galaxies separately. This requires a complicated model for the assembly bias (see e.g., Xu et al. 2020). However, bringing in complementary observables can increase our ability to constrain more sophisticated models. With more statistics, such as higher order moments of the redshift space correlation function, void statistics, and galaxy lensing, we can take a greater leap in constraining both galaxy assembly bias and cosmological parameters. By constructing multiple GP-based emulators for galaxy correlation function and excess surface density of galaxy–galaxy lensing, McLaughlin et. al. (in preparation) explicitly investigate the impact on cosmological inference from galaxy assembly bias. Similar to the method used here, the local density of dark matter halo is used as the secondary halo property within the HOD framework. However that work extends the modeling to both centrals and satellites. The result based on CMASS and LOWZ-like mocks shows that the incorporation of assembly in the model can help reduce the bias for cosmological inference, and can become more important at small scales below  $1 h^{-1} \text{ Mpc}$ .

### 7.4. Measurement of structure growth rate

One of the main goals of measuring galaxy clustering at non-linear scale is to precisely measure the growth rate of structure, quantified by the parameter combination  $f\sigma_8$ . In our model, the parameters that impact this quantity are  $\Omega_m$ ,  $\sigma_8$ , and  $\gamma_f$ . Based on the fiducial analysis, our constraints on  $\Omega_m$  are uncorrelated with other parameters and they are generally in agreement with the Planck results, regardless of prior assumptions. There is a clear degeneracy between  $\sigma_8$  and  $\gamma_f$ , seen in the previous figures, which we focus on in Figure 8 by plotting the results from all three redshift slices together. The degeneracy curve traced out by these results is approximately  $\gamma_f \sim \sigma_8^{-2}$ . We note that the results approach the lower limit of  $\sigma_8$  in the cosmological prior, which likely influences the exact shape of this degeneracy curve. The Planck  $\Lambda\text{CDM+GR}$  value lies just outside the  $1 - 2\sigma$  constraints in this plane. The right panel shows the posterior probability of the parameter combination  $\sigma_8^2\gamma_f$  for all three redshift bins, with the Planck value indicated with the vertical lines. Expressed using this parameter combination, the results from the three different redshift bins are in good agreement with one another, and in tension with the Planck value.

The growth of structure parameter,  $f$ , is determined by both the matter density and the amplitude of the halo velocity field. Thus, we one can think of our constraint on  $f\sigma_8$  as  $\gamma_f\Omega_m^{0.55}(z)\sigma_8(z)$ . In detail, we use the CAMB software to compute the exact value of  $f$  given the redshift and cosmology. After marginalizing all the cosmological, HOD, and assembly bias parameters, are constraints are:

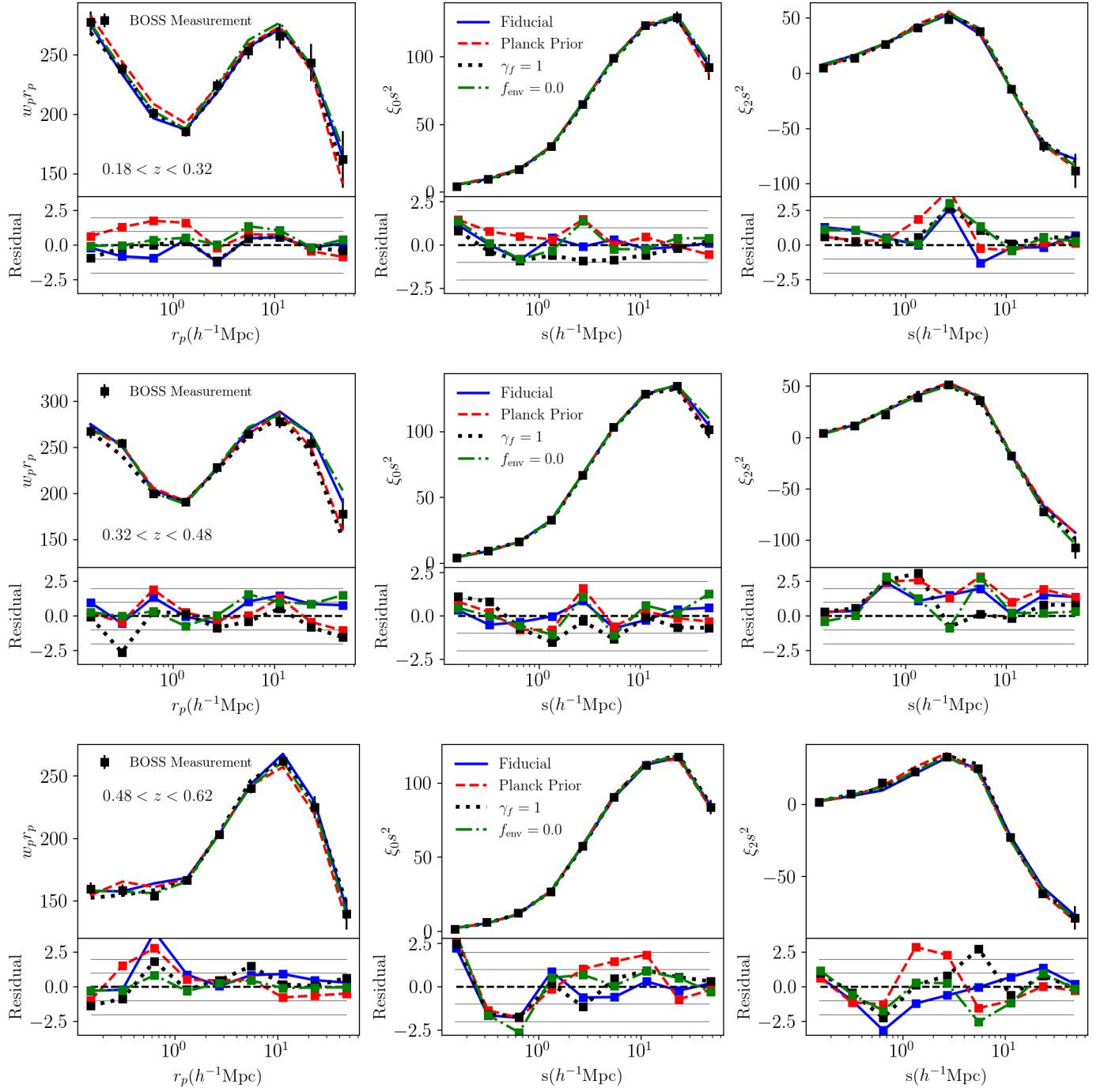
$$f\sigma_8(z=0.25) = 0.413 \pm 0.031 \quad (9)$$

$$f\sigma_8(z=0.4) = 0.470 \pm 0.026 \quad (10)$$

$$f\sigma_8(z=0.55) = 0.396 \pm 0.022 \quad (11)$$

They correspond to fractional errors of 7.8%, 5.6% and 5.5%, respectively, matching the precision expected from Zhai et al. (2019). The corresponding constraints from Planck 2018 are  $0.473 \pm 0.006$ ,  $0.478 \pm 0.005$ , and  $0.474 \pm 0.004$ . Considering the uncertainties of both measurements, we find our measurements are lower than Planck by  $1.9\sigma$ ,  $0.3\sigma$  and  $3.4\sigma$  for the three redshifts. We will discuss these differences in the following section.

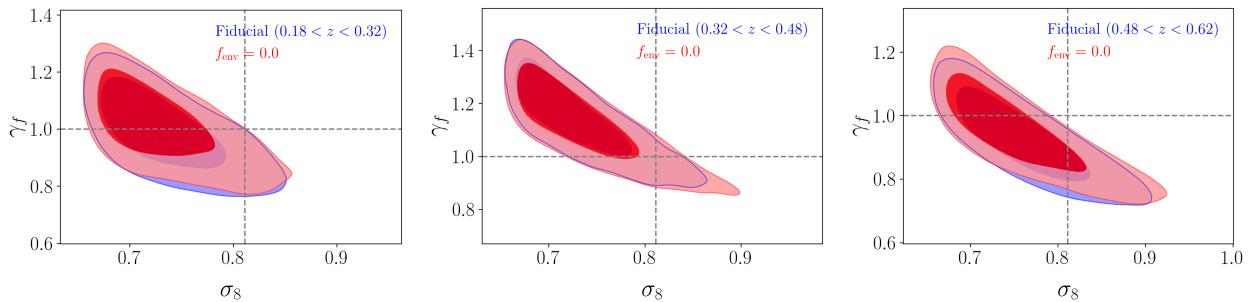
In Figure 9, we display our measurements of  $f\sigma_8$  for each of our three redshift bins, and we compare these values with other results in literature, as well as the prediction from a flat  $\Lambda\text{CDM}$  model using Planck 2018 results. The measurements are collected from clustering analyses of galaxies from surveys including 6dFGS (Beutler et al. 2012), GAMA (Blake et al. 2013), SDSS-I/II main galaxy sample (Howlett et al. 2015, MGS), WiggleZ (Blake et al. 2012), VIPERS (de la Torre et al. 2013) and eBOSS-LRG (Bautista et al. 2021). Note that all these analyses assume GR as the underlying gravity. We also include measurements using BOSS galaxies, using either large-scale or small-scale clustering data. On large scales, Alam et al. (2017) gives the consensus constraints on  $f\sigma_8$  and BAO distance scales over the redshift range of CMASS+LOWZ using RSD multipoles. Their mea-



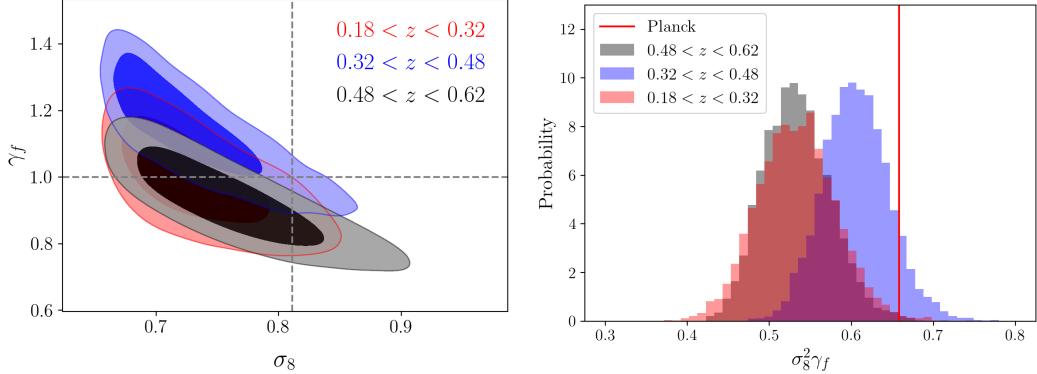
**Figure 6.** Best-fit model of  $w_p$  (left),  $\xi_0$  (middle) and  $\xi_2$  (right) for three subsamples, as well as the residuals with respect to the BOSS measurements. Different lines correspond to the tests as explained in the text. The bottom panels of each plot show the residual normalized by the observed uncertainty, and the horizontal thin lines correspond to  $1$  and  $2\sigma$  level.

**Table 4.** Constraints on the cosmological parameters, HOD parameters and assembly bias parameters using clustering measurement of BOSS galaxies. The training priors for the cosmological parameters are adopted throughout the analysis and the numbers in the parenthesis represent the ratio of 68% interval compared with the range of training space.

| Parameter             | $0.18 < z < 0.32$        | $0.32 < z < 0.48$        | $0.48 < z < 0.62$        |
|-----------------------|--------------------------|--------------------------|--------------------------|
| $\Omega_m$            | $0.315 \pm 0.019$ (40%)  | $0.313 \pm 0.018$ (37%)  | $0.313 \pm 0.018$ (37%)  |
| $\Omega_b$            | $0.052 \pm 0.0035$ (30%) | $0.053 \pm 0.0036$ (31%) | $0.052 \pm 0.0038$ (33%) |
| $\sigma_8$            | $0.733 \pm 0.043$ (22%)  | $0.726 \pm 0.047$ (24%)  | $0.76 \pm 0.052$ (27%)   |
| $h$                   | $0.659 \pm 0.022$ (32%)  | $0.657 \pm 0.021$ (31%)  | $0.663 \pm 0.024$ (35%)  |
| $n_s$                 | $0.963 \pm 0.017$ (50%)  | $0.964 \pm 0.017$ (50%)  | $0.963 \pm 0.018$ (53%)  |
| $N_{\text{eff}}$      | $3.61 \pm 0.378$ (45%)   | $3.582 \pm 0.372$ (45%)  | $3.319 \pm 0.387$ (47%)  |
| $w$                   | $-0.87 \pm 0.105$ (24%)  | $-0.825 \pm 0.112$ (27%) | $-0.874 \pm 0.131$ (32%) |
| $\gamma_f$            | $0.998 \pm 0.104$ (21%)  | $1.155 \pm 0.119$ (24%)  | $0.931 \pm 0.095$ (19%)  |
| $\log M_{\text{sat}}$ | $14.3 \pm 0.07$ (9%)     | $14.44 \pm 0.127$ (17%)  | $14.43 \pm 0.21$ (28%)   |
| $\alpha$              | $1.2 \pm 0.148$ (16%)    | $0.978 \pm 0.154$ (17%)  | $0.839 \pm 0.231$ (26%)  |
| $\log M_{\text{cut}}$ | $11.64 \pm 0.99$ (53%)   | $11.78 \pm 0.96$ (52%)   | $12.44 \pm 1.34$ (73%)   |
| $\sigma_{\log M}$     | $0.435 \pm 0.156$ (35%)  | $0.336 \pm 0.163$ (36%)  | $0.7 \pm 0.145$ (32%)    |
| $\eta_{\text{vc}}$    | $0.154 \pm 0.092$ (26%)  | $0.135 \pm 0.083$ (24%)  | $0.301 \pm 0.094$ (27%)  |
| $\eta_{\text{vs}}$    | $1.0 \pm 0.088$ (10%)    | $1.131 \pm 0.107$ (12%)  | $1.074 \pm 0.112$ (12%)  |
| $\eta_{\text{con}}$   | $1.227 \pm 0.342$ (38%)  | $0.589 \pm 0.189$ (21%)  | $0.606 \pm 0.24$ (27%)   |
| $f_{\text{max}}$      | $0.819 \pm 0.114$ (25%)  | $0.738 \pm 0.13$ (29%)   | $0.82 \pm 0.124$ (27%)   |
| $f_{\text{env}}$      | $-0.054 \pm 0.091$ (30%) | $0.022 \pm 0.109$ (36%)  | $-0.092 \pm 0.084$ (28%) |
| $\delta_{\text{env}}$ | $1.261 \pm 0.481$ (64%)  | $0.961 \pm 0.411$ (55%)  | $1.155 \pm 0.45$ (60%)   |
| $\sigma_{\text{env}}$ | $0.551 \pm 0.257$ (57%)  | $0.588 \pm 0.248$ (55%)  | $0.577 \pm 0.25$ (56%)   |



**Figure 7.** Constraint on parameter  $\sigma_8$  and  $\gamma_f$  from the fiducial constraint (blue) and forcing  $f_{\text{env}} = 0$  (red), i.e., assuming a model with no assembly bias. The dashed line denotes the Planck observation for  $\sigma_8$  and  $\gamma_f = 1$ .



**Figure 8. Left:** Constraint on parameters  $\sigma_8$  and  $\gamma_f$  for all three BOSS subsamples in the fiducial case. **Right:** One-dimensional distribution of the parameter combinations  $\sigma_8^2 \gamma_f$ . The vertical red line shows the measurement from Planck 2018 observation with  $\gamma_f = 1$ .

surement and uncertainty of  $f\sigma_8$  shows a clear dependence on redshift. We linearly interpolate their results to our redshift and compare the constraints. At  $z=0.4$  and  $0.55$ , our fiducial result is consistent within  $0.4\sigma$  and  $1.3\sigma$ , respectively, indicating the internal consistency of the BOSS analysis.<sup>5</sup>

On small scales, there are a number of other studies. Lange et al. (2021) perform the measurements using cosmological evidence modelling approach for BOSS galaxies at  $z = 0.25$  and  $0.4$ , marginalized over AEMULUS cosmological models. Chapman et al. (2021) applies the emulator approach to model the eBOSS LRG at  $z = 0.7$  and extract the measurement of linear growth rate that is close to our method. Reid et al. (2014) present a  $2.5\%$  measurement of  $f\sigma_8$  for the CMASS galaxies, but we note that this analysis is at fixed cosmology and thus the error is likely underestimated (see the discussion in Zhai et al. 2019).

We also extract constraints on structure growth assuming  $\gamma_f = 1$ , which is consistent with most of the other analyses using GR as underlying gravity model. The results are

$$f\sigma_8(z=0.25) = 0.416 \pm 0.022 \quad (12)$$

$$f\sigma_8(z=0.4) = 0.448 \pm 0.025 \quad (13)$$

$$f\sigma_8(z=0.55) = 0.401 \pm 0.019 \quad (14)$$

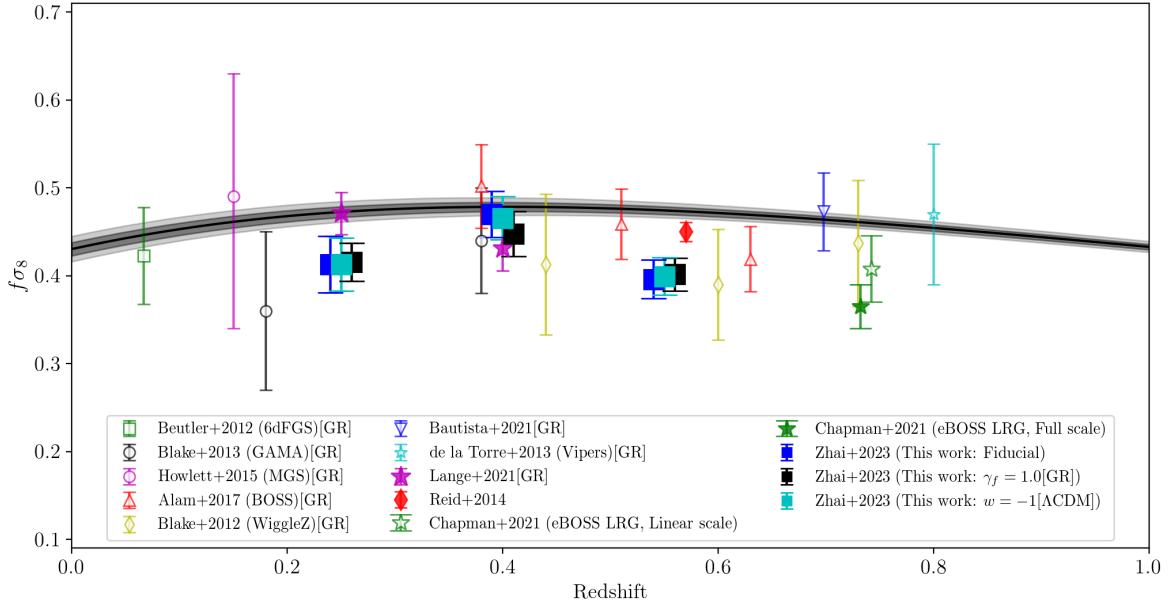
corresponding to a fractional error of  $5.2\%$ ,  $5.7\%$  and  $4.7\%$  respectively. This result has similar accuracy as Lange et al. (2021) where the authors claim a five percent measurement of  $f\sigma_8$  at  $z = 0.25$  and  $0.4$ . Our constraints for the samples at  $z = 0.25$  and  $z = 0.4$  are somewhat tighter because we include the projected correlation function,  $w_p$ , in the analysis to anchor the galaxy bias in real space. Although  $w_p$  has limited cosmological sensitivity, it can strengthen the constraints on the HOD parameters themselves, which can yield tighter cosmological parameters by breaking degeneracies be-

tween cosmology and bias that exist in the RSD data. The tension of our measurements with Planck is  $2.5\sigma$ ,  $1.3\sigma$  and  $3.7\sigma$  for three redshifts. In general, the  $\gamma_f = 1$  prior raises the values of  $f\sigma_8$ , closer to the Planck values. However, the reduced uncertainties make the tension more significant. We also present the measurement of  $f\sigma_8$  with  $w = -1$  in Figure 9. This prior yields quite consistent measurement as our fiducial analysis but can shrink the uncertainty slightly. It is known that a cosmology with  $w \neq -1$  predicts different growth history of structure compared with  $\Lambda$ CDM cosmology (Lue et al. 2004). In our analysis, assuming this prior  $w = -1$  mainly changes degeneracy for some of the parameters, for instance the contour plot for  $\Omega_m$  (Figure 4). However, it doesn't alter the estimate of  $f\sigma_8$  significantly due to the velocity scaling parameter which is flexible to model a wide range of growth history.

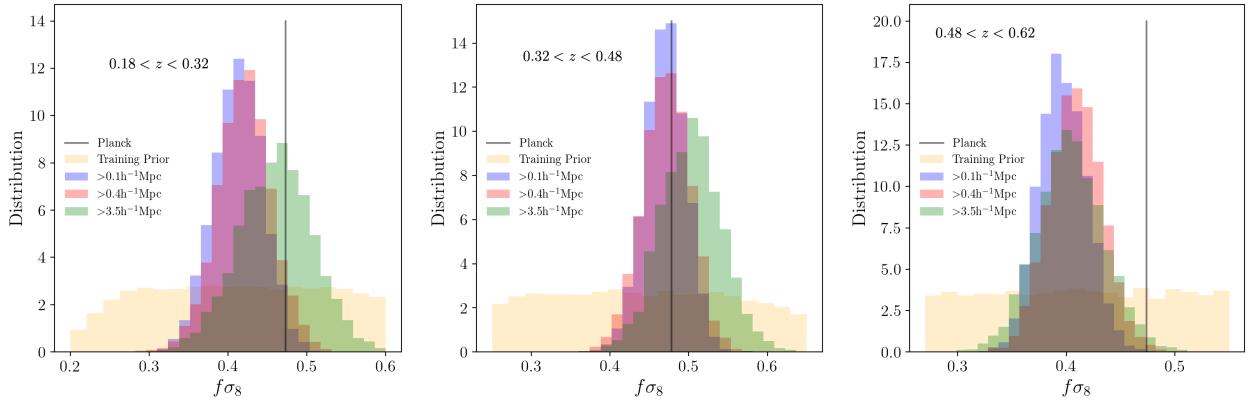
### 7.5. Scale Dependence

In addition, we investigate the impact of scales considered in the clustering analysis. Figure 10 compares the posterior probabilities of  $f\sigma_8$  for our fiducial analysis, which has a minimum scale of  $0.1 h^{-1}\text{Mpc}$ , to two other analyses that increase the minimum scale used. Here we choose minimum scales of  $0.4 h^{-1}\text{Mpc}$  and  $3.5 h^{-1}\text{Mpc}$ . The former choice is motivated by the scale of the fiber collision effect; the latter choice is motivated by the transition between the one-halo and two-halo terms in the galaxy clustering signal. The results for a minimum scale of  $0.4 h^{-1}\text{Mpc}$  are consistent with those of the fiducial scale, both in terms of the values of  $f\sigma_8$  and their uncertainties. However, the results for  $3.5 h^{-1}\text{Mpc}$  are substantially different and the offset decreases with redshift. Although the new posteriors overlap with the fiducial, the best-fit values increase and the errors, as expected, widen. The tension with the Planck results is substantially alleviated, although the high- $z$  result is still low by  $\sim 2\sigma$ . These results agree with similar tests in Chapman et al. (2021), which imply that the lower values of  $f\sigma_8$ , and the tension with Planck, is driven by the fully non-linear regime.

<sup>5</sup> We note that the BOSS measurements are correlated due to overlapping redshift bins, thus the fact that our results are lower than the Alam et al. (2017) constraints at both redshifts is not necessarily indicative of a systematic bias.



**Figure 9.** Measurement of the growth rate of structure from our analysis using BOSS DR12 galaxies, as well as a compilation of the results in literature. The black line with shaded area is the prediction from the Planck 2018 release assuming a flat  $\Lambda$ CDM cosmology. The data include 6dFGS (Beutler et al. 2012), GAMA (Blake et al. 2013), SDSS-I/II main galaxy sample (Howlett et al. 2015, MGS), WiggleZ (Blake et al. 2012), Vipers (de la Torre et al. 2013) and eBOSS-LRG (Bautista et al. 2021). In addition, we also display measurements using BOSS galaxies similar to our work: DR12 final consensus results (Alam et al. 2017), BOSS CMASS RSD analysis (Reid et al. 2014) and BOSS LOWZ small scale analysis (Lange et al. 2021). Note that our results of  $f\sigma_8$  is quoted from measurement of  $f\sigma_8\gamma_f$ . We omit  $\gamma_f$  in the y-axis to be in line with the measurements used for cosmological implications. In other words, except our analysis (including Chapman et al. 2021) and Reid et al. (2014), all the other studies implicitly assume  $\gamma_f = 1$ , i.e., the gravitational interaction is described by GR. The symbols are split into two categories: large scale (open) and small scale (filled) measurements. Note that our measurements and eBOSS results at the same redshifts are shifted slightly for plotting purpose.



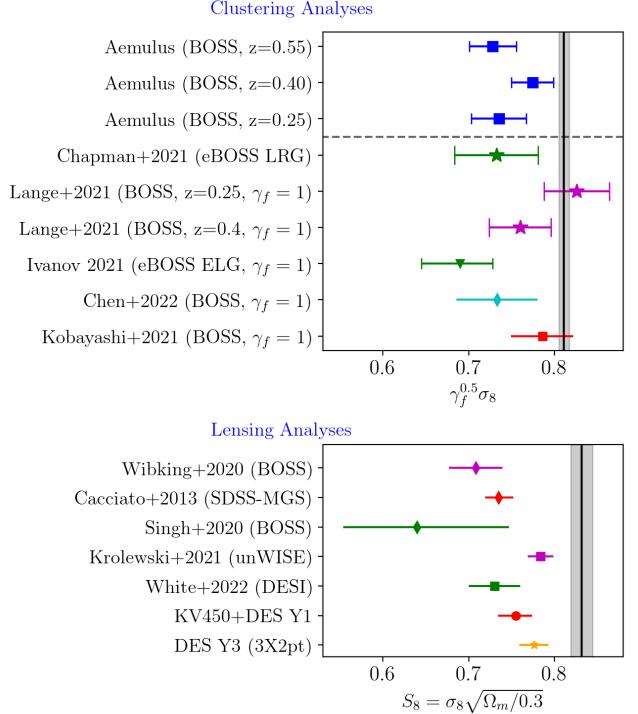
**Figure 10.** Scale dependence of the constraint on  $f\sigma_8$  from BOSS galaxies. Results are shown for three redshift ranges respectively as indicated in each panel. The range covered by the training prior is marked as yellow area.

### 7.6. Comparison with other studies using small-scale clustering

As shown in Figure 9, the small-scale, BOSS LOWZ analysis of Lange et al. (2021) is in good agreement with our constraints on  $f\sigma_8$  at  $z = 0.4$ , but it is higher than ours at  $z = 0.25$ . Even though these analyses use the same set of galaxies at this redshift, there are a number of differences in both the modeling and in the data that may drive this difference. Lange et al. (2021) employs a novel statistical method using the AEMULUS simulations without explicitly constructing an emulator. Rather, they use the likelihood at each AEMULUS cosmology to fit the likelihood function of  $f\sigma_8$ . On the data side, the galaxy statistics used include anisotropic clustering in redshift space of moments up to hexadecapole, but not  $w_p$ . As shown in Chapman et al. (2021), including  $w_p$  reduces  $f\sigma_8$  relative to the multipoles alone. In addition, their analysis is restricted to scales above  $0.4h^{-1}\text{Mpc}$ , which roughly corresponds to the fiber collision scale of BOSS LOWZ galaxies. Last, on the data side, is that Lange et al. (2021) only uses the NGC, whereas our analysis uses both NGC and SGC data. On the modeling side, there are significant differences as well. Lange et al. (2021) assumes  $\gamma_f = 1$  (GR only), which we find also marginally changes the value of  $f\sigma_8$ . Although one single explanation does not seem able to explain the difference in these two analyses at  $z = 0.25$ , the cumulative effect of all the differences can explain a significant amount.

As a direct application of the emulator approach, Chapman et al. (2021) measure small-scale galaxy clustering of eBOSS LRGs at  $z = 0.7$  to constrain the growth rate of structure. This analysis is closest in spirit to our work, with the exception of our incorporation of assembly bias. They use the emulator of Zhai et al. (2019), updated to match the redshift and number density of eBOSS, and include the  $f_{\max}$  parameter discussed above. Chapman et al. (2021) present two constraints on  $f\sigma_8$ : one using clustering data down to  $0.1 h^{-1}\text{Mpc}$ , and another restricted to scales about  $7 h^{-1}\text{Mpc}$ . The full-scale analysis yields a value of  $f\sigma_8$  that, like our BOSS measurements, is significantly below the Planck value. The larger-scale analysis lies in between, with errors such that it is consistent with both the small-scale result and the Planck value. Chapman et al. (2021) also perform a number of notable tests that help to validate the emulator approach taken in this paper. Since measurements of the galaxy correlation function assume a cosmology in order to convert redshift to distance, this choice may lead to the so-called Alcock–Paczynski (AP, Alcock & Paczynski 1979) effect. Chapman et al. (2021) examine this effect for small-scale clustering, finding that the impact on parameter constraints is negligible. Therefore we do not explicitly model this effect in our analysis. In addition, the eBOSS galaxies at higher redshift can experience significant errors in the galaxy redshifts, leading to a biased measurement of  $f\sigma_8$  of up to  $0.5\sigma$ . Since BOSS galaxies are at lower redshift, the amplitude of redshift errors are significantly smaller. However, to be certain that there is no impact on our results, in Appendix H, we examine this by running parameter recovery tests using test data that incorporate red-

shift uncertainty. Our results show that the impact of redshift uncertainty is negligible for the BOSS galaxies.



**Figure 11.** **Top:** measurement of  $\gamma_f^{0.5} \sigma_8$  from our BOSS galaxy analysis (blue square), compared with other works using similar galaxy statistics at both linear and non-linear scales. **Bottom:** latest measurements of  $S_8 = \sigma_8 \sqrt{\Omega_m / 0.3}$  using galaxy lensing statistics. The vertical lines in both panels correspond to the Planck measurements (Planck Collaboration et al. 2020).

In Figure 11, we project our measurement of  $\gamma_f^{0.5} \sigma_8$  and compare with other analysis in literature, including the aforementioned works in Lange et al. (2021), Chapman et al. (2021), Ivanov (2021), Kobayashi et al. (2021) and Chen et al. (2022) using redshift-space distortion, and galaxy lensing analyses from Singh et al. (2020), Wibking et al. (2020), Krolewski et al. (2021), White et al. (2022), Asgari et al. (2020), and Abbott et al. (2022). In the top panel, we express the RSD measurement by parameter combination  $\gamma_f^{0.5} \sigma_8$ , which basically measures the amplitude of matter perturbation. Compared with the Planck result, all clustering analysis using low redshift galaxies, with the exception of the  $z \sim 0.25$  bin from Lange et al. (2021), give lower estimates of the perturbation amplitude than Planck (including Kobayashi et al. (2021) which is slightly more consistent with Planck). Note that the eBOSS LRG (Chapman et al. 2021) result with  $\gamma_f = 1$  has a consistent measurement with Planck, but with a high  $\chi^2$ . In the bottom panel, we compile some of the latest galaxy lensing measurements of  $S_8 = \sigma_8 \sqrt{\Omega_m / 0.3}$ . The overall results of our clustering analysis are in agreement with these lensing analyses that the lensing amplitude is also lower than

Planck prediction (Leauthaud et al. 2017). A more robust analysis can combine the galaxy lensing measurement and RSD signals to improve the accuracy and the constraint on any deviation from GR or  $\Lambda$ CDM and also incorporate more flexible model for galaxy–halo connection (Zu 2020).

## 8. DISCUSSION AND CONCLUSION

Galaxy clustering at small scales has been demonstrated to have a significant amount of cosmological constraining power, especially for the parameters that govern the growth and amplitude of structure. This paper extends the emulator approach for modeling galaxy clustering developed in Zhai et al. (2019) to analyze the clustering of BOSS galaxies. In addition to the standard cosmological parameters of  $\Omega_m$  and  $\sigma_8$ , we introduce the parameter  $\gamma_f$  to scale the velocity field of dark matter halos to mimic the effect of modified gravity. At all redshifts covered by BOSS, our constraints on the parameter combination  $f\sigma_8$  are below those predicted by the  $\Lambda$ CDM+GR model assuming the current Planck cosmology, with varying levels of statistical significance at each redshift bin. This result is similar to that of the “lensing is low” phenomenon (Leauthaud et al. 2017) in which the galaxy–galaxy lensing signal of the BOSS galaxies is lower than that predicted assuming a Planck cosmology by 30% (Leauthaud et al. 2017; Wibking et al. 2020). As shown in Figure 11, there are a growing number of results using small-scale clustering and lensing that are in tension with the current Planck cosmology.

There are three primary ways that this tension between our galaxy clustering results and the Planck constraints can be ameliorated. These include: (1) new physics that imparts deviations from general relativity, (2) cosmological solutions in the form of massive neutrinos, and (3) astrophysical solutions rooted in galaxy formation processes.

The first of these solutions, in which gravity deviates from GR in order to change the prediction for the halo velocity field at a fixed matter density, is not favored. The first reason is that this class of explanation does not necessarily resolve the tension between the Planck cosmology and the aforementioned lensing results. For example, models with weaker gravity than GR boost the lensing signal in the two-halo term (Leauthaud et al. 2017), but would reduce the value of  $f\sigma_8$  (Samushia et al. 2014). This resolves one tension at the cost of amplifying the other. But the second reason is that, by our own analysis, such a solution is not favored by the redshift-space distortions. In our fiducial analysis,  $\gamma_f = 1$  (i.e., gravity is GR) is within  $\sim 1\sigma$ . If we adopt a Planck prior for all cosmological parameters except  $\gamma_f$ , we do find that  $\gamma_f < 1$  to high significance, but this model is not a good fit to the data. However, when allowing the cosmology to be free but adopting a prior of  $\gamma_f = 1$ , our model is a good descriptor of the BOSS clustering data.

The second solution addresses the tensions created by both lensing and clustering results simultaneously. From Figure 11, the collected results from both of these probes indicate the need for a lower value of  $\sigma_8$  than derived from CMB data. The values of these two constraints on the amplitude of

matter fluctuations, one at  $z < 1$  and the other at  $z \approx 1, 100$ , can be reconciled to some degree through the presence of massive neutrinos. The presence of massive neutrinos suppresses the growth of structure, which manifests in a scale-dependent manner with smaller scales being more highly affected. Given the current constraints on sum of neutrino masses to be  $< 0.12$  eV, it is unlikely that massive neutrinos can fully resolve the tension between the results listed above and the Planck constraints on the clustering amplitude, but it goes in the right direction to resolve both lensing and RSD results, and it could alleviate a significant amount of the tension.

The last of the primary solutions involves the galaxy–halo connection. In this paper, we have attempted to make our galaxy bias model as flexible as possible, constructing a model with 11 free parameters. However, the tension between our results and the Planck cosmology is primarily driven by the clustering signal at  $\lesssim 3 h^{-1}\text{Mpc}$ , a scale that probes galaxy pairs within a single host halo as well as the transition between 1-halo and 2-halo galaxy pairs. Thus, a systematic error in the HOD approach cannot be ruled out. The tests presented in this paper use abundance matching techniques to create the test data. These test many of the assumptions in our model, such as spherical halos with isothermal velocity distributions and NFW density profiles for satellite galaxies. However, subhalo abundance matching does not incorporate baryonic effects that may influence the spatial distribution of galaxies. Currently, hydrodynamic simulations are not large enough to present statistically robust tests of the emulator. But it may be possible to incorporate the impact of baryons on the spatial distribution of galaxies in larger, dark matter only simulations, providing a more rigorous test (Hearin et al. 2016, 2021). The effects of galaxy formation on the galaxy–halo connection, and thus galaxy clustering, may present itself either through assembly bias or in an occupation function that is not well-represented with our current model.

The chief observational systematic, that of fiber collisions between BOSS spectra, is not likely to be a dominant source of bias. We note that the discrepancy with the Planck cosmology persists if we exclude data below  $0.4 h^{-1}\text{Mpc}$ , the scale at which collisions become significant in the highest redshift bin. We also note that the method of correcting for collisions employed in this paper is distinct from the method used in the Chapman et al. (2021) emulator analysis of eBOSS LRG clustering, but both studies are consistent in finding tension with Planck when including the smallest scales. On the other hand, we have incorporated the emulator uncertainty throughout the analysis which is non-negligible in the total covariance, implying that we can have another boost in the constraining power if the emulator error can be reduced significantly or totally removed. This can be non-trivial. In our pilot study (Zhai et al. 2019), we use the analytical model of  $w_p$  (Tinker et al. 2005; Tinker et al. 2012) with full control of the uncertainty of the input training data and test data. With similar coverage of the parameter space as in this paper, the intrinsic emulator error can be reduced to sub-percent

level across a wide range of scales, thus much lower than the data uncertainty. However the simulation-based emulator is worse and the performance varies with scale. This may indicate factors other than the GP algorithm that can weaken the emulator performance, including but not limited to the modeling of the training error and the specifics of the simulations such as volume, resolution and small scale density field. A more thorough exploration for the emulator error can help to answer this question and will be useful for future studies.

For the last two of the proposed solutions listed above, the AEMULUS Project is in excellent position to make significant progress. The natural next step in development of simulations for emulator development is to create new simulations that incorporate neutrinos as an active particle species along with the dark matter. This will properly model the differential growth of structure induced by such particles, giving our next generation emulators the ability to incorporate and constrain the sum of the neutrino masses. These simulations are currently underway (DeRose et al, in preparation). Adding new parameters to the galaxy bias model, be they applicable to the mean occupation function or related to assembly bias, is a straightforward process. The addition of new parameters in a model necessitates increased constraining power from data. The AEMULUS approach, however, is open-ended. Non-standard galaxy clustering statistics, such as void statistics, marked statistics, and kNN statistics, can be emulated with equal efficacy to redshift-space distortions and weak lensing. These statistics bring in complementary information that can be used to constrain new freedom in HOD models (Tinker et al. 2008; Walsh & Tinker 2019; Vakili & Hahn 2019; Wang et al. 2019; Szewciw et al. 2021). The next generation of AEMULUS emulators will include complementary, non-standard statistics (Storey-Fisher et al, in preparation).

The measurement of the growth and amplitude of large-scale structure is a crucial test of our cosmological model. It provides complementary information to geometric probes, i.e., the cosmic distance measurements from observation of Type Ia supernovae and baryon acoustic oscillations. The measurement of  $f\sigma_8$  as a function of redshift is able to constrain the growth history of the cosmic density field. For the large families of dark energy and modified gravity models that are proposed to explain cosmic acceleration, the growth measurement is able to distinguish them and pare down the viable parameter space. The accurate and model-independent measurement presented in this paper serves as the latest at-

tempt. In future work, we will explore the cosmological implications of this latest measurement.

The AEMULUS Project aims at providing accurate and unbiased emulators for galaxy statistics at any redshift and number density. This paper represents the first application to the BOSS dataset. The overall performance is consistent with the estimates of our earlier work, and additional improvement is also possible. Ongoing surveys, most notably DESI, will increase the statistical constraining power of observational data and will thus require higher precision and accuracy from theoretical models. The next generation of Aemulus simulations and emulators will allow the analysis of small-scale clustering to continue down the path on which this paper lies.

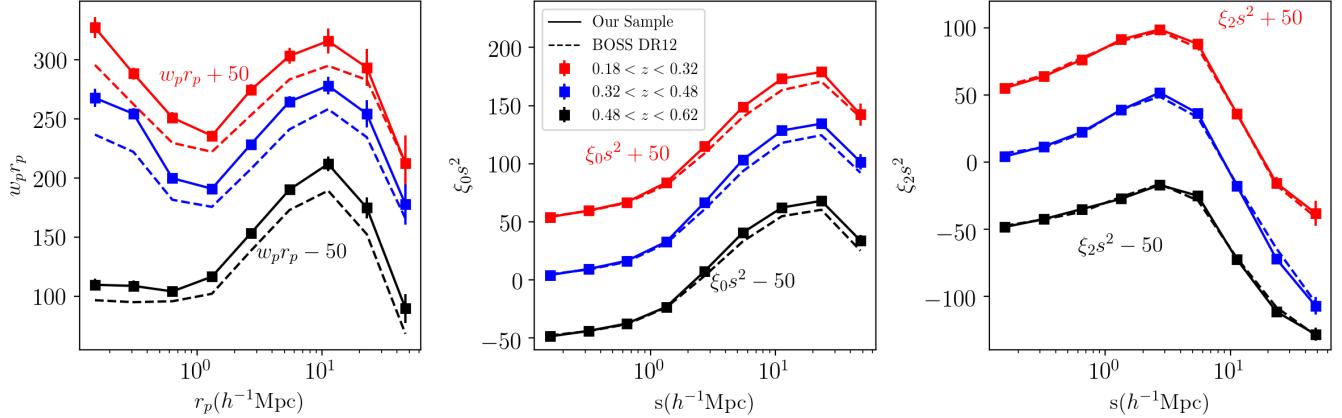
ZZ thanks Michael Chapman, Johannes Lange, Surhud More, and Sukhdeep Singh for kindly providing their data. This work received support from the U.S. Department of Energy under contract number DE-AC02-76SF00515 and from the Kavli Institute for Particle Astrophysics and Cosmology. ZZ is supported in part by NASA grant 15-WFIRST15-0008, Cosmology with the High Latitude Survey Roman Science Investigation Team (SIT). JLT and RHW acknowledge support of NSF grant AST-1211889. JLT also acknowledges support of NSF grant AST-2009291. YYM is supported by NASA through the NASA Hubble Fellowship grant no. HST-HF2-51441.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555. AB is supported by the Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, the U.S. Department of Energy (DOE) Office of Science Distinguished Scientist Fellow Program. This research made use of computational resources at SLAC National Accelerator Laboratory, and the authors thank the SLAC computational team for support. This research used resources of the National Energy Research scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

**Software:** Python, Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011), SciPy (Jones et al. 2001–), George (Ambikasaran et al. 2015), Corrunc (Sinha & Garrison 2020), MultiNest (Feroz et al. 2009; Buchner et al. 2014)

## APPENDIX

### A. LUMINOSITY SELECTION ON GALAXY CLUSTERING

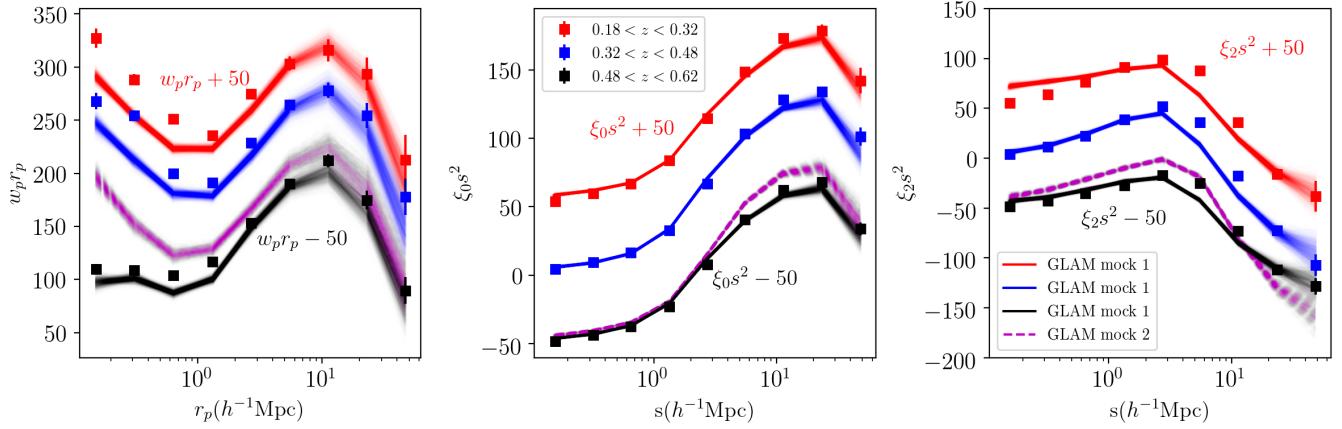
The selection of galaxies based on their luminosity results in a brighter subsample. In Figure 12, we show the 2PCF of these galaxies and in comparison with the original BOSS galaxies. Due to the correlation between galaxy brightness and the host halo mass, our subsample reveals an increased amplitude of correlation function. The impact is significant for  $w_p$  at all scales, and becomes weaker for RSD multipoles on small scales. This change requires the emulator of 2PCF from Zhai et al. (2019) to be re-trained to explain the measurements as described in the text.



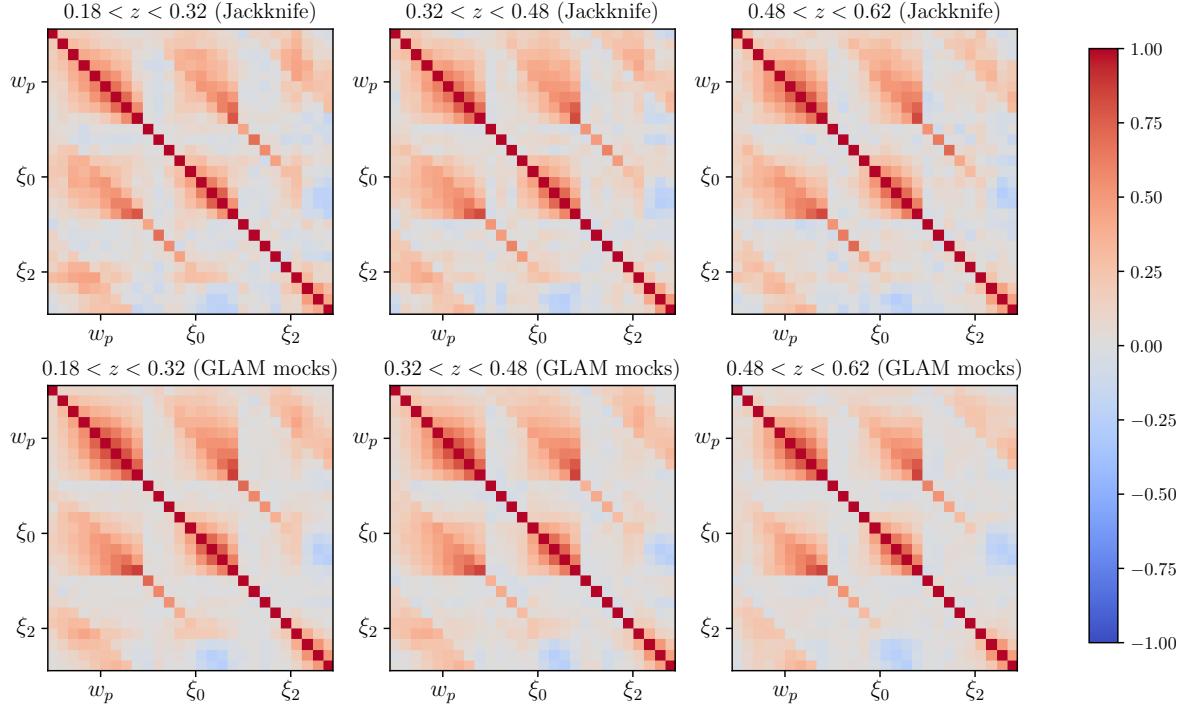
**Figure 12.** The two-point correlation function  $w_p$  (left),  $\xi_0$  (middle) and  $\xi_2$  (right) for the three subsamples of BOSS galaxies with and without luminosity selection. Dashed lines are measurements using all the galaxies, while the solid lines and squares denote our subsample selected by brightness. The low- $z$  and high- $z$  subsamples are shifted vertically for plotting purpose.

## B. CONSTRUCTION OF COVARIANCE MATRIX

The construction of the covariance matrix is described with details in Section 6.1. In this appendix, we present the measurement of the 2PCF and the resultant correlation matrix in Figure 13 and 14 respectively. The top panel of Figure 14 shows the correlation matrix of  $w_p$ ,  $\xi_0$  and  $\xi_2$  using jackknife resampling method of BOSS galaxies, and the bottom panel is from the GLAM mocks. For the high- $z$  subsample, we produce another set of GLAM mocks based on a random HOD. The correlation function is shown as the dashed purple line in Figure 13 with different clustering amplitude. However the resultant correlation matrix is similar to the first GLAM mocks (bottom panel of Figure 14). Based on these galaxy mocks, we construct the covariance matrix in the likelihood analysis as explained in the text, and test the impact of different choices of covariance matrix in Appendix G.



**Figure 13.** Correlation function of the GLAM mocks to construct the covariance matrix. The squares with error bars are measurements from BOSS galaxies. The red, blue, and black lines are obtained by populating the GLAM mocks with an optimized HOD parameter set and the GLAM cosmology. We do not require the mocks to give consistent results as observation, since the GLAM cosmology may not be the true cosmology. In order to further test the covariance matrix, we “randomly” pick a HOD model to populate the GLAM halo catalog for high- $z$  galaxies, as shown in the dashed purple lines. The clustering amplitude shows different behavior but the correlation matrix only shows mild variation.



**Figure 14.** **Top:** Correlation matrix for galaxy statistics  $w_p$ ,  $\xi_0$ , and  $\xi_2$  obtained by jackknife re-sampling of the BOSS galaxies. **Bottom:** The correlation matrix constructed using GLAM mocks.

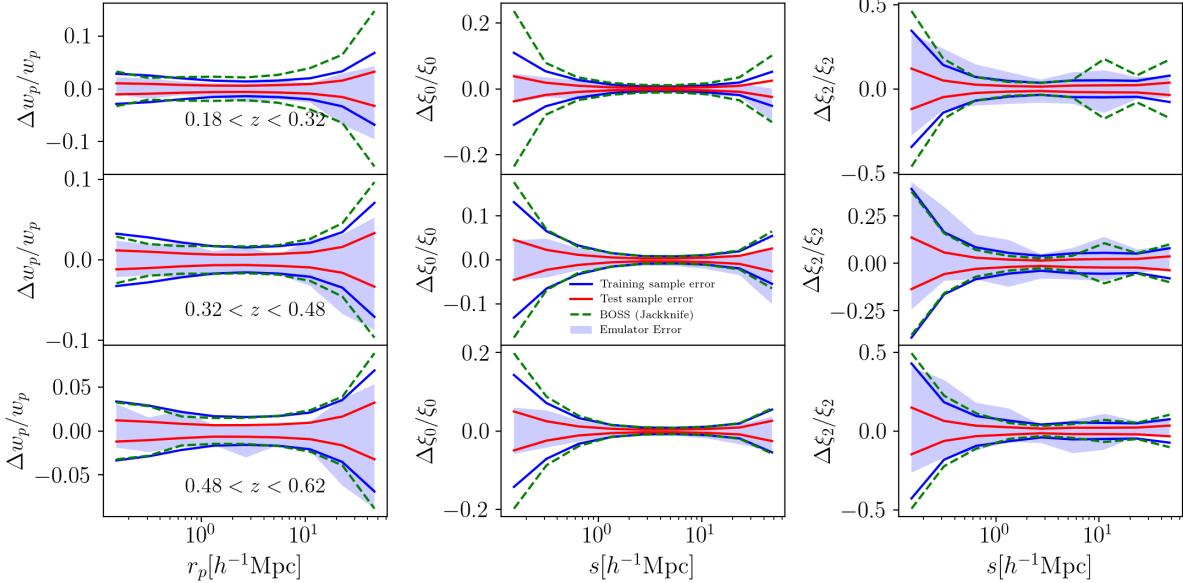
### C. CONSTRUCTION OF THE EMULATOR

The construction of the emulator for  $w_p$ ,  $\xi_0$  and  $\xi_2$  follows the same method as in Zhai et al. (2019), including the modeling of the training error, the choice of kernel function for the Gaussian-Process (GP) and the optimization method. In particular, we define the likelihood function with training sample as Eq (15) of Zhai et al. (2019) for each statistics. The hyper-parameters in the kernel which define the distance metric are for individual model parameters. The training process is to optimize the above likelihood function and the resultant hyper-parameters form the final emulator to make prediction for new model parameter set. Since the HOD parameter space is extended with additional parameters for assembly bias and widened ranges for  $M_{\text{sat}}$  and  $\alpha$ , the sampling of the emulator needs to be improved. Our test shows that using 80 HODs per cosmology is able to provide sufficient sampling for the emulator accuracy, which is a 60% increase compared with Zhai et al. (2019). The 2PCF of the galaxy mocks from both the training sample and test sample is computed by the publicly available code CORRFUNC (Sinha & Garrison 2020). Given the number density of our galaxy mocks, the average cost of one model evaluation using emulators for  $w_p + \xi_0 + \xi_2$  is at the level of 1 CPU-second. In Figure 15, we present the overall accuracy of the emulator for  $w_p$ ,  $\xi_0$  and  $\xi_2$  at three redshifts. The training error and test sample error are also shown for comparison. The emulator performance is quite similar as the pilot study in Zhai et al. (2019). For  $w_p$  and  $\xi_0$ , the emulator accuracy is close or better than sample variance. At scales from 1 to 10  $h^{-1}\text{Mpc}$ , the accuracy is better than 1 or 2%, which enables the use of the clustering measurement at the most informative scales. The emulator performance is worse for  $\xi_2$ , but it is still able to contain useful information for cosmological constraint.

### D. PRIOR FOR THE LIKELIHOOD ANALYSIS

The sampling in the cosmological parameter space is restricted within the area covered by training cosmologies. We incorporate this restriction by defining a prior for cosmological parameters. We first compute the mean  $\mu_{\text{cos}}$  and covariance matrix  $C_{\text{cos}}$  of the seven parameters using AEMULUS training cosmologies. Then for an arbitrary point  $\mu$  in the cosmological parameter space, we define a distance metric

$$\chi^2_{\text{cos}} = [\mu - \mu_{\text{cos}}] C_{\text{cos}}^{-1} [\mu - \mu_{\text{cos}}]. \quad (\text{D1})$$



**Figure 15.** Performance of the emulator for  $w_p$  (left),  $\xi_0$  (middle) and  $\xi_2$  (right) for three redshifts. The solid blue line stands for the training error, while the red line is the error of the test samples. We also show the uncertainty from BOSS measurement as green dashed lines for reference. The test samples have multiple boxes and population to suppress sample variance and shot noise. The shaded area represents the inner 68% distribution of the emulator performance.

We choose a threshold value for  $\chi_{\text{cos}}^2 < 12$  and only allows sampling that can satisfy this condition. This roughly corresponds to  $3\sigma$  level of this multivariate Gaussian distribution. The resulting seven-dimensional ellipsoid is able to enclose the training cosmologies. In Figure 16, we display the projection of the training cosmologies, as well as the prior space.

### E. TEST ON SHAM CATALOGS

We construct the SHAM catalog following the model proposed in Lehmann et al. (2017). This model adopts a velocity proxy to allow continuous transition between two halo properties

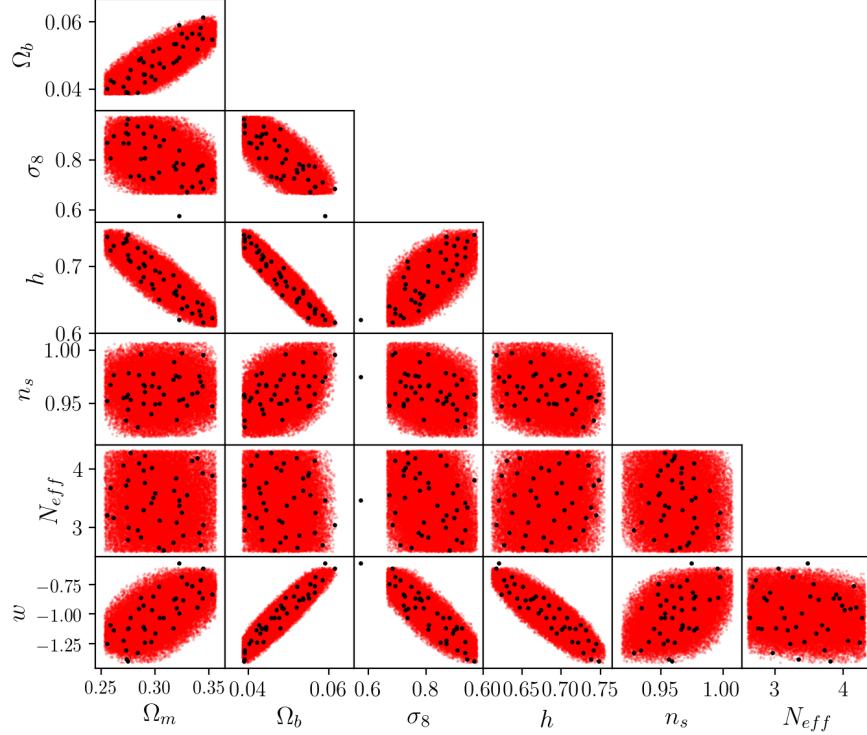
$$v_\alpha = v_{\text{vir}} \left( \frac{v_{\text{max}}}{v_{\text{vir}}} \right)^{\alpha_{\text{sham}}}, \quad (\text{E2})$$

where  $v_{\text{max}}$  is the maximal circular velocity and

$$v_{\text{vir}} = \left( \frac{GM_{\text{vir}}}{R_{\text{vir}}} \right)^{1/2}. \quad (\text{E3})$$

Note that we add subscript ‘‘sham’’ to distinguish  $\alpha_{\text{sham}}$  from the HOD parameter  $\alpha$ . When  $\alpha_{\text{sham}} = 0$ ,  $v_\alpha = v_{\text{vir}}$ , equivalent to matching galaxy by halo mass. When  $\alpha_{\text{sham}} = 1$ ,  $v_\alpha = v_{\text{max}}$ , the matching is based on maximal circular velocity. The typical value of  $\alpha_{\text{sham}}$  is restricted within  $[0, 1]$ , however we can artificially increase the value to increase the dependence of clustering on  $v_{\text{max}}$  and therefore boost the level of assembly bias.

In Figure 17, we present the measurements of  $w_p$ ,  $\xi_0$  and  $\xi_2$  for the SHAM catalogs using different values of scatter and  $\alpha_{\text{sham}}$ . The scatter parameter denotes the standard deviation of the stellar mass of galaxies at a fixed value of halo mass. We first perform the recovery test on the Uchuu SHAM catalog. The covariance matrix in the likelihood analysis is similar to the fiducial model as introduced in Section 6.1: the correlation matrix is from the GLAM mocks and the diagonal elements correspond to the sample variance of the AEMULUS training box. Since the Uchuu simulation has a box size of  $2h^{-1}\text{Gpc}$ , we also use the ratio of the volume to scale the diagonal elements to match the Uchuu volume. These two tests are distinguished as ( $1^{-1}\text{Gpc}$ ) and ( $2h^{-1}\text{Gpc}$ ) respectively. Since this volume factor only applies to the sample variance  $C_{\text{sam}}$ , while the emulator error remains the same, the difference in the finalized covariance matrix is diluted. Figure 18 displays the final constraints on the cosmological parameters and the key assembly bias parameter  $f_{\text{env}}$ . We can see that the input cosmology of the Uchuu simulation is well recovered within  $1\sigma$  using galaxy correlation function at non-linear scale. It shows that the assembly bias induced in the sub-halo matching process doesn’t bias the cosmological constraint, and supports the robustness of our emulator. The two tests with



**Figure 16.** Illustration of the prior space in the likelihood analysis for cosmological parameters. The black dots are the 40 training cosmologies from AEMULUS suite. Since Box023 has an outlier value of  $\sigma_8$ , we just use the other 39 models to get a covariance matrix of the cosmological parameters. Then we define a distance metric  $\chi^2$  and only allow the sampling of the points within a threshold. This defines a 7D ellipsoid with uniform and un-informative distribution. The red dots show the resulting sampling in the prior space, i.e., the ellipsoid imposed with the range of each parameters as shown in the first section of Table 3.

different diagonal elements of the covariance matrix ( $1^{-1}\text{Gpc}$ ) and ( $2h^{-1}\text{Gpc}$ ) show quite similar results, with the ( $2h^{-1}\text{Gpc}$ ) case presenting slightly tighter constraint on the contour plot. It is consistent with expectation that the larger volume can suppress the sample variance, but with a smaller difference due to the intrinsic emulator error since the volume scaling is only applied to the sample variance and the contribution from emulator error remains unchanged. This can make the (volume scaled) sample variance subdominant in the final covariance matrix.

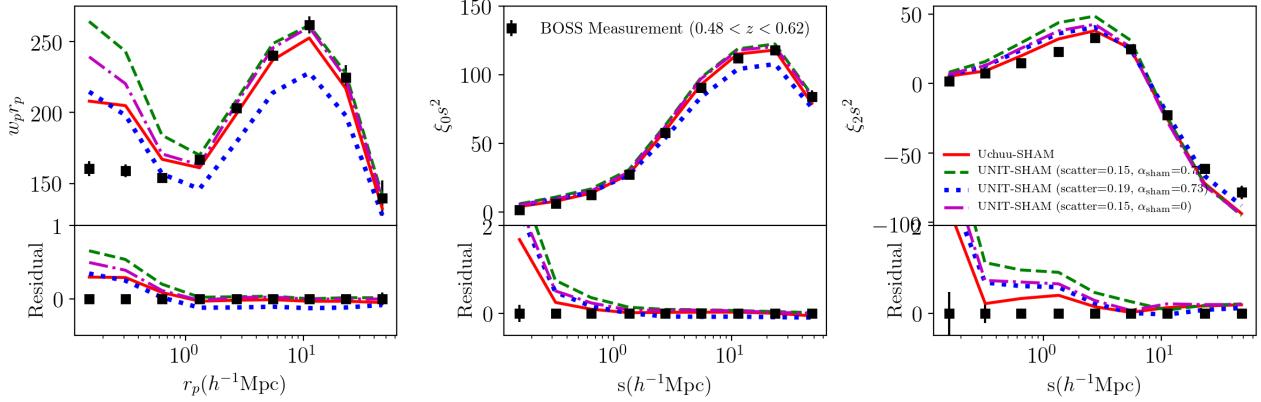
In Figure 19, we present the same recovery tests on the UNIT SHAM catalog. For simplicity, we assume the sample variance of the measurement of correlation function is equivalent to the training box of the AEMULUS suite. The actual “effective volume” of the UNIT boxes is non-trivial to estimate due to the inverse phase technique to reduce cosmic variance. However, the overall findings based on this recovery test is likely to hold. For different values of the scatter and  $\alpha_{\text{sham}}$ , our HOD-based model is also able to recover the input cosmology. The result is quite similar to Uchuu simulation.

#### F. CONSTRAINT ON ALL THE PARAMETERS

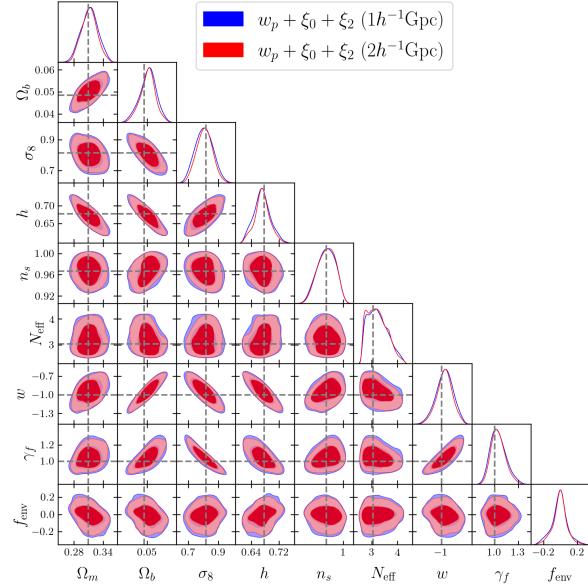
Figure 20 displays the full triangle plot for the constraint on all the parameters in our model using measurements of  $w_p + \xi_0 + \xi_2$ . The results of the three subsamples are shown on top of each other.

#### G. LIKELIHOOD TEST OF COVARIANCE MATRIX

In this section, we show the constraint using different covariance matrices as constructed in Section 6.1. We use the high- $z$  ( $0.48 < z < 0.62$ ) subsample to present our results, but the low- $z$  and med- $z$  subsamples give similar results. For the high- $z$  subsample, we remind the readers that we create another set of GLAM mocks using a “randomly” chosen HOD. Figure 21 presents the finalized constraint using  $w_p + \xi_0 + \xi_2$  for all these different covariance matrices. We find that the overall constraint is stable against difference choices of covariance matrix with some offset for a subset of the parameters. In order to explicitly investigate the impact on resultant measurement of structure growth rate, we extract  $f\sigma_8$  from these analysis and present the distribution in Figure 22. The results shows that different constructions of the covariance matrix give quite consistent measurement of linear growth rate of structure.



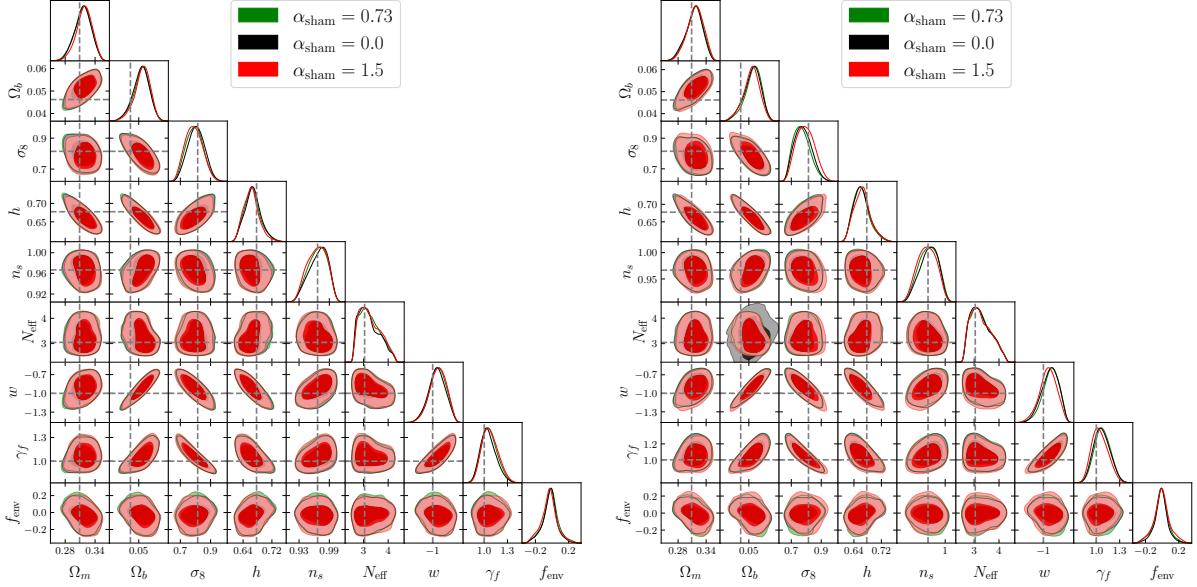
**Figure 17.** Galaxy 2PCF of the SHAM catalogs using Uchuu and UNIT simulations at  $z \sim 0.55$ , in comparison with high- $z$  BOSS galaxies.



**Figure 18.** Recovery test of the Uchuu SHAM catalog with  $\alpha_{\text{sham}} = 1.0$  and scatter=0.15. The constraint uses  $w_p + \xi_0 + \xi_2$  (right). The covariance matrix corresponds to the fiducial choice. Since the Uchuu simulation has larger volume than BOSS, we scale the uncertainties of the sample variance to match a  $1h^{-1}\text{Gpc}$  and  $2h^{-1}\text{Gpc}$  box respectively and compare the constraints. Our result shows that this difference is diluted to some extent by the contribution from the emulator error, however the sample variance corresponding to a larger box still gives a tighter constraint, consistent with expectation. The result shows that the input cosmology can be recovered within  $1\sigma$  for all the cosmological parameters. This demonstrates the robustness of our assembly bias augmented HOD model and the construction of the emulator.

## H. REDSHIFT UNCERTAINTY

One of the systematics in the clustering analysis is redshift uncertainty, or velocity dispersion, which can be estimated by repeat observations. The examination of BOSS galaxies shows that this uncertainty can be described by a Gaussian distribution with some value of standard deviation. The result based on both LOWZ and CMASS reveals a clear dependence of the dispersion on redshift, see Bolton et al. (2012) for a detailed analysis. From a theoretical point of view, this velocity dispersion can be captured through our HOD modeling by the parameters  $\gamma_f$  and velocity bias for central and satellites. In order to explicitly investigate the impact of this systematics on our cosmological measurements. We randomly pick 10 HOD models from our test suite, and add this velocity dispersion to the velocity of the galaxies in the mock and do the recovery test using the same set of emulator. In particular, we choose the high- $z$  subsample ( $0.48 < z < 0.62$ ) in this analysis, add an additional velocity component randomly draw from a Gaussian distribution with a standard deviation of 30 km/s corresponding to the maximum estimate in Bolton et al.



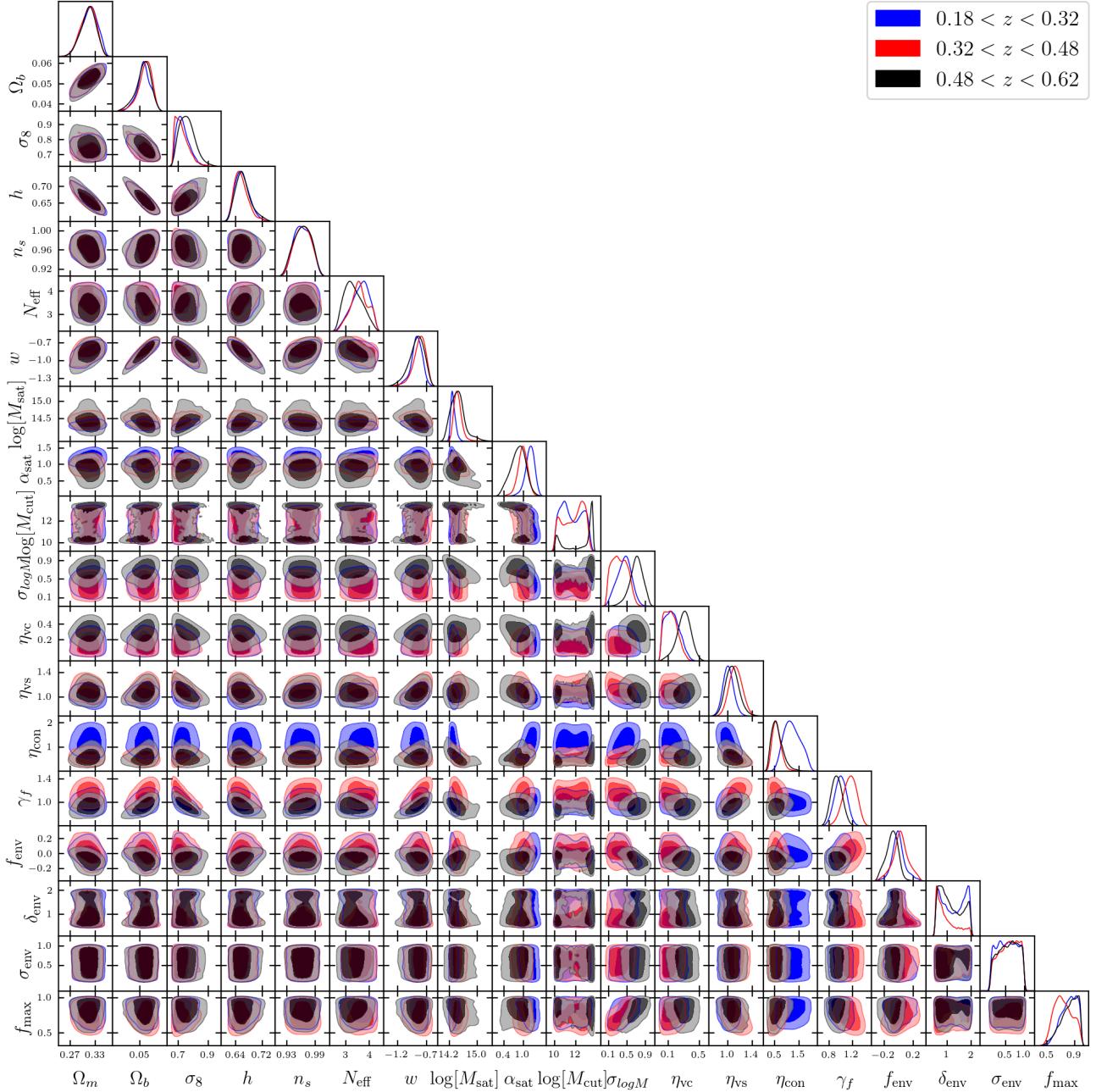
**Figure 19.** Test on the SHAM catalog constructed using UNIT simulation with  $\alpha_{\text{sham}} = 1$  and scatter=0.08. The galaxy statistics employed in the analysis is  $w_p + \xi_0 + \xi_2$  for high- $z$  subsample. The left hand panel shows SHAM catalog constructed with scatter= 0.15, while the right hand panel is for scatter= 0.19. The test is done on SHAM catalogs with different values of  $\alpha_{\text{sham}}$ , which is a measure of the significance of assembly bias. The result shows that our HOD based model can give unbiased recovery of the input cosmology.

(2012) within our redshift range. We perform the recovery tests using galaxy statistics  $w_p + \xi_0 + \xi_2$  for models with and without this additional velocity component, and present the inferred measurement of structure growth rate in Figure 23. The dots with error bars represent the  $1\sigma$  and  $2\sigma$  uncertainties of  $f\sigma_8$  recovered from the mocks compared with the input truth. This result shows that this velocity dispersion has a negligible impact on the final constraint. Since the velocity dispersion is increasing with redshift, the low- $z$  and med- $z$  subsamples can experience a less significant effect. Therefore this result validates the robustness of our measurement.

In Chapman et al. (2021), the velocity dispersion is examined for the eBOSS LRG sample. Since these high-redshift galaxies have an average velocity dispersion of  $91.8 \text{ km s}^{-1}$ , a factor of three higher than our high- $z$  BOSS galaxies, it can impact the cosmological constraint to a higher extent. For full scale measurement, the parameter  $\gamma_f$  can be biased with an offset of  $0.5\sigma$ . Although this component of systematics doesn't bias the BOSS measurements as in this paper, it imposes an additional concern for high-redshift galaxies in the near future, e.g., ELGs from eBOSS, DESI etc. Either a more flexible emulator or forward modeling is needed to achieve unbiased cosmological inference.

## REFERENCES

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, [ApJS, 182, 543](#)
- Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2022, [PhRvD, 105, 023520](#)
- Alam, S., Ata, M., Bailey, S., et al. 2017, [MNRAS, 470, 2617](#)
- Alcock, C., & Paczynski, B. 1979, [Nature, 281, 358](#)
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., & O'Neil, M. 2015, [IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, arXiv:1403.6015 \[math.NA\]](#)
- Angulo, R. E., & Pontzen, A. 2016, [MNRAS, 462, L1](#)
- Aricò, G., Angulo, R. E., Contreras, S., et al. 2020, arXiv e-prints, [arXiv:2011.15018](#)
- Asgari, M., Tröster, T., Heymans, C., et al. 2020, [A&A, 634, A127](#)
- Bautista, J. E., Paviot, R., Vargas Magaña, M., et al. 2021, [MNRAS, 500, 736](#)
- Beutler, F., Blake, C., Colless, M., et al. 2012, [MNRAS, 423, 3430](#)
- Bird, S., Rogers, K. K., Peiris, H. V., et al. 2019, [JCAP, 2019, 050](#)
- Blake, C., Brough, S., Colless, M., et al. 2012, [MNRAS, 425, 405](#)
- Blake, C., Baldry, I. K., Bland-Hawthorn, J., et al. 2013, [MNRAS, 436, 3089](#)
- Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, [AJ, 144, 144](#)
- Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, [A&A, 564, A125](#)



**Figure 20.** 1D and 2D contours of the parameters for our fiducial constraint using  $w_p + \xi_0 + \xi_2$ .

Chapman, M. J., Mohammad, F. G., Zhai, Z., et al. 2021, arXiv e-prints, arXiv:2106.14961

Chen, S.-F., Vlah, Z., & White, M. 2022, *JCAP*, 2022, 008

Chuang, C.-H., Yepes, G., Kitaura, F.-S., et al. 2019, *MNRAS*, 487, 48

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, *MNRAS*, 362, 505

Colless, M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, 328, 1039

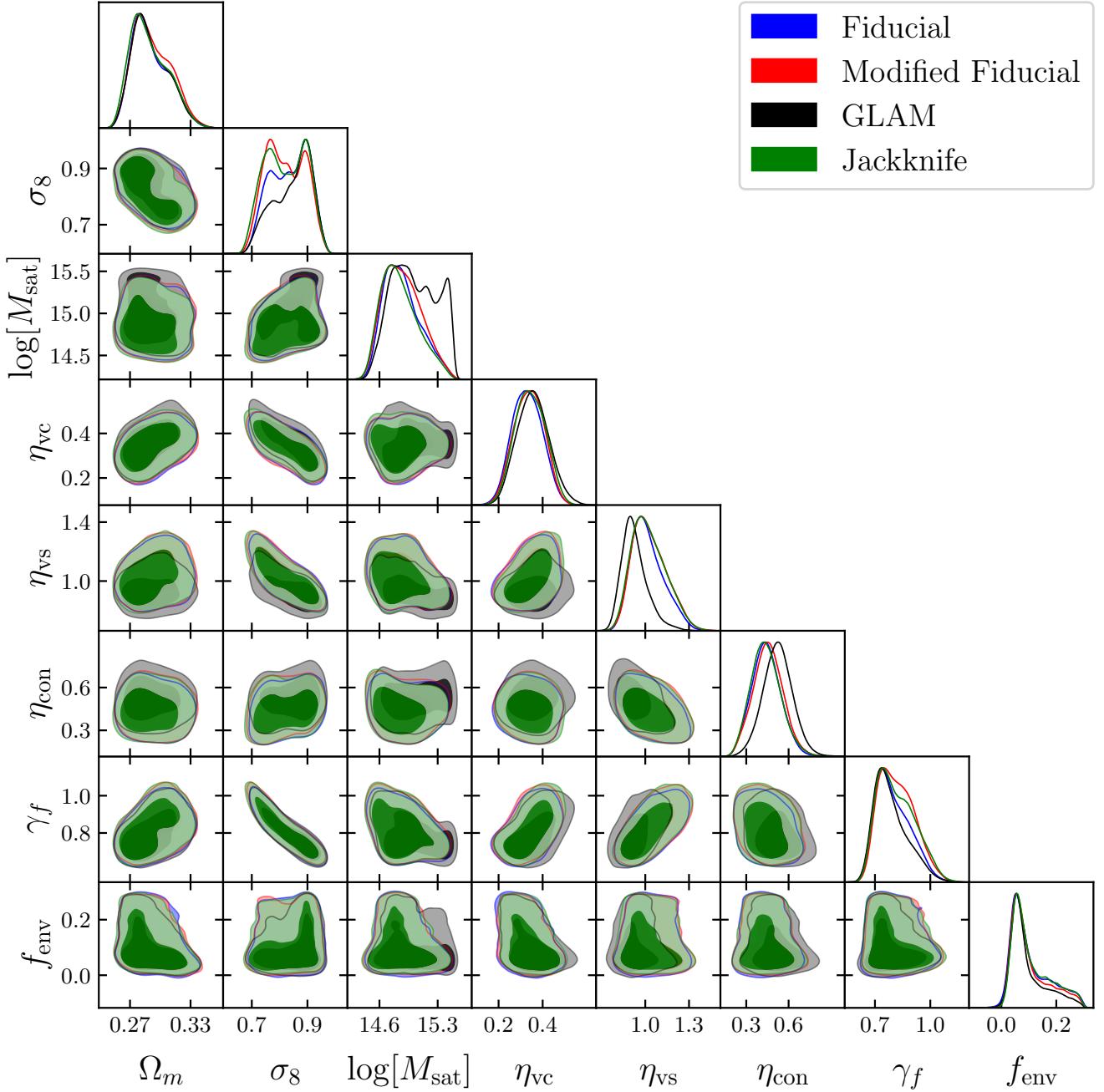
Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, *ApJ*, 647, 201

Contreras, S., Angulo, R., & Zennaro, M. 2020, arXiv e-prints, arXiv:2005.03672

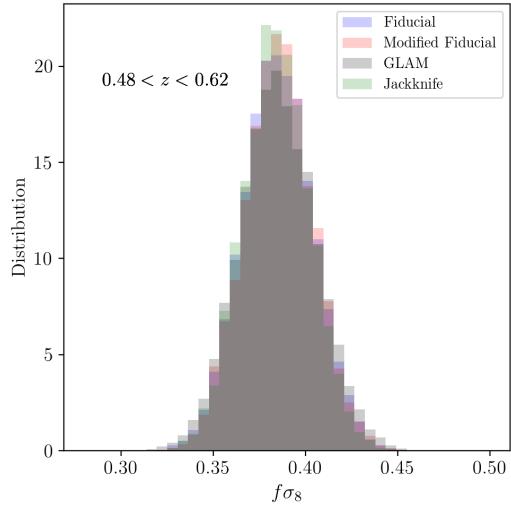
Davis, M., & Peebles, P. J. E. 1983, *ApJ*, 267, 465

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10

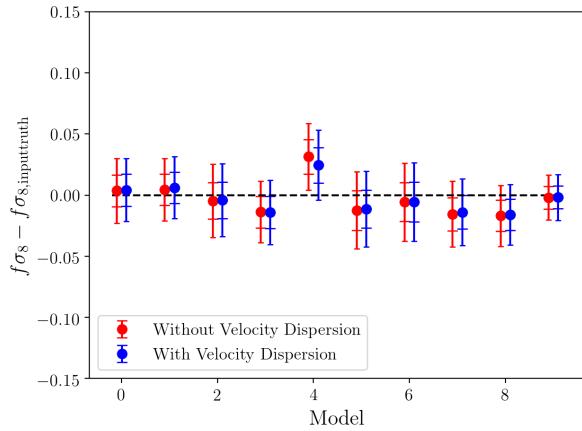
Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *AJ*, 151, 44



**Figure 21.** Test of the different constructions of covariance matrix. The result shows constraint on a subset of the cosmological, HOD and assembly bias parameters using  $w_p + \xi_0 + \xi_2$  for high- $z$  ( $z \sim 0.55$ ) subsample. The consistency of these results demonstrate that the construction of the covariance matrix for sample variance doesn't have a significant impact on the cosmological inference.



**Figure 22.** Projection of  $f\sigma_8$  using different constructions of the covariance matrix, labeled as Figure 21. The results shows that the different constructions of the covariance matrix give quite consistent measurement of  $f\sigma_8$ .



**Figure 23.** Impact of velocity dispersion on the measurement of structure growth rate. The tests are done with ten randomly chosen models. The result shows the difference of  $f\sigma_8$  recovered from the mock with respect to the input truth. The red dots with errorbars are result without additional velocity dispersion, while the blue dots are from mocks with additional velocity component. For plotting purpose, the points for the same model are shifted slightly along the x-axis and the errorbars represent 1 and 2  $\sigma$ .

- de Jong, R. S., Barden, S. C., Bellido-Tirado, O., et al. 2016, in Proc. SPIE, Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI, 990810
- de la Torre, S., Guzzo, L., Peacock, J. A., et al. 2013, *A&A*, 557, A54
- DeRose, J., Wechsler, R. H., Tinker, J. L., et al. 2019, *ApJ*, 875, 69
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, ArXiv e-prints, arXiv:1611.00036 [astro-ph.IM]
- Dressler, A., Spergel, D., Mountain, M., et al. 2012, arXiv e-prints, arXiv:1210.7809
- Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, *MNRAS*, 401, 1429
- Euclid Collaboration, Knabenhans, M., Stadel, J., et al. 2019, *MNRAS*, 484, 5509
- Feroz, F., Hobson, M. P., & Bridges, M. 2009, *MNRAS*, 398, 1601
- Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. 1994, *MNRAS*, 266, 50
- Gao, L., Springel, V., & White, S. D. M. 2005, *MNRAS*, 363, L66
- Giblin, B., Cataneo, M., Moews, B., & Heymans, C. 2019, *MNRAS*, 490, 4826
- Green, J., Schechter, P., Baltay, C., et al. 2012, arXiv e-prints, arXiv:1208.4012
- Guo, H., Zehavi, I., & Zheng, Z. 2012, *ApJ*, 756, 127
- Guo, H., Zheng, Z., Zehavi, I., et al. 2015, *MNRAS*, 446, 578
- Han, J., Li, Y., Jing, Y., et al. 2019, *MNRAS*, 482, 1900
- Harker, G., Cole, S., Helly, J., Frenk, C., & Jenkins, A. 2006, *MNRAS*, 367, 1039
- Hearin, A. P., Chaves-Montero, J., Becker, M. R., & Alarcon, A. 2021, *The Open Journal of Astrophysics*, 4, 7
- Hearin, A. P., Zentner, A. R., van den Bosch, F. C., Campbell, D., & Tollerud, E. 2016, *MNRAS*, 460, 2552
- Heitmann, K., Higdon, D., White, M., et al. 2009, *ApJ*, 705, 156
- Heitmann, K., Lawrence, E., Kwan, J., Habib, S., & Higdon, D. 2014, *ApJ*, 780, 111
- Heitmann, K., White, M., Wagner, C., Habib, S., & Higdon, D. 2010, *ApJ*, 715, 104
- Hoshino, H., Leauthaud, A., Lackner, C., et al. 2015, *MNRAS*, 452, 998
- Howlett, C., Ross, A. J., Samushia, L., Percival, W. J., & Manera, M. 2015, *MNRAS*, 449, 848
- Hunter, J. D. 2007, *Computing in Science Engineering*, 9, 90
- Ishiyama, T., Prada, F., Klypin, A. A., et al. 2020, arXiv e-prints, arXiv:2007.14720
- Ivanov, M. M. 2021, *PhRvD*, 104, 103514
- Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python, [Online; [scipy.org](http://scipy.org)]
- Klypin, A., & Prada, F. 2018, *MNRAS*, 478, 4602
- Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, *MNRAS*, 457, 4340
- Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, *ApJ*, 740, 102
- Kobayashi, Y., Nishimichi, T., Takada, M., & Miyatake, H. 2021, arXiv e-prints, arXiv:2110.06969
- Kobayashi, Y., Nishimichi, T., Takada, M., Takahashi, R., & Osato, K. 2020, *PhRvD*, 102, 063504
- Kokron, N., DeRose, J., Chen, S.-F., White, M., & Wechsler, R. H. 2021, *MNRAS*, 505, 1422
- Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, *ApJ*, 609, 35
- Krolewski, A., Ferraro, S., & White, M. 2021, *JCAP*, 2021, 028
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Lange, J. U., Hearin, A. P., Leauthaud, A., et al. 2021, arXiv e-prints, arXiv:2101.12261
- Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193
- Laureijs, R., Gondoin, P., Duvet, L., et al. 2012, in Proc. SPIE, Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, 84420T
- Lawrence, E., Heitmann, K., White, M., et al. 2010, *ApJ*, 713, 1322
- Leauthaud, A., Bundy, K., Saito, S., et al. 2016, *MNRAS*, 457, 4021
- Leauthaud, A., Saito, S., Hilbert, S., et al. 2017, *MNRAS*, 467, 3024
- Lehmann, B. V., Mao, Y.-Y., Becker, M. R., Skillman, S. W., & Wechsler, R. H. 2017, *ApJ*, 834, 37
- Lue, A., Scoccimarro, R., & Starkman, G. 2004, *Physical Review D*, 69
- Mao, Y.-Y., Williamson, M., & Wechsler, R. H. 2015, *ApJ*, 810, 21
- Mao, Y.-Y., Zentner, A. R., & Wechsler, R. H. 2018, *MNRAS*, 474, 5143
- McClintock, T., Rozo, E., Becker, M. R., et al. 2019a, *ApJ*, 872, 53
- McClintock, T., Rozo, E., Banerjee, A., et al. 2019b, arXiv e-prints, arXiv:1907.13167
- Miyatake, H., Kobayashi, Y., Takada, M., et al. 2020, arXiv e-prints, arXiv:2101.00113
- Nishimichi, T., Takada, M., Takahashi, R., et al. 2019, *ApJ*, 884, 29
- Pedersen, C., Font-Ribera, A., Rogers, K. K., et al. 2020, arXiv e-prints, arXiv:2011.15127
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, ArXiv e-prints, arXiv:1502.01589
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, *A&A*, 641, A6
- Ramachandra, N., Valogiannis, G., Ishak, M., & Heitmann, K. 2020, arXiv e-prints, arXiv:2010.00596
- Reid, B., Ho, S., Padmanabhan, N., et al. 2016, *MNRAS*, 455, 1553
- Reid, B. A., Seo, H.-J., Leauthaud, A., Tinker, J. L., & White, M. 2014, *MNRAS*, 444, 476
- Rogers, K. K., Peiris, H. V., Pontzen, A., et al. 2019, *JCAP*, 2019, 031

- Salcedo, A. N., Maller, A. H., Berlind, A. A., et al. 2018, *MNRAS*, **475**, 4411
- Salcedo, A. N., Zu, Y., Zhang, Y., et al. 2020, arXiv e-prints, arXiv:2010.04176
- Samushia, L., Reid, B. A., White, M., et al. 2014, *MNRAS*, **439**, 3504
- Sheth, R. K., & Tormen, G. 2004, *MNRAS*, **350**, 1385
- Shi, J., & Sheth, R. K. 2018, *MNRAS*, **473**, 2486
- Singh, S., Mandelbaum, R., Seljak, U., Rodríguez-Torres, S., & Slosar, A. 2020, *MNRAS*, **491**, 51
- Sinha, M., & Garrison, L. H. 2020, *MNRAS*, **491**, 3022
- Skilling, J. 2004, in *American Institute of Physics Conference Series*, Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395
- Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757
- Szewciw, A. O., Beltz-Mohrmann, G. D., Berlind, A. A., & Sinha, M. 2021, arXiv e-prints, arXiv:2110.03701
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, *PASJ*, **66**, R1
- Tinker, J. L., Conroy, C., Norberg, P., et al. 2008, *ApJ*, **686**, 53
- Tinker, J. L., Weinberg, D. H., Zheng, Z., & Zehavi, I. 2005, *ApJ*, **631**, 41
- Tinker, J. L., Sheldon, E. S., Wechsler, R. H., et al. 2012, *ApJ*, **745**, 16
- Tinker, J. L., Brownstein, J. R., Guo, H., et al. 2017, *ApJ*, **839**, 121
- Vakili, M., & Hahn, C. 2019, *ApJ*, **872**, 115
- Vale, A., & Ostriker, J. P. 2004, *MNRAS*, **353**, 189
- van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *Computing in Science Engineering*, **13**, 22
- Villarreal, A. S., Zentner, A. R., Mao, Y.-Y., et al. 2017, *MNRAS*, **472**, 1088
- Walsh, K., & Tinker, J. 2019, *MNRAS*, **488**, 470
- Walther, M., Armengaud, E., Ravoux, C., et al. 2020, arXiv e-prints, arXiv:2012.04008
- Wang, K., Mao, Y.-Y., Zentner, A. R., et al. 2019, *MNRAS*, **488**, 3541
- Wang, Y., Zhai, Z., Alavi, A., et al. 2021, arXiv e-prints, arXiv:2110.01829
- Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, *ApJ*, **568**, 52
- Wechsler, R. H., & Tinker, J. L. 2018, *ARA&A*, **56**, 435
- Wechsler, R. H., Zentner, A. R., Bullock, J. S., Kravtsov, A. V., & Allgood, B. 2006, *ApJ*, **652**, 71
- White, M., Zhou, R., DeRose, J., et al. 2022, *JCAP*, **2022**, 007
- Wibking, B. D., Weinberg, D. H., Salcedo, A. N., et al. 2020, *MNRAS*, **492**, 2872
- Wibking, B. D., Salcedo, A. N., Weinberg, D. H., et al. 2017, ArXiv e-prints, arXiv:1709.07099
- Xu, X., Zehavi, I., & Contreras, S. 2020, arXiv e-prints, arXiv:2007.05545
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, **120**, 1579
- Yuan, S., Garrison, L. H., Hadzhiyska, B., Bose, S., & Eisenstein, D. J. 2021, arXiv e-prints, arXiv:2110.11412
- Yuan, S., Hadzhiyska, B., Bose, S., Eisenstein, D. J., & Guo, H. 2020, arXiv e-prints, arXiv:2010.04182
- Zentner, A. R., Hearin, A., van den Bosch, F. C., Lange, J. U., & Villarreal, A. 2019, *MNRAS*, **485**, 1196
- Zhai, Z., Tinker, J. L., Hahn, C., et al. 2017, *ApJ*, **848**, 76
- Zhai, Z., Tinker, J. L., Becker, M. R., et al. 2019, *ApJ*, **874**, 95
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, *ApJ*, **633**, 791
- Zu, Y. 2020, arXiv e-prints, arXiv:2010.01143