

# Five per cent measurements of the growth rate from simulation-based modelling of redshift-space clustering in BOSS LOWZ

Johannes U. Lange<sup>1,2,★</sup>, Andrew P. Hearin<sup>3</sup>, Alexie Leauthaud<sup>1</sup>, Frank C. van den Bosch<sup>4</sup>, Hong Guo<sup>5</sup> and Joseph DeRose<sup>6,7</sup>

<sup>1</sup>Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA

<sup>2</sup>Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, Stanford, CA 94305, USA

<sup>3</sup>High Energy Physics Division, Argonne National Laboratory, Argonne, IL 60439, USA

<sup>4</sup>Department of Astronomy, Yale University, New Haven, CT 06511, USA

<sup>5</sup>Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, Shanghai 200030, China

<sup>6</sup>Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA

<sup>7</sup>Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, CA 95064, USA

Accepted 2021 October 20. Received 2021 September 29; in original form 2021 January 23

## ABSTRACT

We use a simulation-based modelling approach to analyse the anisotropic clustering of the BOSS LOWZ sample over the radial range  $0.4 h^{-1} \text{ Mpc}$  to  $63 h^{-1} \text{ Mpc}$ , significantly extending what is possible with a purely analytic modelling framework. Our full-scale analysis yields constraints on the growth of structure that are a factor of two more stringent than any other study on large scales at similar redshifts. We infer  $f\sigma_8 = 0.471 \pm 0.024$  at  $z \approx 0.25$ , and  $f\sigma_8 = 0.430 \pm 0.025$  at  $z \approx 0.40$ ; the corresponding  $\Lambda\text{CDM}$  predictions of the Planck cosmic microwave background (CMB) analysis are  $0.470 \pm 0.006$  and  $0.476 \pm 0.005$ , respectively. Our results are thus consistent with Planck, but also follow the trend seen in previous low-redshift measurements of  $f\sigma_8$  falling slightly below the  $\Lambda\text{CDM} + \text{CMB}$  prediction. We find that small- and large-radial scales yield mutually consistent values of  $f\sigma_8$ , but there are  $1\text{--}2.5\sigma$  hints of small scales ( $< 10 h^{-1} \text{ Mpc}$ ) preferring lower values for  $f\sigma_8$  relative to larger scales. We analyse the constraining power of the full range of radial scales, finding that most of the multipole information about  $f\sigma_8$  is contained in the scales  $2 h^{-1} \text{ Mpc} \lesssim s \lesssim 20 h^{-1} \text{ Mpc}$ . Evidently, once the cosmological information of the quasi-to-nonlinear regime has been harvested, large-scale modes contain only modest additional information about structure growth. Finally, we compare predictions for the galaxy–galaxy lensing amplitude of the two samples against measurements from SDSS and assess the lensing-is-low effect in light of our findings.

**Key words:** cosmological parameters – dark matter – large-scale structure of Universe.

## 1 INTRODUCTION

The standard cosmological model postulates that the evolution of the Universe can be described by the interaction of (dark) matter, radiation, and dark energy in the context of general relativity. In particular, the simplest model that can accurately describe most if not all observations today is the  $\Lambda$  cold dark matter ( $\Lambda\text{CDM}$ ) cosmological model. Although the standard model has been remarkably successful in explaining numerous observational data sets in isolation, the growing diversity of high-precision cosmological measurements offers an opportunity to conduct stringent self-consistency tests of  $\Lambda\text{CDM}$ .

A very powerful combination of observations are the cosmic microwave background (CMB) coupled with measurements of the low-redshift Universe, as these can be used to test  $\Lambda\text{CDM}$  predictions for the state of the Universe at very different times of its evolution. Observations of the CMB alone place tight constraints on  $\Lambda\text{CDM}$  parameters, enabling precise, testable predictions for the late-time

evolution of the Universe. We can generically divide these predictions into two categories: the expansion history of the Universe, and the growth of large-scale structure. Much attention has been paid to the apparent tension between the late-time expansion rate inferred from the CMB in comparison to direct observations (see e.g. Verde, Treu & Riess 2019, and references therein). Similarly, there is a mounting evidence that growth of structure predictions from the CMB do not match what is seen in observations (Abbott et al. 2018; Hikage et al. 2019; Asgari et al. 2021). Confirming and quantifying the level of disagreement in both aspects of the prediction is crucial to narrowing down theoretical explanations for  $\Lambda\text{CDM}$  tensions (Blinov et al. 2019; Keeley et al. 2019; Vattis, Koushiappas & Loeb 2019; Di Valentino, Melchiorri & Silk 2020; Hill et al. 2020; Ivanov et al. 2020; Kreisch, Cyr-Racine & Doré 2020; D’Amico et al. 2021; Handley 2021; Muir et al. 2021).

The observed clustering properties of galaxies reflect the underlying cosmological matter field in which the galaxies evolve. The cosmological constraining power of galaxy clustering becomes especially stringent when spectroscopic redshift information is available, as redshift-space distortions (RSDs) can be used to probe the matter velocity field in addition to the density field. The primary

\* E-mail: [julange@stanford.edu](mailto:julange@stanford.edu)

constraint obtained from the analysis of redshift-space clustering is the cosmological parameter combination  $f\sigma_8$ , where  $f$  is the growth rate of structure obtained from linear perturbation theory, and  $\sigma_8$  quantifies the normalization of the matter power spectrum. The advent of large-scale galaxy surveys has opened up the road for such constraints, including from the the WiggleZ survey (Blake et al. 2011), the *Sloan Digital Sky Survey* (Beutler et al. 2012; Samushia, Percival & Raccanelli 2012), the *Galaxy and Mass Assembly* survey (Blake et al. 2013), the *Baryon Oscillation Spectroscopic Survey* (BOSS; see e.g. Parejko et al. 2013; Alam et al. 2017), and the *Extended Baryon Oscillation Spectroscopic Survey* (eBOSS; see e.g. Bautista et al. 2021; de Mattia et al. 2021; Hou et al. 2021).

Theoretical predictions for the redshift-space clustering of galaxies are notoriously challenging. To leading order in perturbation theory, the RSD signal is determined by the Kaiser effect (Sargent & Turner 1977; Kaiser 1987), i.e. the coherent infall of galaxies towards over-densities. However, even on very large scales, the precision of present-day galaxy surveys requires the RSD prediction to include a correction that accounts for the virial motion of satellites inside haloes (Peacock & Dodds 1994; Scoccimarro 2004), the so-called *Fingers of God* effect, a fundamentally nonlinear phenomenon that cannot be predicted from perturbation theory.

It is well known that the constraining power of RSD analyses improves dramatically when including information from the quasi-to-nonlinear regime (e.g. Zhai et al. 2019); the potential gains are so substantial that there now exists a substantial literature dedicated to the development of models that can extend the range of scales used in the analysis. This long-standing effort includes various extensions of perturbation theory (Taruya, Nishimichi & Saito 2010; Carlson, Reid & White 2013; Matsubara 2014; Wang, Reid & White 2014), the Gaussian streaming model (Reid & White 2011), the Zel'dovich streaming model (White 2014), Effective Field Theory (Lewandowski, Perko & Senatore 2015), and approaches based on distribution functions (Seljak & McDonald 2011; Okumura, Seljak & Desjacques 2012). Whilst this effort to extend linear theory has certainly improved the predictive power of contemporary analytical frameworks, the reliability of such efforts currently remains limited to scales larger than  $\sim 30 h^{-1}$  Mpc (see White et al. 2015, for a review).

In a promising pilot study, Reid et al. (2014) highlighted the potential constraining power of extending the analysis of RSDs to smaller scales by performing an analysis on the BOSS CMASS sample. In the absence of reliable analytic models for redshift-space clustering on all scales, the authors used a single simulation with a re-scaling of the bulk velocity of haloes to constrain the growth rate. Reid et al. (2014) presented a 2.5 per cent constraint on  $f\sigma_8$ , a more than factor of two improvement over other studies using larger scales and on par with constraints on  $S_8 = \sigma_8 \sqrt{\Omega_m}/0.3$ , the equivalent cosmological quantity probed by large-scale structure probes involving gravitational lensing. Recently, Zhai et al. (2019) made another significant step towards accurate constraints from non-linear scales by constructing a high-precision Gaussian process (GP) emulator (also see Kwan et al. 2015; Nishimichi et al. 2019) to predict galaxy redshift-space clustering for BOSS CMASS. Whilst Zhai et al. (2019) argue that the re-scaling method of Reid et al. (2014) is subject to systematic errors and yields cosmological posteriors that are too narrow, they similarly show that including data from highly non-linear scales has the potential to significantly tighten constraints on the growth of structure.

On sufficiently large scales, the statistical relation between observed galaxies and the underlying density field can be described in terms of a perturbative bias expansion, with uncertainty in

this relationship encoded by the coefficients of the expansion (see Desjacques, Jeong & Schmidt 2018, for a comprehensive review). Due to the non-linear physics of gravitational collapse, this formalism breaks down irreparably on sufficiently small scales, and one needs to directly model the relation between galaxies and the dark matter haloes that host them, colloquially known as the ‘galaxy–halo connection’. There are two commonly used frameworks to model this relationship. In subhalo abundance matching (SHAM; see e.g. Conroy, Wechsler & Kravtsov 2006; Lehmann et al. 2017), one assumes a correspondence between all dark matter haloes, both field haloes and subhaloes accreted by other haloes, and galaxies. In contrast, in the halo occupation distribution (HOD; see e.g. Seljak 2000; Berlind & Weinberg 2002; Zheng, Coil & Zehavi 2007; Hearin et al. 2016; Sinha et al. 2018) framework, or equivalently, the conditional luminosity function (see e.g. Yang, Mo & van den Bosch 2003; van den Bosch et al. 2007) framework, one only models the relation between galaxies and field haloes but now allows field haloes to host multiple galaxies.

The goal of this paper is to perform the first consistent simulation-based cosmological RSD analysis of the full range of scales accessible with current large-scale structure data sets. Particularly, for this work, we analyse anisotropic redshift-space clustering in the scale range from  $0.4 h^{-1}$  Mpc to  $63 h^{-1}$  Mpc. Our analysis is based on the cosmological evidence modelling (CEM) framework introduced in Lange et al. (2019c); in comparison to the widely used GP emulation technique (e.g. Heitmann et al. 2010), the CEM provides improved flexibility to robustly quantify Bayesian posterior uncertainties with more realistically complex models of the galaxy–halo connection. We test our method using mock data sets created by populating dark matter haloes in numerical simulations with galaxies using the SHAM method, which we subsequently analyse using our HOD model applied to a different set of simulations. We demonstrate that our method is able to yield tight, unbiased cosmological constraints, which is a non-trivial result given that the SHAM model is substantially different from the HOD model. Having demonstrated that the HOD model is unbiased and sufficiently general, we next apply it to two carefully constructed galaxy samples from the BOSS LOWZ survey. After marginalizing over uncertainties in the galaxy–halo connection, we obtain two independent measurements of  $f\sigma_8$ , each with an accuracy of  $\sim 5$  per cent, a factor of two improvement over previous results. Finally, we quantify how our cosmological constraints depend on both small and large scales and explore predictions for gravitational lensing.

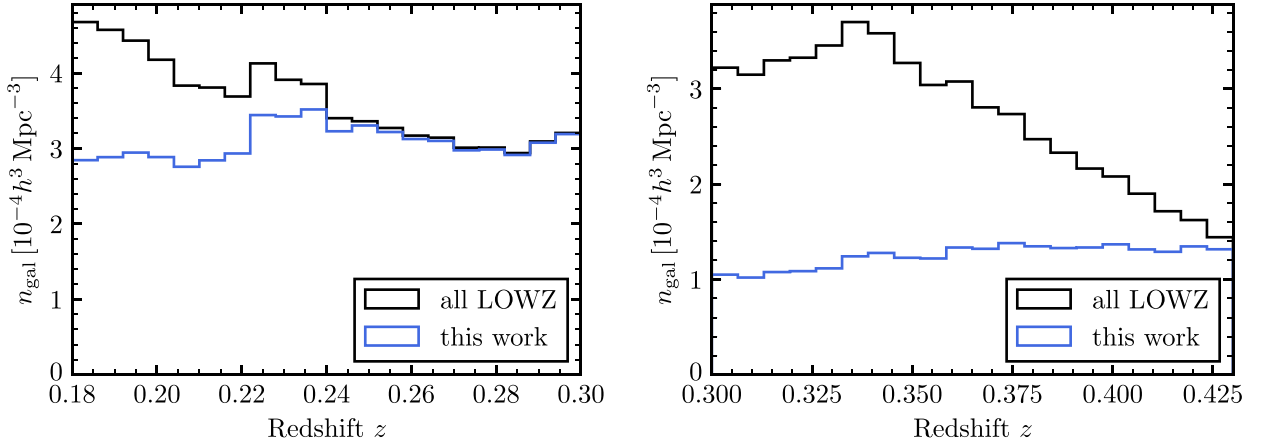
This paper is organized as follows. In Section 2, we describe our observational data set from the BOSS LOWZ survey and the summary statistics we extract from it. We describe our modelling framework, including the HOD model and simulations used, in Section 3. Section 4 shows that our modelling framework is able to recover accurate cosmological constraints from mock measurements. Our cosmological constraints from applying the analysis framework to the observational data are described in Section 5. Finally, we discuss our results in Section 6 and list our conclusions in Section 7.

Throughout this work, all observational measurements are made assuming a spatially flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.307$ .

## 2 OBSERVATIONS

### 2.1 Sample selection

We select galaxies from the BOSS LOWZ catalogue. Spectroscopic target selection for LOWZ requires that objects fulfil the following selection cuts:



**Figure 1.** The comoving galaxy number density of LOWZ galaxies in the redshift range  $0.18 \leq z < 0.3$  (left) and  $0.3 \leq z < 0.43$  (right). We show the distribution of the LOWZ parent sample (black) and our subset (blue).

$$r_{\text{cmod}} < 13.5 + c_{\parallel} / 0.3 \quad (1)$$

$$|c_{\perp}| < 0.2 \quad (2)$$

$$16 < r_{\text{cmod}} < 19.6. \quad (3)$$

In the above equations,  $r_{\text{cmod}}$  refers to the observed cmodel magnitudes and  $c_{\parallel}$  and  $c_{\perp}$  are colours defined as follows:

$$c_{\parallel} = 0.7(g_{\text{mod}} - r_{\text{mod}}) + 1.2(r_{\text{mod}} - i_{\text{mod}} - 0.18) \quad (4)$$

and

$$c_{\perp} = r_{\text{mod}} - i_{\text{mod}} - (g_{\text{mod}} - r_{\text{mod}}) / 4.0 - 0.18. \quad (5)$$

Note that these colours are measured using model magnitudes. We refer the reader to Eisenstein et al. (2001) for a detailed motivation of these selection cuts and outline only the most salient points here. First, passively evolving galaxies form a locus in the  $g - r$  versus  $r - i$  plane. Equation (2) uses the position in the  $g - r$  versus  $r - i$  plane to select galaxies at  $z \lesssim 0.45$ . Additionally, equation (1) results in a roughly redshift-independent cut on absolute magnitude. Finally, equation (3) ensures that galaxies too faint to be targeted by the SDSS spectrograph are not selected. Ultimately, the above criteria for targets result in a luminous red galaxy (LRG) sample in the redshift range  $z \sim 0.15 - 0.50$ . We seek to select from this parent sample two approximately volume-limited samples of LRGs with median redshifts of  $z \sim 0.25$  and  $z \sim 0.40$ , corresponding to the two Aemulus simulation outputs in the BOSS LOWZ redshift range.

We note that the above selection is based on apparent magnitudes. It is thus not expected that the target selection would be perfectly uniform in redshift. However, such a non-uniform sample selection violates the assumption of a redshift-independent HOD model that we will assume in the modelling in Section 5. To create volume-limited samples of red galaxies, we have to impose further cuts.

We start by calculating  $k$ -corrected, absolute magnitudes at  $z_0 = 0.25$  and  $z_0 = 0.4$  for all BOSS LOWZ galaxies. The  $k$ -corrections are calculated from the model magnitudes using KCORRECT and, in the following, all absolute magnitudes are expressed in the AB magnitude system. Next, we select galaxies in the redshift ranges  $(0.18, 0.30]$  and  $(0.30, 0.43]$ . Finally, we impose more stringent cuts on the absolute magnitudes ( $M_{r,\text{cmod},z_0}$ ) and colours ( $c_{\parallel,z_0}$  and  $c_{\perp,z_0}$ ):

$$M_{r,\text{cmod},0.25} < -25.874 + c_{\parallel,0.25} / 0.3 \quad (6)$$

$$-0.216 < c_{\perp,0.25} < 0.162 \quad (7)$$

$$M_{r,\text{cmod},0.25} < -20.412 \quad (8)$$

and

$$M_{r,\text{cmod},0.40} < -26.899 + c_{\parallel,0.40} / 0.3 \quad (9)$$

$$-0.154 < c_{\perp,0.40} < 0.112 \quad (10)$$

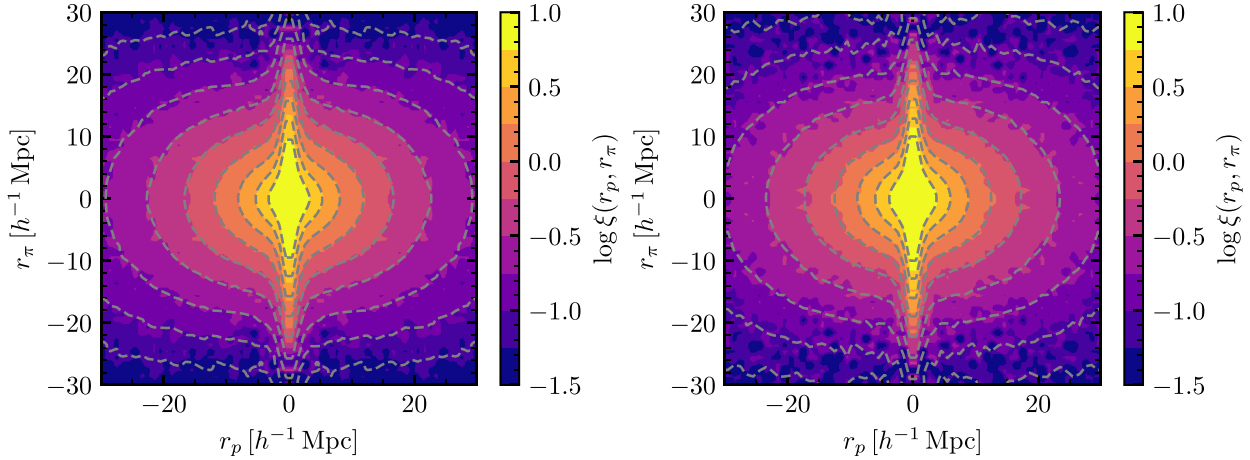
$$M_{r,\text{cmod},0.40} < -21.209 \quad (11)$$

for the  $0.18 < z \leq 0.30$  and  $0.30 < z \leq 0.43$  samples, respectively. These cuts were calculated based on the average  $k$ -corrections as a function of redshift for the respective samples. In the absence of scatter in the  $k$ -corrections, all galaxies fulfilling these more stringent cuts also fulfil the general LOWZ target selection cuts.<sup>1</sup>

Because these more stringent cuts are based on rest-frame properties, this should result in roughly volume-limited samples of red galaxies. In Fig. 1, we show the impact of our selection cuts on the comoving galaxy number density in both redshift ranges considered. Especially for the second sample, our selection cuts reduce the total number of galaxies in this range by a factor of  $\sim 2$ . The main reason for the reduction is the cut in absolute  $r$ -band magnitude. The original BOSS LOWZ selection effectively applied a very high cut on absolute magnitude at  $z = 0.43$  and applying this across the redshift range reduces the number of galaxies at  $z = 0.30$ . At the same time, we see that our more stringent selection results in a roughly redshift-independent galaxy number density, as expected for a volume-limited sample without intrinsic number density evolution.

As described in Ross et al. (2017), there are slight differences in the photometric calibration of samples in the Northern Galactic Cap (NGC) and the Southern Galactic Cap. This results in slightly different target selections between the two hemispheres. Therefore, in this paper, we only select galaxies from the NGC. Overall, we have 79 018 galaxies with a median redshift of 0.251 for the first sample and 72 056 galaxies with a median redshift of 0.377 for the second sample. The corresponding comoving number densities are  $(2.80 \pm$

<sup>1</sup>We ignored that the LOWZ target selection requires  $r_{\text{cmod}} > 16$ , thereby excluding objects that are too bright. However, there are only few physical objects for which this is true.



**Figure 2.** The anisotropic two-point correlation function for the  $0.18 \leq z < 0.3$  sample (left) and the  $0.3 \leq z < 0.43$  sample (right). Colours indicate the measurements and dashed lines are predictions of the best-fitting models discussed in Section 5 of where the colour transitions occur.

$0.05) \times 10^{-4} h^3 \text{ Mpc}^{-3}$  and  $(1.18 \pm 0.02) \times 10^{-4} h^3 \text{ Mpc}^{-3}$ , respectively. The redshift ranges,  $0.18 < z \leq 0.30$  and  $0.30 < z \leq 0.43$ , are narrow and correspond to time ranges of 1.2 and 1.1 Gyr, respectively. Thus, we do not expect strong redshift evolution effects within each redshift bin.

## 2.2 Clustering measurements

Following Guo et al. (2015b), we measure the 3D redshift-space two-point correlation function  $\xi(s, \mu)$  through the Landy–Szalay estimator (Landy & Szalay 1993), where  $s$  is the redshift-space separation of the galaxy pairs and  $\mu$  is the cosine of the angle between  $s$  and the line of sight. We choose logarithmic  $s$  bins with a width  $\Delta \log s = 0.2$  from 0.1 to  $63.1 h^{-1} \text{ Mpc}$ , and linear  $\mu$  bins of width  $\Delta \mu = 0.05$  from  $-1$  to  $1$ .

The multipole moments of order  $\ell$  are then defined via

$$\xi_\ell(s) = \frac{2\ell + 1}{2} \int_{-1}^1 L_\ell(\mu) \xi(s, \mu) d\mu, \quad (12)$$

(Hamilton 1992). In the above equation,  $L_\ell$  represents the Legendre polynomial of order  $\ell$ . Note that contrary to Reid et al. (2014), we integrate  $L_\ell$  over the entire  $\mu$ -range. Measuring  $\xi_\ell(s)$  on small scales  $s$  is complicated by the presence of so-called fibre collisions, the fact that two spectroscopic fibres cannot be placed close to each other on a BOSS spectroscopic plate.

To accurately measure  $\xi(s, \mu)$ , we correct for the fibre collision effect of the BOSS LOWZ sample using the method of Guo, Zehavi & Zheng (2012). After being updated with the latest SDSS Data Release 16 (Ahumada et al. 2020), the overall fraction of fibre collided galaxies is just 1.4 per cent for galaxies in the NGC of the BOSS LOWZ sample. The small-scale measurements can then be very accurately recovered, with minor fibre collision corrections.

The anisotropic two-point correlation functions are shown in Fig. 2 for visualization purposes. However, throughout this work, our modelling is performed using the multipole moments.

## 2.3 Covariance matrix

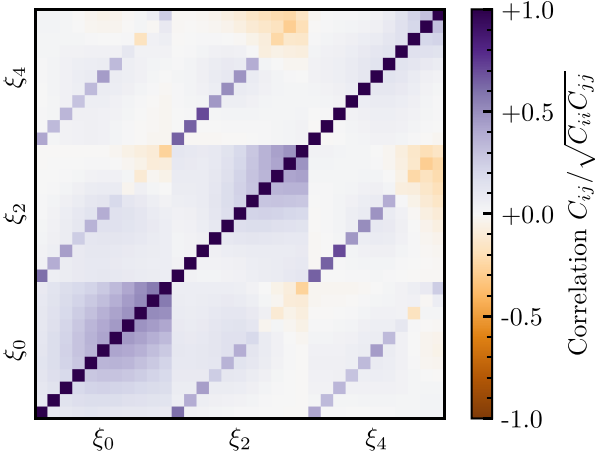
Uncertainties on the measurements are derived from jackknife-resampling of 74 roughly equal-area regions of the NGC. Note that

this number of jackknife samples is comparable to the number of measurement points. Thus, without corrections, the inverse of the covariance matrix is expected to be a biased estimate of the precision matrix (Hartlap, Simon & Schneider 2007; Taylor, Joachimi & Kitching 2013). In principle, an unbiased estimate of the precision matrix can be obtained by applying the so-called Hartlap factor (Hartlap et al. 2007). However, we opt for a more physically motivated correction to the covariance matrix estimator.

Let us look at the covariance matrix of two multipole moments  $\xi_i$  and  $\xi_k$ . We expect the following statements to hold regardless of whether we study the covariance within the same multipole moment,  $i = k$ , or different moments,  $i \neq k$ . First, it is expected that the multipole moments of the same bin in  $s$  are heavily correlated because, as shown in equation (12), they are derived from the same  $\xi$  measurements at fixed  $s$  with different  $\mu$  weights given by the Legendre polynomials. These correlations correspond to the diagonal elements of the covariance matrix and we choose not to modify those. Similarly, multipole moments should be less correlated if they do not belong to the same bin in  $s$ , i.e. off-diagonal elements, since they are derived from distinct measurements of  $\xi(s, \mu)$ . Additionally, it is expected that neighbouring off-diagonal elements of the covariance matrix that are close in scale show similar levels of correlation. Thus, we smooth the off-diagonal elements of the correlation matrix with a 2D Gaussian kernel.<sup>2</sup> The resulting correlation matrix is shown in Fig. 3 for the  $z = 0.25$  sample. We choose a width of 2 for the Gaussian kernel but the exact value does not significantly affect our results. Additionally, as described later, we have also performed the entire analysis on the data using the standard Hartlap correction and find very similar results compared to the smoothing procedure. We note that the approach employed here has been used in a very similar fashion in Mandelbaum et al. (2013), who also find that such a smoothing of the correlation matrix gives virtually the same results as applying the Hartlap factor to the noisy precision matrix.

When fitting models to the data we exclude scales below  $400 h^{-1} \text{ kpc}$ . The signal to noise of the data on these small scales is poor, especially for the hexadecapole moments, and since the jackknife sub-samples have very few or no galaxy pairs at these

<sup>2</sup>Before applying the smoothing, we replace each diagonal element with the average from the four neighbouring, off-diagonal elements.



**Figure 3.** The assumed covariance matrix for the redshift-space clustering measurements. The scale  $s$  increases from left to right and bottom to top. We see that multipole moments within the same  $s$ -bin are strongly correlated. In contrast, measurements at different scales  $s$  only show significant correlation at larger scales. The covariance matrix here is for the sample at  $z \sim 0.25$ . The covariance matrix for the higher-redshift sample,  $z \sim 0.4$ , is qualitatively similar. See the text for details regarding the construction of the covariance matrix. We ignore cross-correlation between the clustering measurements and the galaxy number density.

separations, the corresponding estimates of the precision matrix are unreliable.

### 3 THEORETICAL MODELLING

Our modelling is based on comparisons of the observational data described in the previous sections with predictions from simulations. In the following, we describe the simulations used and how we construct mock galaxy samples from them.

#### 3.1 Simulations

Our modelling is based on the publicly available Aemulus simulations (DeRose et al. 2019). Aemulus is a suite of 75 cosmological, dark matter-only simulations with a volume of  $(1050 h^{-1} \text{ Mpc})^3$  and  $1400^3$  particles resulting in a particle mass resolution of  $m_p = 3.51 \times 10^{10} \sqrt{\Omega_{m,0}/0.3} M_\odot/h$ . Our analysis is based on the first 40 Aemulus simulations that all have different cosmologies. The cosmological parameters of all simulations are listed in the appendix and probe the  $4\sigma$  posterior ranges of the combined Baryon Acoustic Oscillation plus CMB analysis by Anderson et al. (2014).

The Aemulus simulation suite is specifically designed for the study of non-linear clustering in BOSS CMASS. The BOSS LOWZ and CMASS samples target luminous red galaxies of similar number densities and redshifts. Thus, the convergence analysis presented in DeRose et al. (2019) should also be applicable to this work. Generally, particle resolution is the main limiting factor regarding the convergence of the Aemulus simulations. At fixed CMASS-like HOD, galaxy redshift-space correlation functions are converged to within  $\sim 2.5$  per cent. This is better than the accuracy of our measurements except for  $\xi_0$  over the range  $1 h^{-1} \text{ Mpc} < s < 20 h^{-1} \text{ Mpc}$ . Given that the analysis of DeRose et al. (2019) is performed at fixed HOD, resolution effects could bias our galaxy–halo connection parameters. At the same time, it is not clear that small changes to the halo mass function or mass–concentration relation due to resolution effects

would bias  $f\sigma_8$  constraints significantly since they could easily be degenerate with galaxy–halo connection parameters. Indeed, in Section 4, we perform mock tests on galaxy mock catalogues derived from the UNIT simulations which have  $\sim 25$  times better mass resolution than the Aemulus simulations. Since we do not find biases in the inferred  $f\sigma_8$  values from these mock tests, this serves as indication that the resolution of the Aemulus simulations is sufficient for our application regarding obtaining cosmological parameter constraints.

#### 3.2 Halo catalogues

Dark matter haloes in the simulations were identified with the ROCKSTAR phase-space halo finder (Behroozi, Wechsler & Wu 2013). From the halo catalogue, we use field haloes with a mass of at least  $100 m_p$ . We will use halo catalogues extracted at  $z = 0.25$  and  $z = 0.40$ , as appropriate. Halo masses  $M$  and radii  $r_h$  are defined using an over-density threshold of  $\Delta = 200$  times the background density of the Universe at the redshift of the halo catalogues. Halo concentrations  $c$  are derived from the maximum circular velocity  $V_{\text{max}}$  values provided in the catalogues assuming a spherically symmetric Navarro-Frenk-White (NFW) profile (Navarro, Frenk & White 1996).

#### 3.3 Halo occupation

We use an HOD model to occupy the ROCKSTAR halo catalogues with galaxies. Specifically, the average number of galaxies living in a halo of mass  $M$  and  $V_{\text{max}}$  is assumed to be split into a central and a satellite galaxy component:

$$\langle N_{\text{gal}} | M, V_{\text{max}} \rangle = \langle N_{\text{cen}} | M, V_{\text{max}} \rangle + \langle N_{\text{sat}} | M, V_{\text{max}} \rangle. \quad (13)$$

The number of centrals at a given halo mass  $M$  (averaged over  $V_{\text{max}}$ ) is parametrized by

$$\langle N_{\text{cen}} | M \rangle = \frac{f_\Gamma}{2} \left( 1 + \text{erf} \left[ \frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right] \right), \quad (14)$$

where  $f_\Gamma$ ,  $\log M_{\text{min}}$ , and  $\sigma_{\log M}$  are free parameters. The above equation describes a function that smoothly transitions from 0 to  $f_\Gamma \leq 1$  with  $\langle N_{\text{cen}} | M \rangle = f_\Gamma / 2$  at  $\log M_{\text{min}}$ . The rate of the transition is characterized by  $\sigma_{\log M}$ . Naively, one might expect that more massive haloes host brighter galaxies and thereby  $\lim_{\log M \rightarrow \infty} \langle N_{\text{cen}} | M \rangle = 1$  instead of  $\lim_{\log M \rightarrow \infty} \langle N_{\text{cen}} | M \rangle = f_\Gamma \leq 1$ . However, the BOSS LOWZ target selection misses some bright galaxies due to colour and magnitude cuts (Leauthaud et al. 2016). Thus, we can expect the central occupation number of high-mass haloes to be below unity (see e.g. Hoshino et al. 2015; Leauthaud et al. 2016; Guo, Yang & Lu 2018).

The average number of satellites is given by

$$\langle N_{\text{sat}} | M \rangle = \left( \frac{M - M_0}{M_1} \right)^\alpha \quad (15)$$

with  $M_0$ ,  $M_1$ , and  $\alpha$  being free parameters. The satellite number thus has a power-law dependence on halo mass with a lower cut-off characterized by  $M_0$ , particularly  $\langle N_{\text{sat}} | M \rangle = 0$  for  $M \leq M_0$ . Note that we do not implement a satellite galaxy analogue of  $f_\Gamma$  that accounts for colour cuts in LOWZ centrals since such a free parameter would be degenerate with  $M_1$ .

In order to allow for galaxy assembly bias, we utilize the decorated HOD framework of Hearin et al. (2017) to parametrize the number of galaxies as a function of  $V_{\text{max}}$  at a fixed halo mass  $M$ . In this framework, the average number of galaxies is derived by modifying

the average number at a fixed halo mass based on whether  $V_{\max}$  is above or below the median  $V_{\max}$  of all haloes at that mass,

$$\langle N_{\text{gal}} | M, V_{\max} \rangle = \langle N_{\text{gal}} | M \rangle \pm \delta N_{\text{gal}}. \quad (16)$$

In the above equation,  $\delta N_{\text{gal}}$  is added if  $V_{\max}$  is larger than the median and subtracted otherwise. This parametrization is not purely *ad hoc* but reflects models of galaxy formation. For example, dark matter haloes that formed earlier, i.e. those with high  $V_{\max}$ , have less substructure and therefore likely fewer satellites (Zentner et al. 2005; Jiang & van den Bosch 2017). In this case, we would expect  $\delta N_{\text{sat}} < 0$ , at least for satellite galaxies. In the following, we will again separate contributions from central and satellite galaxies. For centrals, we use

$$\delta N_{\text{cen}} = A_{\text{cen}} (0.5 - |0.5 - \langle N_{\text{cen}} | M \rangle|), \quad (17)$$

which fulfils the constraint  $0 \leq \langle N_{\text{cen}} | M, V_{\max} \rangle \leq 1$  for  $-1 \leq A_{\text{cen}} \leq 1$  as a free parameter. In contrast, for satellites we have

$$\delta N_{\text{sat}} = A_{\text{sat}} \langle N_{\text{sat}} | M \rangle, \quad (18)$$

where  $-1 \leq A_{\text{sat}} \leq 1$  is another free parameter. With these average numbers specified, we assume that the number of centrals follows a Bernoulli distribution and satellites a Poisson distribution.

### 3.4 Central galaxy phase-space coordinates

Central galaxies are assumed to coincide spatially with the halo centre defined as the average position of particles surrounding the dark matter halo phase-space density peak. Additionally, we assign a bulk velocity corresponding to the average velocity of all particles within 10 per cent of the halo radius  $r_h = R_{200b}$  around the density peak. We refer the reader to Behroozi et al. (2013) for details regarding the phase-space positions of haloes. The definition of the bulk velocity used here is very similar to the  $v_{\text{dens}}$  definition in Reid et al. (2014), which was shown to yield much better fits to available BOSS CMASS data than using the centre-of-mass velocity of each halo. Similarly, as shown in Ye et al. (2017), this definition traces the velocity of central galaxies in the Illustris hydrodynamical simulation much better than the centre-of-mass velocity.

Finally, following Reid et al. (2014) and Guo et al. (2015a), we add an additional Gaussian scatter to the line-of-sight component of the bulk velocity. The scatter has width  $\sigma$ ,

$$\sigma = \frac{\alpha_c V_{\text{vir}}}{\sqrt{3}}. \quad (19)$$

Here,  $V_{\text{vir}} = \sqrt{GM/r_h}$  is the circular speed at the halo radius and  $\alpha_c$  is a free parameter. Our definition of central velocity bias is the same as in Reid et al. (2014), where  $\alpha_c$  is called  $\gamma_{\text{cen}}$ , and similar to Guo et al. (2015a) with the difference being that the scatter is scaled by  $V_{\text{vir}}$  instead of the particle velocity dispersion. Physically, central velocity bias could originate from the central galaxy oscillating inside the dark matter halo or the dark matter halo not being fully relaxed (van den Bosch et al. 2005). When defining central velocity bias with respect to the halo core as given in ROCKSTAR, hydrodynamical simulations predict  $\alpha_c \lesssim 0.1$  (Ye et al. 2017).

### 3.5 Satellite galaxy phase-space coordinates

We model satellite galaxy positions and velocities using parametrized analytical forms instead of placing them on resolved subhaloes or dark matter particles (see e.g. Reid et al. 2014; Guo et al. 2015a, b). The number density of satellite galaxies within their host haloes is

assumed to follow an NFW profile,

$$n(r) \propto \frac{1}{r/r_s (1 + r/r_s)^2}. \quad (20)$$

In the above equation,  $r_s$  is the scale radius and is expressed via  $r_s = r_h/c_{\text{sat}}$ , where  $r_h$  is the host halo radius. We allow the concentration parameter of satellites  $c_{\text{sat}}$  to be different than that of the dark matter  $c_{\text{dm}}$  via the free parameter  $\eta$ :

$$c_{\text{sat}} = \eta c_{\text{dm}}. \quad (21)$$

The average velocity of satellites is assumed to be the bulk velocity of the dark matter halo. Additionally, we add a stochastic Gaussian scatter to the line-of-sight velocity of each satellite. The amount of scatter depends on the distance  $r$  from the halo centre and reflects satellite trajectories inside the dark matter halo. The width of the scatter is derived from solving the spherically symmetric Jeans equation without orbital anisotropy (van den Bosch et al. 2004) with an additional multiplicative factor  $\alpha_s$ :

$$\sigma^2 = \alpha_s^2 \frac{G\eta^2 c^2 M}{r_h g(c)} \left( \frac{r}{r_h} \right) \left( 1 + \frac{\eta cr}{r_h} \right)^2 \int_{\eta cr/r_h}^{\infty} \frac{g(y/\eta) dy}{y^3 (1+y)^2}, \quad (22)$$

where  $g(x) = \ln(1+x) - x/(1+x)$ . The definition of satellite velocity bias is the same as in Reid et al. (2014), Guo et al. (2015a, b), and Zhai et al. (2019), with  $\alpha_s$  also being called  $\gamma_{\text{HV}}$  or  $\eta_{\text{vs}}$ , in the sense that it is a scaling of the satellite velocities. However,  $\alpha_s$  is not the same as the ratio of the velocity dispersion of satellites and to that of dark matter particles. By solving the Jeans equation, this ratio is different from unity for  $\eta \neq 1$  and  $\alpha_s = 1$ . Instead,  $\alpha_s \neq 1$  describes a deviation from the prediction of the Jeans equation and can indicate that the kinematics of satellite galaxies is anisotropic, that satellite systems and/or dark matter haloes are not spherical, that satellite galaxies are not equilibrated within the halo potential, or any combination thereof. In general we thus have three free parameters,  $\alpha_c$ ,  $\alpha_s$ , and  $\eta$ , to describe the phase-space coordinates of galaxies.

### 3.6 Clustering predictions

We use the `s_mu_tpcf` and `tpcf_multipole` functions from `halotools.mock_observables` to predict redshift-space multipoles for mock galaxy populations. Particularly, we use the distant observer approximation whilst choosing one of the three axes of the simulation volume as the line of sight. In practice, we average results from projecting along each of the three simulation axes. The predictions do not take into account observational systematics like fibre collisions or survey boundary effects since we assume the measurements to be corrected for those. However, when making clustering predictions, we correct for the Alcock–Paczynski (AP) effect (Alcock & Paczynski 1979). The AP effect describes the phenomenon that inferred comoving coordinates of galaxies are derived from angular positions and redshift and thus depend on the reference cosmology assumed for this conversion. We take the AP effect into account by rescaling the phase-space coordinates of the mock galaxy population. Specifically, the line-of-sight coordinate is scaled via

$$\pi_{\text{ref}} = \pi_{\text{sim}} \frac{E(z|C_{\text{sim}})}{E(z|C_{\text{ref}})} \quad (23)$$

and the perpendicular coordinates are scaled via

$$r_{\text{p,ref}} = r_{\text{p,sim}} \frac{d_{\text{com}}(z|C_{\text{ref}})}{d_{\text{com}}(z|C_{\text{sim}})}. \quad (24)$$

**Table 1.** Prior limits for all galaxy–halo connection parameters assumed when calculating the evidence. We always assume flat priors within the limits stated above.

Parameter	Minimum	Maximum
$\log M_{\min}$	12.5	14.0
$\log M_0$	12.0	15.0
$\log M_1$	13.5	15.0
$\sigma_{\log M}$	0.1	1.0
$\alpha$	0.5	2.0
$f_{\text{r}}$	0.5	1.0
$A_{\text{cen}}$	−1.0	1.0
$A_{\text{sat}}$	−1.0	1.0
$\log \eta$	−log 3	+log 3
$\alpha_{\text{c}}$	0.0	0.4
$\alpha_{\text{s}}$	0.8	1.2

In the above two equations,  $\pi_{\text{ref}}$  and  $r_{\text{p, ref}}$  are the coordinates that would be inferred using a reference cosmology  $\mathcal{C}_{\text{ref}}$  different from the simulation cosmology  $\mathcal{C}_{\text{sim}}$ . The comoving distance  $d_{\text{com}}$  is expressed in  $h^{-1}$  Mpc and  $E(z) = \sqrt{\Omega_{\text{m}}(1+z)^3 + (1 - \Omega_{\text{m}})(1+z)^{3(1+w_0)}}$ . Consistent with Section 2, we use a spatially flat  $\Lambda$ CDM cosmology with  $\Omega_{\text{m}} = 0.307$  and  $w_0 = -1$  as our reference cosmology.

We make use of TABCORR<sup>3</sup> to speed up the predictions for mock observables. TABCORR implements the method described in Zheng & Guo (2016), whereby correlation functions between haloes as a function of  $M$  and  $V_{\text{max}}$  are tabulated. Afterwards, these correlation functions are convolved with  $\langle N_{\text{gal}} | M, V_{\text{max}} \rangle$  to get fast and accurate predictions for galaxy correlation functions. We refer the reader to Neistein & Khochfar (2012), Reid et al. (2014), Zheng & Guo (2016), and Lange et al. (2019c) for details.

We use 100 bins in  $\log M$  going from  $100 m_p = 3.51 \times 10^{12} (\Omega_{\text{m}}/0.3) h^{-1} M_{\odot}$  to the maximum halo mass in the simulation and two bins for  $V_{\text{max}}$  at fixed  $M$  corresponding to  $V_{\text{max}}$  values above and below the median. We have tested that the number of halo mass bins is sufficient given the precision of our observations. Halo correlation functions are tabulated for mock central–central, central–satellite, and satellite–satellite pairs. When calculating these auto- and cross-correlation functions, each halo is assigned one central and a Poisson number of satellite galaxies with a mean of  $M/10^{13} h^{-1} M_{\odot}$ , significantly larger than the number of satellites we expect to find per halo.

Finally, the correlation functions are tabulated for fixed values of the galaxy phase-space parameters  $\alpha_{\text{c}}$ ,  $\alpha_{\text{s}}$ , and  $\eta$ . To implement smooth variations for these parameters, we need to rely on interpolation. We first construct a sample of 200 points in  $\alpha_{\text{c}}$ ,  $\alpha_{\text{s}}$ , and  $\log \eta$  filling the prior space described in Table 1. The sample points are the centres of clusters obtained from  $k$ -means clustering of the prior space.<sup>4</sup> The space on which the  $k$ -means clustering algorithm is run extends slightly beyond the prior space to ensure that the sample points fully encompass it. We tabulate correlation functions for all 200 sample points and obtain predictions for arbitrary values of  $\alpha_{\text{c}}$ ,  $\alpha_{\text{s}}$ , and  $\log \eta$  through linear, barycentric interpolation. We test the accuracy of this approach through leave-one-out cross-validation, i.e. we test the accuracy of the prediction for a point on the grid when excluding it from the interpolation process. When choosing a model that reproduces the observational data, we find that all observables are predicted to within at least  $0.2\sigma$  accuracy

where  $\sigma$  is the observational uncertainty. The actual uncertainty due to interpolation is likely smaller because random points of the posterior generally have smaller distances to the 200 data points than a data point to its nearest neighbours. Thus, we can neglect errors related to interpolation because they would likely be negligible, i.e.  $\sqrt{1\sigma^2 + 0.2^2\sigma^2} \approx 1.02\sigma$ .

### 3.7 Cosmological evidence modelling

Throughout this paper, we use the CEM framework (Lange et al. 2019c) to derive cosmological constraints. The posterior constraint on cosmological parameters  $\mathcal{C}$  is obtained by marginalizing the full posterior over the galaxy–halo connection parameters  $\mathcal{G}$  described in Sections 3.3, 3.4 and 3.5:

$$P(\mathcal{C}|\mathbf{D}) \propto P(\mathcal{C}) \int \mathcal{L}(\mathbf{D}|\mathcal{C}, \mathcal{G}) P(\mathcal{G}) d\mathcal{G} = P(\mathcal{C}) \mathcal{Z}(\mathbf{D}|\mathcal{C}). \quad (25)$$

In the above equations,  $\mathbf{D}$  denotes the observational data,  $\mathcal{L}$  the likelihood and  $P(\mathcal{C})$  and  $P(\mathcal{G})$  the priors on cosmology and the galaxy–halo connection, respectively. Throughout this work, we assume flat priors, i.e.  $P(\mathcal{C}) \equiv P(\mathcal{G}) \equiv 1$  with ranges described in Table 1 for  $\mathcal{G}$  and the volume probed by the Aemulus simulation suite for  $\mathcal{C}$ . The integral  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$ , which we call cosmological evidence but is also known as the Bayesian evidence or marginal likelihood, can be calculated for each simulation of the Aemulus simulation suite, as detailed in Section 4.3. Thus, we have 40 samples of  $\mathcal{Z}(\mathbf{D}|\mathcal{C}) \propto P(\mathcal{C}|\mathbf{D})$  for different parameter combinations of  $\mathcal{C}$ . After developing and fitting a model for  $P(\mathcal{C}|\mathbf{D})$  to the 40 samples, we then have an estimate for the full cosmological posterior. Compared to the widely used emulation method (Nishimichi et al. 2019; Zhai et al. 2019; Wibking et al. 2020), this shifts the problem from emulating predictions for multi-dimensional observables  $\hat{\mathbf{D}}$  as a function of  $\mathcal{C}$  and  $\mathcal{G}$  to emulating a single number, the evidence, as a function of  $\mathcal{C}$ . The advantage of this approach is that, in principle, we can use arbitrarily complex models for  $\mathcal{G}$ . In contrast, the standard emulation approach will suffer from a decrease in the emulator accuracy with increasing complexity of  $\mathcal{G}$ . Additionally, the observables, i.e.  $\xi_{0,2,4}$ , can be complicated functions of  $\mathcal{G}$  and  $\mathcal{C}$ . Thus, to emulate them, one needs sophisticated, non-parametric approaches with large degrees of freedom like GP emulation. However, for the cosmological evidence, as we show in Section 4.4, simple parametric forms like multi-dimensional Gaussian functions are often sufficient. This has the potential of increasing the accuracy of the CEM method over the standard emulation technique.

## 4 TESTS ON MOCK CATALOGUES

Before analysing the data described in Section 2, we conduct mock tests to ensure that our analysis method can yield unbiased cosmology results. The tests we seek to conduct go beyond the investigations carried out by Zhai et al. (2019) and Lange et al. (2019c). These studies showed that unbiased cosmology results can be obtained when mock observations are created from the same model that is used to analyse the data. This result is non-trivial given the difficulty in predicting mock observables for arbitrary cosmology and HOD parameters. However, these findings do not show that unbiased cosmology results can be obtained for arbitrary galaxy populations. The main reason is that realistic galaxy populations will violate several of the assumptions made in Sections 3.3, 3.4, and 3.5 (see e.g. Beltz-Mohrmann, Berlind & Szecewicz 2020; Hadzhiyska et al. 2020). What remains unclear is to what extent these differences can bias our findings regarding cosmology here.

<sup>3</sup><https://github.com/johannesulf/TabCorr>

<sup>4</sup><https://scikit-learn.org/stable/modules/clustering.html#k-means>

#### 4.1 SHAM galaxy model

To explore to what extent complex galaxy populations can bias our cosmological inference, we utilize mock catalogues based on a SHAM model (see Wechsler & Tinker 2018, for a review). Specifically, we use the SHAM model first introduced by Lehmann et al. (2017). Briefly, this model populates all dark matter (sub-)haloes above a threshold in  $V_\alpha$  with galaxies. Here,  $V_\alpha$  is defined via the virial velocity  $V_{\text{vir}}$  and the maximum circular velocity  $V_{\text{max}}$  of a dark matter (sub-)halo,

$$V_\alpha = V_{\text{vir}} \left( \frac{V_{\text{max}}}{V_{\text{vir}}} \right)^\alpha. \quad (26)$$

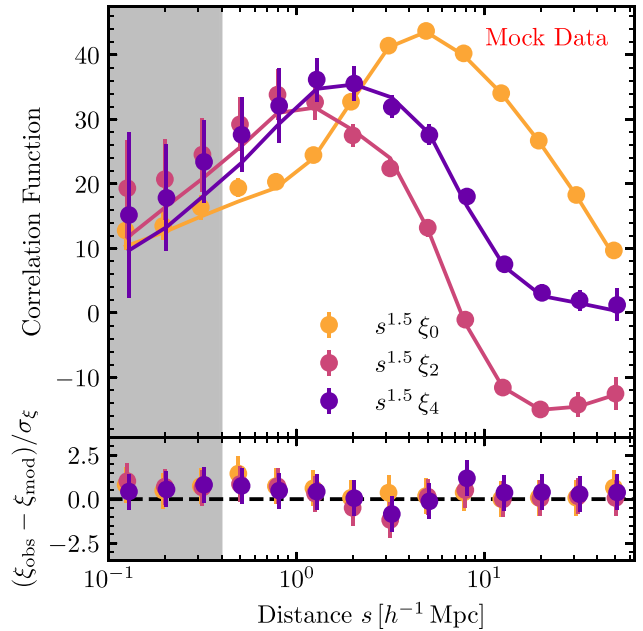
The free variable  $\alpha$  determines the degree to which the matching depends on the concentration of the dark matter halo. For  $\alpha \neq 0$ ,  $V_\alpha$  depends on both halo mass and concentration. Ultimately, the average occupation and moments constructed using this SHAM model will deviate from the parametrized forms assumed for our HOD model, as described in Section 3.3. Furthermore, satellite galaxies in this model follow the dynamics of resolved subhaloes in simulations. Thus, we expect them to violate some of the assumptions made in Section 3.5, such as spherical symmetry or orbital isotropy. However, since the SHAM is performed on dark matter-only simulations, the mock catalogues do not include the effects of baryonic feedback on the matter distribution and dynamics. We leave the study of these effects on RSD measurements for future work.

#### 4.2 Mock measurements

Our mock measurements are based on applying the SHAM model described above to the UNIT simulations (Chuang et al. 2019). Specifically, we use four simulations with a fixed power-spectrum normalization but complementary phases. Each of these simulations has  $4096^3$  dark matter particles in a volume of  $(1 h^{-1} \text{ Gpc})^3$ . We study the simulation snapshots at redshift  $z = 0.39$ . The cosmological parameters of these simulations follow the Planck Collaboration (2016) CMB analysis, i.e.  $\Omega_m = 0.3089$ ,  $h = H_0/100 = 0.6774$ ,  $n_s = 0.9667$ , and  $\sigma_8 = 0.8147$ . Thus, the  $f\sigma_8(z = 0.39)$  value we seek to recover is 0.476.

The mock galaxy sample is constructed by setting the SHAM model parameter  $\alpha$  to 0.73 as motivated by the findings of Lehmann et al. (2017). We have also tested our results with  $\alpha = 0$  and  $\alpha = 1.5$ , finding nearly identical cosmological constraints, as described below. Instead of applying a hard cut-off in  $V_\alpha$  for each (sub)halo, we use a smooth transition. For field haloes, the probability to host a galaxy is 50 per cent at  $V_\alpha \approx 700 \text{ km s}^{-1}$  and increases roughly linear with a rate of  $\sim 0.1$  per cent  $\text{km}^{-1} \text{ s}$ . This simulates scatter between halo and galaxy properties. For subhaloes, we require a slightly higher  $V_\alpha \approx 760 \text{ km s}^{-1}$  for a 50 per cent chance to host a satellite. We choose these slightly different  $V_\alpha$  values for centrals and satellites in order for the mock clustering measurements to roughly match the observations of the  $0.3 < z \leq 0.43$  sample. Additionally, they are chosen such that the mock catalogues reproduce the number density in observations of the same sample.

Mock measurements are obtained for the same statistics as for the observations. We calculate the redshift-space clustering by projecting each of the four simulations along each of the three spatial axes,  $x$ ,  $y$ , and  $z$ . As our final mock data vector, we take the mean of the 12 mock measurements. As the covariance matrix, we use the same covariance matrix as for the  $0.3 < z \leq 0.43$  BOSS sample. The resulting mock measurements and uncertainties are shown in Fig. 4. In the same figure, we also show the best fit to the mock data using



**Figure 4.** Mock measurements of the multipole moments of redshift-space correlation function. The measurements are based on applying a SHAM model to the UNIT simulations. Since the input cosmology is known, we use these measurements to test our analysis pipeline. The best-fitting model marginalized over all the Aemulus simulations and HOD parameters is shown by the solid line. All quantities are expressed in units of  $h^{-1} \text{ Mpc}$ . As indicated by the grey band, scales  $s = 400 h^{-1} \text{ kpc}$  are excluded from the fit.

the HOD model. Specifically, the fit shown is the best fit obtained over all 40 Aemulus  $z = 0.40$  simulations outputs after marginalizing over all HOD parameters listed in Table 1. Because the covariance matrix reflects our observational volume of  $\sim 0.6 (\text{Gpc}/h)^3$  instead of the  $4 (\text{Gpc}/h)^3$  volume of the mock data, the best-fitting  $\chi^2 = 8.2$  is significantly smaller than the number of measurement points.

#### 4.3 Evidence calculation

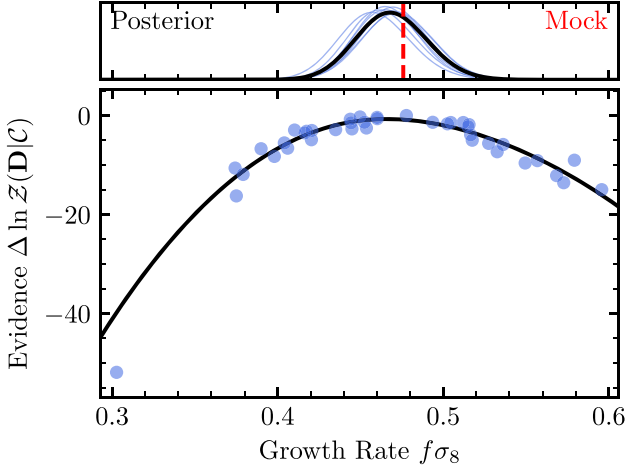
We use the CEM method to extract cosmological information from the mock data using the Aemulus simulation suite. As a first step, we define the likelihood via

$$\ln \mathcal{L}(\mathcal{C}, \mathcal{G}) = \frac{(n_{\text{gal}} - \hat{n}_{\text{gal}})^2}{2\sigma_{n_{\text{gal}}}^2} + \frac{1}{2}(\xi - \hat{\xi})^T \Sigma^{-1}(\xi - \hat{\xi}). \quad (27)$$

In the above equation,  $\xi$  contains the monopole, quadrupole, and hexadecapole measurements, i.e.  $\xi = [\xi_0, \xi_2, \xi_4]$ .  $\Sigma$  is the covariance matrix of the LOWZ clustering measurements and  $\sigma_{n_{\text{gal}}}$  is the observational error on the number density. Finally, to calculate the integral in equation (25), we use MULTINEST (Feroz, Hobson & Trotta 2010) with 3000 live points and a target efficiency of 4 per cent. We have confirmed that our results are converged by repeating analyses with 6000 live points, as well as a target efficiency of 1 per cent and getting indistinguishable cosmology and galaxy–halo connection parameter constraints. In all cases, constant efficiency mode and importance nested sampling are turned off.

#### 4.4 Evidence modelling

In the lower panel of Fig. 5, we show a scatter plot of the evidence  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$  as a function of  $f\sigma_8(z = 0.40)$  for each of the 40 regular



**Figure 5.** CEM analysis of mock data. The lower panels show the cosmological evidence  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$  as a function of  $f\sigma_8$  for the 40 regular simulations of the Aemulus simulation suite. The evidence is calculated with respect to the mock observations described in Section 4. The line shows the best-fitting skew normal to the evidence values. The upper panel shows as thin lines random draws from the posterior of skew normal distributions fitted to the evidence values. Finally, the solid black line in the upper panel is the superposition of all skew normal distributions of the posterior serving as the effective posterior constraint on  $f\sigma_8$ . This posterior constraint on  $f\sigma_8$  compares favourably to the input value (red dashed) in the mocks that we seek to recover.

simulations of the Aemulus simulation suite. From linear theory, we expect observations of the redshift-space correlation function on large, linear scales to primarily constrain  $f\sigma_8$ . Thus, it is natural to expect the same to hold true on non-linear scales. And indeed, Fig. 5 shows that the evidence is a strong function of  $f\sigma_8$ . As discussed in Lange et al. (2019c), because the cosmological evidence is directly proportional to the posterior on cosmology, once we have a reliable model for  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$ , we also have obtained posterior constraints on cosmology.

#### 4.4.1 1D models

For the moment, we follow Lange et al. (2019c) and assume that  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$  is a function of  $f\sigma_8$  only. We would naively expect the posterior constraint on  $f\sigma_8$  to be similar to a normal distribution. A skew normal distribution is a natural extension by allowing non-zero skewness. This additional freedom is found in Lange et al. (2019c) to be necessary to fit  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$  for other mock RSD data. Thus, this functional form is primarily empirically motivated. The skew normal distribution is parametrized by

$$\mathcal{Z}(\mathbf{D}|\mathcal{C}) \approx \hat{\mathcal{Z}}(\mathbf{D}|f\sigma_8) \propto \left[ 1 + \operatorname{erf} \left( \frac{\alpha(f\sigma_8 - \mu)}{\sqrt{2}\sigma} \right) \right] \exp \left[ -\frac{(f\sigma_8 - \mu)^2}{2\sigma^2} \right]. \quad (28)$$

Here,  $\mu$ ,  $\sigma$ , and  $\alpha$  are free parameters to be obtained by fitting the distribution shown in the lower panel of Fig. 5. We note that some simulations may not be able to yield reasonable fits to the data, thus resulting in very low evidence values. Since the cosmological parameters of such simulations would be ruled out anyway, we do not want them to influence the fit to the parametric evidence form  $\hat{\mathcal{Z}}$  since this form is purely empirically motivated and might not hold over large ranges in  $f\sigma_8$ . Thus, we remove all simulations for which the best-fitting predicted evidence  $\ln \hat{\mathcal{Z}}$  value is 25 below the maximum best-fitting evidence. This roughly corresponds to simulations that

are more than  $7\sigma$  away from the maximum posterior. In the present case, this only removes one simulation for which  $f\sigma_8 \sim 0.3$ .

From Fig. 5, it is also apparent that any reasonable parametric form for  $\hat{\mathcal{Z}}$  will not perfectly fit the evidence values of all simulations. Some amount of scatter between the predicted and the measured  $\mathcal{Z}(\mathbf{D}|f\sigma_8)$  is expected and can be due to a variety of reasons. For example, part of the scatter could be due to our approximation that the evidence values only depend on  $f\sigma_8$ . However, even if the evidence value was purely a function of  $f\sigma_8$ , a non-zero scatter would be expected due to cosmic variance. Each simulation probes a finite cosmological volume. Thus, the model predictions of each simulation will be affected by the random realizations of the phases and amplitudes of the perturbation modes in the initial conditions of each simulation. Furthermore, sampling noise when probing the phase-space properties of galaxies (see Sections 3.4 and 3.5), as well as errors when interpolating between  $\alpha_s$ ,  $\alpha_c$ , and  $\eta$  values could further add to the scatter. For the moment, we assume that the main source of the scatter in  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$  is truly random, i.e. cosmic variance, of the simulations themselves. We present further evidence for this assumption in the appendix. If the assumption holds, Lange et al. (2019c) advocated for the following functional form for the evidence scatter,

$$\Delta \ln \mathcal{Z} \approx \sqrt{\frac{r_{\text{sim}}^2 N_{\text{data}}}{2} + r_{\text{sim}} \chi_{\text{min}}^2}. \quad (29)$$

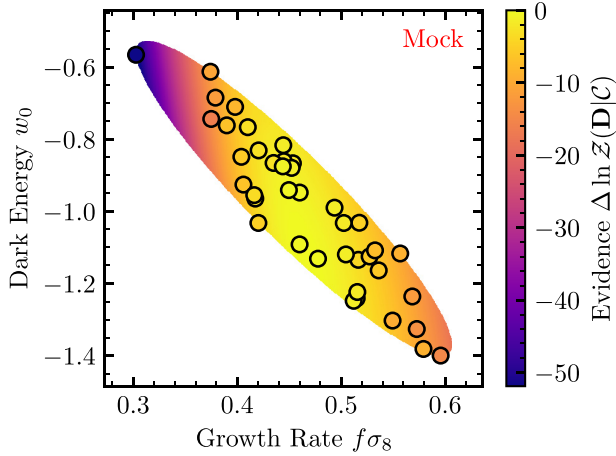
Here,  $N_{\text{data}}$  is the number of data points,  $\chi_{\text{min}}^2$  the best-fitting  $\chi^2$  for each simulation, and  $r_{\text{sim}}$  another free parameter. In essence, the above functional form allows for arbitrary scatter values that increase with increasing  $\chi^2$  of each simulation.

We choose flat priors for  $\mu$ ,  $\delta = \alpha/\sqrt{1+\alpha^2}$  and  $r_{\text{sim}}$  with ranges  $[0.35, 0.60]$ ,  $(-1, 1)$  and  $[0, 1]$ , respectively. Similarly, the scatter  $\sigma$  is parametrized via  $\tilde{\sigma} = \sigma/(\sigma_{\text{sim}} + \sigma)$  using a flat prior with range  $[0, 1]$ , where  $\sigma_{\text{sim}}$  is the scatter of  $f\sigma_8$  values of the Aemulus simulations. When fitting  $\mathcal{Z}(\mathbf{D}|f\sigma_8)$ , we obtain a posterior sample of skew normal distributions that could fit the evidence values in Fig. 5. In order to account for uncertainties in the fit, we choose as the final cosmological posterior, the average of the normalized skew normal distributions in our posterior sample. The upper panel of Fig. 5 shows as thin blue lines random draws from the posterior and as a thick black line the total posterior. Overall, we infer  $f\sigma_8 = 0.469 \pm 0.021$ , comparing favourably to the input value  $f\sigma_8 = 0.476$ .

#### 4.4.2 2D models

In the 1D model for  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$ ,  $\mathcal{Z}(\mathbf{D}|\mathcal{C}) \approx \hat{\mathcal{Z}}(\mathbf{D}|f\sigma_8)$ , we have implicitly assumed that any deviation from the best-fitting model originates entirely from random errors associated with cosmic variance in each simulation volume. However, it is possible that scatter around the best-fitting  $\mathcal{Z}(\mathbf{D}|f\sigma_8)$  is a sign that the evidence depends on parameters besides  $f\sigma_8$ . In particular, there could be a degeneracy between  $f\sigma_8$  and other cosmological parameters, such as the anisotropic deformation parameter  $\epsilon$  of the AP effect. Neglecting such a degeneracy could, in principle, result in artificially narrow or otherwise biased posterior constraints on  $f\sigma_8$ . Here, we test for such a degeneracy by modelling the evidence with 2D models, i.e.  $\mathcal{Z}(\mathbf{D}|\mathcal{C}) = \mathcal{Z}(\mathbf{D}|f\sigma_8, \gamma)$  with  $\gamma$  being a suitable second cosmological parameter.

As the functional form for  $\mathcal{Z}(\mathbf{D}|f\sigma_8, \gamma)$ , we choose a 2D skew-normal (Azzalini & Valle 1996), a natural extension to the one-dimensional case. We fix the skewness over the second cosmological parameter,  $\alpha_\gamma$ , to 0, choose a flat prior for  $\mu_\sigma$  in the range sampled



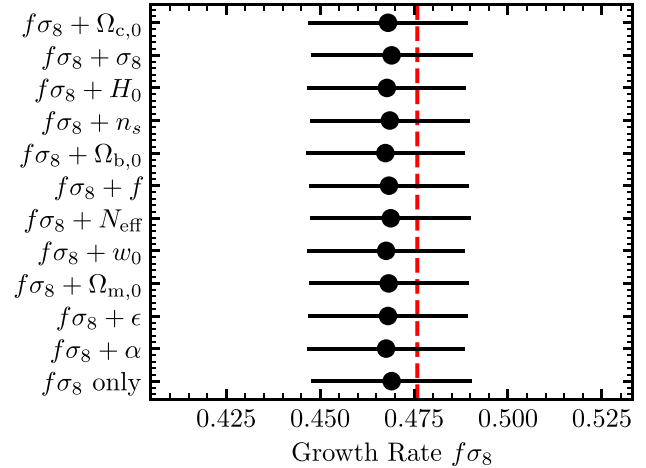
**Figure 6.** Cosmological evidence as a function of  $f\sigma_8$  and the dark energy equation-of-state parameter  $w_0$  for the mock data. Each point corresponds to one of the 40 Aemulus simulations. The background colour in the ellipse is the best-fitting model to the points.

by the Aemulus simulations and add a flat prior for  $r$ , the correlation between  $f\sigma_8$  and  $\gamma$  in the range  $[-1, +1]$ . The scatter of the second parameter,  $\sigma_\gamma$ , has a similar prior as the one for  $f\sigma_8$ , i.e. we choose flat prior for  $\tilde{\sigma}_\gamma = \sigma_\gamma/(\sigma_{\gamma,\text{sim}} + \sigma_\gamma)$  in the range  $[0, 1]$ . Ultimately, this evidence model has seven free parameters.

In Fig. 6, we show a 2D analysis in which the second parameter  $\gamma$  is the dark energy equation-of-state parameter  $w_0$ . As in the previous figure, each point corresponds to one simulation of the Aemulus simulation suite. The colour indicates the evidence value associated with each simulation. The background colour now shows the best-fitting 2D model. In this case, this model predicts that the evidence changes primarily along the  $f\sigma_8$ -direction, as expected. We can now, in principle, infer constraints on  $f\sigma_8$  by marginalizing over the  $w_0$  direction and proceeding in the same way as for the 1D evidence model. However, doing so naively would result in rather weak constraints on  $f\sigma_8$  because the available data cannot rule out a weak degeneracy between  $f\sigma_8$  and  $w_0$  such that the marginalised  $f\sigma_8$  posterior is very different for  $w_0$  far away from  $-1$ . Therefore, we explicitly impose a prior related to the cosmological parameter space probed by the simulations. We do so by constructing the minimum bounding ellipse around simulation parameters in the  $f\sigma_8 - w_0$  plane, as shown in Fig. 6. After imposing this strict prior on  $f\sigma_8$  and  $w_0$ , we infer a marginalized constraint of  $f\sigma_8 = 0.468 \pm 0.021$ , in very good agreement with the 1D result.

#### 4.5 Accuracy of cosmological constraints

Fig. 7 shows how the marginalized constraints on  $f\sigma_8$  depend on modelling choices. Specifically, the lowest error bar shows the result for the 1D model and all other error bars results of 2D models. In this figure,  $\epsilon$  and  $\alpha$  refer to the distortion parameters of the AP effect (Padmanabhan & White 2008). Most importantly, we see that all analysis methods result in very similar marginalized  $f\sigma_8$  constraints. All shifts in the mean of the posterior are negligible and the uncertainty on  $f\sigma_8$  is virtually unaffected. Additionally, we find that these 2D models are not statistically significantly preferred over the 1D model. Regardless of which second cosmological parameter is considered, the evidence ratios in favour of a 2D model do not exceed 2. All of this suggests that our modelling of the cosmological evidence is robust, even when assuming it only



**Figure 7.** Marginalized constraints on  $f\sigma_8$  for the mock as a function of the model for the evidence. The lowest error bar shows the  $f\sigma_8$  constraint if the evidence is assumed to only depend on  $f\sigma_8$ . All other error bars show results if the evidence is assumed to depend on  $f\sigma_8$  and an additional cosmological parameter. The input value from mock data set is shown by the red dashed line.

depends on  $f\sigma_8$ . We emphasize that this finding is likely dependent on explicitly imposing a prior on the cosmological parameters probed by the Aemulus simulations. For example, one might expect that  $f\sigma_8$  and  $\epsilon = \left( \frac{d_{\text{com}}(z|C_{\text{ref}})E(z|C_{\text{ref}})}{d_{\text{com}}(z|C)E(z|C)} \right)^{1/3} - 1$  are degenerate since both cosmological parameters determine anisotropies in the redshift-space correlation function. However, the Aemulus prior limits  $\epsilon$  to  $\sim \pm 0.005$  at fixed  $f\sigma_8$ . Thus, within the Aemulus prior, we are not able to confidently detect such a degeneracy and, even if it was present, would not significantly affect the constraint on  $f\sigma_8$ . But we stress that our constraints on  $f\sigma_8$  should not be regarded as model-independent. Instead, they are valid for cosmologies probed by or close to those of the Aemulus simulation suite.

For simplicity, we choose the result from the 1D modelling as the final result on  $f\sigma_8$ . Note that our mock observations are constructed using a volume of  $4(h^{-1} \text{ Gpc})^3$  whereas the covariance matrix corresponds to a volume of  $\sim 0.6(h^{-1} \text{ Gpc})^3$ . Furthermore, for the SHAM mock catalogues, we project on to the three simulation axes. Since most of our constraining power comes from small scales, as shown later, this effectively triples the volume (Smith et al. 2021). Thus, we should expect to be able to recover the input  $f\sigma_8 = 0.476$  to within roughly  $\pm 0.021/\sqrt{3 \times 4/0.6} = 0.005$ . At the same time, there is an additional uncertainty in the best-fitting evidence model  $\hat{Z}(f\sigma_8)$ . For example, the maximum evidence value is inferred to occur at  $f\sigma_8 = 0.468 \pm 0.004$ . Ultimately, the difference between the maximum  $f\sigma_8$  evidence and the input is 0.008, too small to confidently claim a detection of a bias given uncertainties in the mock observables and evidence model. Larger simulations, both for the mocks and for the modelling, would be needed to robustly detect a possible bias. Overall, if our recovered values for  $f\sigma_8$  are biased, the bias is likely lower than statistical uncertainties for  $f\sigma_8$  from current observations.

We have repeated the above exercise with  $\alpha = 0$  and  $\alpha = 1.5$  when constructing the SHAM mock. Changing  $\alpha$  impacts the galaxy assembly bias strength and the satellite fraction of the mock sample (Lehmann et al. 2017). Consequently, we find that the multipoles change significantly. For example, the quadrupole  $\xi_2$  on small scales changes by  $\sim 20$  per cent. Encouragingly, despite these significant changes in the observables, we find nearly identical  $f\sigma_8$  constraints,  $f\sigma_8 = 0.470 \pm 0.021$  and  $f\sigma_8 = 0.467 \pm 0.022$ . The small shift

in  $f\sigma_8$  makes sense given that the three mocks are derived from the same halo catalogue. Thus, the mock observations constructed using different values for  $\alpha$  are likely highly correlated. In addition to changing  $\alpha$ , we performed a similar mock test for an analogue to the  $0.18 < z \leq 0.30$  sample that we derived from the  $z = 0.25$  UNIT simulation outputs and analyse with the  $z = 0.25$  Aemulus outputs. Using  $\alpha = 0.73$  for this low-redshift mock, we find  $f\sigma_8 = 0.477 \pm 0.022$  for an input values of 0.469. As before, the recovery of  $f\sigma_8$  is satisfactory given the volume of the mock observations and uncertainties in the evidence model. Overall, our tests with SHAM mocks do not indicate any significant biases in the recovery of  $f\sigma_8$ .

#### 4.6 Impact of redshift evolution

The observations described in Section 2 cover a range of redshifts whilst the Aemulus simulation snapshots we use to model them are coming from a single redshift. To mitigate the impact of potential biases, we construct an observational galaxy sample that is roughly volume-limited and use simulation snapshots that correspond to approximately the mean redshift of galaxies in the sample. None the less, we want to test here whether having observations taken at slightly different redshifts than the simulation output can bias our inferences. If our modelling was sensitive to only  $f\sigma_8$ , we would not expect redshift-evolution effects to be very relevant given that the change in  $f\sigma_8$  within each redshift bin is expected to be very small, of order  $\sim 0.01$ , smaller than the observational uncertainties.

To test this line of argument explicitly, we performed a mock test for the  $0.30 < z \leq 0.43$  sample using  $\alpha = 0.73$ . However, instead of constructing the mock galaxy population from the  $z = 0.4$  snapshot of the UNIT simulations, we derive it from the  $z = 0.25$  snapshot. This simulates a mismatch between the redshift of the data,  $z = 0.25$ , and the redshift assumed for the modelling,  $z = 0.39$ . We infer  $f\sigma_8 = 0.466 \pm 0.021$  which is lower by 0.003 than the mock constraint based on the comparable mock coming from the  $z = 0.4$  snapshot. This shift to slightly lower  $f\sigma_8$  is expected given that  $f\sigma_8$  reduces by 0.007 going from  $z = 0.4$  to  $z = 0.25$  for the cosmology of the UNIT simulation. We conclude that redshift evolution effects are likely only relevant to the extent that  $f\sigma_8$  changes within a redshift bin. Since those are negligible compared to our observational uncertainties, redshift evolution inside each redshift bin is unlikely to have a significant effect on our results.

#### 4.7 Effects beyond SHAM mocks

The test performed here are a first, highly non-trivial check that unbiased cosmological constraints can be obtained from the all-scale analysis of RSD data and goes beyond the results presented in Zhai et al. (2019) and Lange et al. (2019c). However, our analysis does not conclusively prove that unbiased cosmological constraints can always be achieved. Ultimately, further tests should be performed on a larger and more diverse sample of mock catalogues, ideally extracted from hydrodynamical simulations. In the following, we describe several effects that should be tested more thoroughly in the future.

##### 4.7.1 Baryonic feedback

Since our modelling is based on dark matter-only simulations, we do not directly model the impact of baryonic feedback processes. For example, baryonic feedback is known to affect the matter distribution on small scales, particularly within a dark matter halo (Jing et al. 2006; Rudd, Zentner & Kravtsov 2008; van Daalen et al. 2011; Chisari

et al. 2018; Villaescusa-Navarro et al. 2021). Since we marginalize over central and satellite phase-space parameters, some of the impact of baryonic feedback might be modelled through this approach.

The impact of baryons on the redshift-space clustering was discussed in Hellwing et al. (2016). Particularly, the authors show that in the EAGLE hydrodynamical simulation, the impact of baryons on the dark matter redshift-space monopole and quadrupole power spectra can be as large as  $\sim 4$  per cent up to  $k = 20 h^{-1} \text{ Mpc}$ . However, the effect is primarily due to the suppression of the *Fingers of God* effect, something we implicitly model via the parameter  $\alpha_s$ . At the same time, Hellwing et al. (2016) show that the impact of baryons on the peculiar velocities of dark matter haloes hosting massive galaxies,  $M_* > 3.5 \times 10^{10} M_\odot$  is much weaker. Particularly, for matched haloes in the dark matter-only and the hydrodynamical run, the velocity offset is consistent with zero and the  $1\sigma$  scatter between the peculiar velocities is only  $\sim 7 \text{ km s}^{-1}$ , significantly below typical halo velocities.

We conclude that baryons are unlikely to impact our conclusions in this work at a significant level. However, the back-reaction of baryons on the dark matter density and velocity field varies widely between different implementations of baryonic feedback (Springel et al. 2018). A detailed study of baryonic feedback on cosmological constraints from small-scale redshift space distortions is beyond the scope of this paper but warranted for future studies that use even higher precision measurements.

##### 4.7.2 Galaxy–halo connection

Our modelling is based on a complex galaxy–halo model that is tested on non-trivial mock catalogues. None the less, it is, in principle, possible that our parametrized galaxy–halo model is not flexible enough to yield unbiased cosmology constraints from more complex galaxy models or real data.

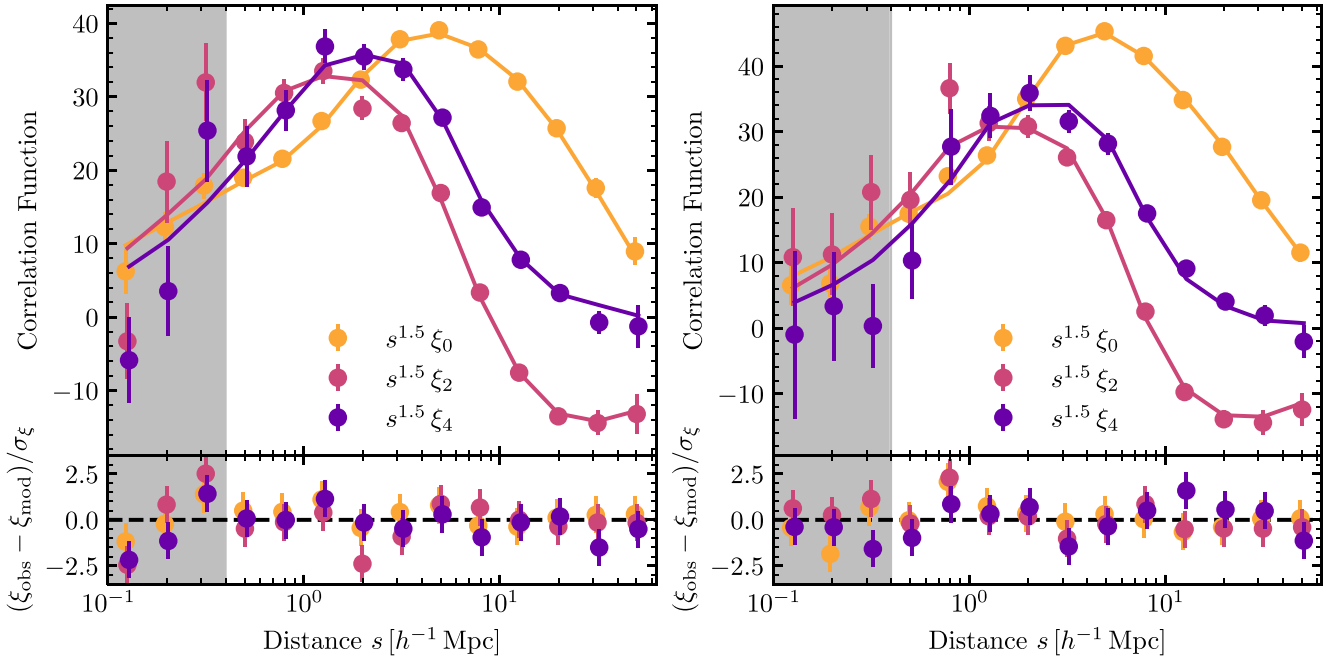
For example, both our HOD model and our SHAM mock implement galaxy assembly bias through a correlation of galaxy number with  $V_{\text{max}}$ . However, Hadzhiyska et al. (2020) and Xu, Zehavi & Contreras (2021) have recently shown in theoretical models of galaxy formation that  $V_{\text{max}}$  is insufficient to capture the full galaxy assembly bias effect. Similarly, Yuan et al. (2021) argue that assembly bias with respect to both concentration and large-scale over-density is needed to get a good fit to the redshift-space clustering of BOSS CMASS galaxies. However, even if overdensity-based assembly bias was needed to model the data, it is not clear that this biases  $f\sigma_8$  constraints significantly since neglecting  $V_{\text{max}}$ -based assembly bias does not significantly alter the  $f\sigma_8$  posterior constraints (Lange et al. 2019c). Additionally, as described in Section 5, we obtain a very good fit to the data without over-density-based assembly bias, contrary to Yuan et al. (2021).

## 5 APPLICATION TO OBSERVATIONS

After performing successful tests on mock catalogues, we now model the actual observations. In Fig. 8, we show the measurements of the multipole moments of the redshift-space correlation functions in the two redshift bins. We also show as solid lines the respective best-fitting models.

### 5.1 Goodness of fit

Using scales from  $400 h^{-1} \text{ kpc}$  to  $63 h^{-1} \text{ Mpc}$ , we find  $\chi^2 = 21.1$  and 22.8 for the  $z = 0.25$  and  $z = 0.40$  samples, respectively. The



**Figure 8.** Measurements of the monopole (blue), quadrupole (purple), and hexadecapole (orange) moments of the redshift-space correlation function. The left-hand side shows the measurements for the  $z = 0.25$  sample and the right-hand side the observations for the  $z = 0.40$  sample. All quantities are expressed in units of  $h^{-1}$  Mpc. We also show as solid lines the best-fitting HOD model from the Aemulus simulation suite. The lower panels show the difference between the observations and the best-fitting models in units of the observational uncertainty. Scales below  $s = 400 h^{-1}$  kpc are excluded from the fit.

number of measurements points is 34. Additionally, we have 11 galaxy-halo connection parameters and 7 cosmological parameters. However, the number of degrees of freedom of the model is likely smaller than  $11 + 7$ . Most importantly, we argued in Section 4 that there is effectively only one cosmological parameter determining the fit such that the effective number of degree of freedom of the model is  $\lesssim 12$ . Hence, the number of degree of freedom of the fit is  $\gtrsim 22$ , in which case the  $\chi^2$  values indicate a good fit of the model to the data.

## 5.2 Constraints on $f\sigma_8$

We calculate the cosmological evidence with respect to the two observations for all 40 simulations of the Aemulus simulation suite. The results are listed in Appendix A, thereby allowing the reader to derive cosmological constraints for different models for the evidence  $\mathcal{Z}(\mathbf{D}|\mathcal{C})$ .

In Fig. 9, we show the distribution of  $\ln \mathcal{Z}$  as a function of  $f\sigma_8$  for both observations and all simulations. In the same figure, we also show the derived constraints on  $f\sigma_8$  if the evidence is assumed to only depend on  $f\sigma_8$ . We find  $f\sigma_8 = 0.471 \pm 0.024$  and  $0.430 \pm 0.025$  for redshifts 0.25 and 0.40, respectively. Similar to Section 4, we find that the  $f\sigma_8$  constraints are virtually unchanged if the evidence is assumed to depend on additional cosmological parameters like  $w_0$  or  $\epsilon$ .

## 5.3 Robustness to analysis choices

In the following, we will analyse how our constraints on  $f\sigma_8$  change if we apply reasonable variations to the default analysis. All results are tabulated in Table 2.

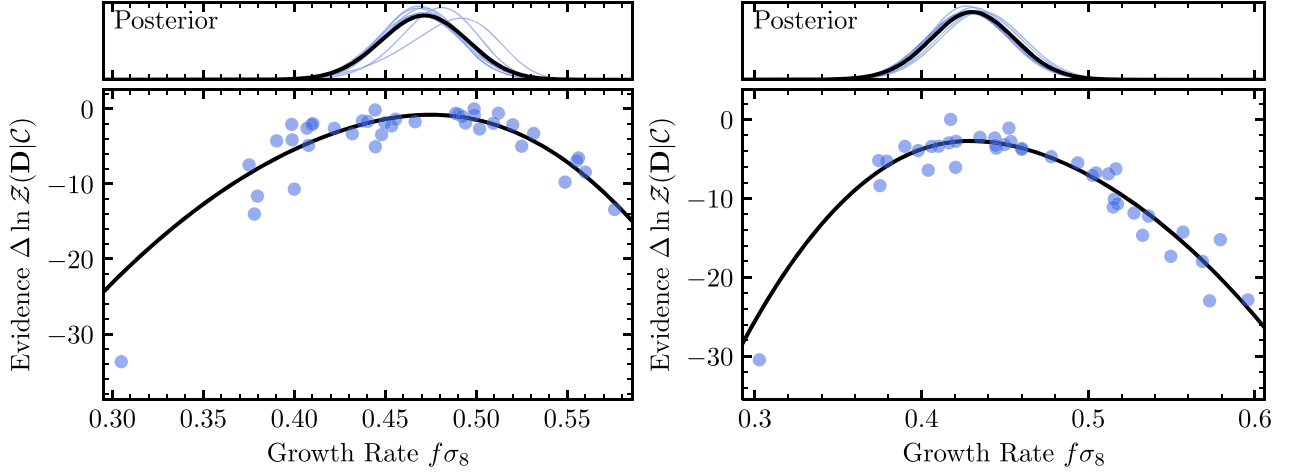
### 5.3.1 Scale dependence

By default, we use scales within  $400 h^{-1} \text{ kpc} < s < 63 h^{-1} \text{ Mpc}$  in our analysis. Here, we study how different scale cuts, both on the lower and the upper limit in  $s$ , affect our cosmological constraints. In Fig. 10, we show the constraints on  $f\sigma_8$  for the two galaxy samples and different scales cuts.

For the low-redshift sample, excluding smaller scales shifts the posterior constraint on  $f\sigma_8$  upward, reaching  $0.538 \pm 0.040$  when excluding scales below  $6.3 h^{-1} \text{ Mpc}$ . Similarly, excluding large scales drives  $f\sigma_8$  down, reaching  $0.415 \pm 0.029$  when limiting the analysis to scales below  $10 h^{-1} \text{ Mpc}$ . That small- and large-scale cuts drive the constraints in opposite direction seems to be a consequence of small scales ( $< 10 h^{-1} \text{ Mpc}$ ) preferring lower values for  $f\sigma_8$  than larger scales. The  $> 6.3 h^{-1} \text{ Mpc}$  and the  $< 10.0 h^{-1} \text{ Mpc}$  samples, with the exception of  $6.3 h^{-1} \text{ Mpc} < s < 10 h^{-1} \text{ Mpc}$ , use completely different scales. Here, we find a  $\sim 2.5\sigma$  discrepancy between the two scale cuts for the  $0.18 < z \leq 0.3$  galaxy sample. We also observe a similar albeit weaker behaviour for the  $0.3 < z \leq 0.43$  sample. Whilst the smaller scales also prefer lower values for  $f\sigma_8$ , the discrepancy is only of the order of  $\sim 1\sigma$ . Additionally, we do not observe comparable trends in the mock tests in Section 4 for any of the five mock samples analysed. These trends are discussed further in the discussion section.

### 5.3.2 Impact of hexadecapole

By default, we use the monopole  $\xi_0$ , quadrupole  $\xi_2$ , and hexadecapole  $\xi_4$  of the redshift-space correlation function as observables. We repeat the analysis here without the hexadecapole  $\xi_4$  to see how sensitive our analysis is with respect to this observable. We now find  $f\sigma_8 = 0.470 \pm 0.026$  and  $0.433 \pm 0.023$  for the two samples, indicating only a negligible shift and little impact on the precision of



**Figure 9.** Cosmological evidence as a function of  $f\sigma_8$  for the BOSS LOWZ data. The left-hand panel shows the results for the  $z = 0.25$  sample and the right-hand side the findings for the  $z = 0.40$  sample. Blue points in the lower panels correspond to each of the 40 regular Aemulus simulations. Similarly, the black line is the best-fitting model to the evidence. The upper panels show draws from the posterior of the evidence model as thin blue lines and the total effective  $f\sigma_8$  posteriors as the black lines.

**Table 2.** Dependence of the cosmological constraints to variations in the analysis choices. In order, we show the results for the default choices, when excluding the hexadecapole, when using the Hartlap et al. (2007) correction to the covariance matrix, when including the projected correlation function and when applying a variety of different scale cuts. For the degrees of freedom, d.o.f., we set  $\text{d.o.f.} = N_{\text{data}} - 12$ , where  $N_{\text{data}}$  is the number of data points and 12 is the number of free parameters in the model, i.e. 11 galaxy–halo connection parameters and one effective cosmological parameter,  $f\sigma_8$ .

Method	$N_{\text{data}}$	$f\sigma_8$	$\chi^2$	$\chi^2/\text{d.o.f.}$	$f\sigma_8$	$\chi^2$	$\chi^2/\text{d.o.f.}$
			$0.18 < z \leq 0.3$			$0.3 < z \leq 0.43$	
Default	34	$0.471 \pm 0.024$	21.1	0.96	$0.430 \pm 0.025$	22.8	1.04
no $\xi_4$	23	$0.470 \pm 0.026$	11.6	1.05	$0.433 \pm 0.023$	9.0	0.82
Hartlap	34	$0.464 \pm 0.023$	18.2	0.83	$0.430 \pm 0.027$	17.0	0.77
with $w_p$	45	$0.460 \pm 0.021$	37.7	1.14	$0.425 \pm 0.023$	34.9	1.06
$w_p$ -only	12	$0.402 \pm 0.097$	5.5	N/D	$0.399 \pm 0.088$	2.4	N/D
$s > 1.0 h^{-1} \text{ Mpc}$	28	$0.474 \pm 0.024$	17.0	1.06	$0.437 \pm 0.024$	12.1	0.76
$s > 2.5 h^{-1} \text{ Mpc}$	22	$0.499 \pm 0.029$	6.3	0.63	$0.446 \pm 0.024$	5.8	0.58
$s > 6.3 h^{-1} \text{ Mpc}$	16	$0.538 \pm 0.040$	3.1	0.77	$0.471 \pm 0.038$	2.5	0.62
$s < 25 h^{-1} \text{ Mpc}$	28	$0.462 \pm 0.026$	17.7	1.11	$0.443 \pm 0.025$	18.9	1.18
$s < 10 h^{-1} \text{ Mpc}$	22	$0.415 \pm 0.029$	12.3	1.23	$0.409 \pm 0.034$	11.4	1.14

the constraints. Generally, only fitting the monopole and quadrupole moments reliably predicts the hexadecapole. Thus, including  $\xi_4$  has a little impact on the cosmological constraints.

### 5.3.3 Dependence on covariance matrix

Our default analysis is performed using the covariance matrix described in Section 2.3. This covariance matrix uses a physically motivated way to correct for biases in the covariance matrix estimate due to a finite number of jackknife samples. A commonly used alternative is the Hartlap et al. (2007) correction,

$$\hat{\Phi} = \frac{N_{\text{jackknife}} - N_{\text{data}} - 2}{n_{\text{jackknife}} - 1} \Sigma^{-1}, \quad (30)$$

where  $\hat{\Phi}$  is used as the precision matrix instead of  $\Sigma^{-1}$  in equation (27). In the above equation,  $n_{\text{jackknife}} = 74$  is the number of jackknife samples and  $N_{\text{data}} = 34$  is the number of measurements.

We have repeated the entire analysis using the Hartlap et al. (2007) correction but find very similar results. For the  $z = 0.25$  and  $z = 0.40$  samples, we now find  $f\sigma_8 = 0.464 \pm 0.023$  and

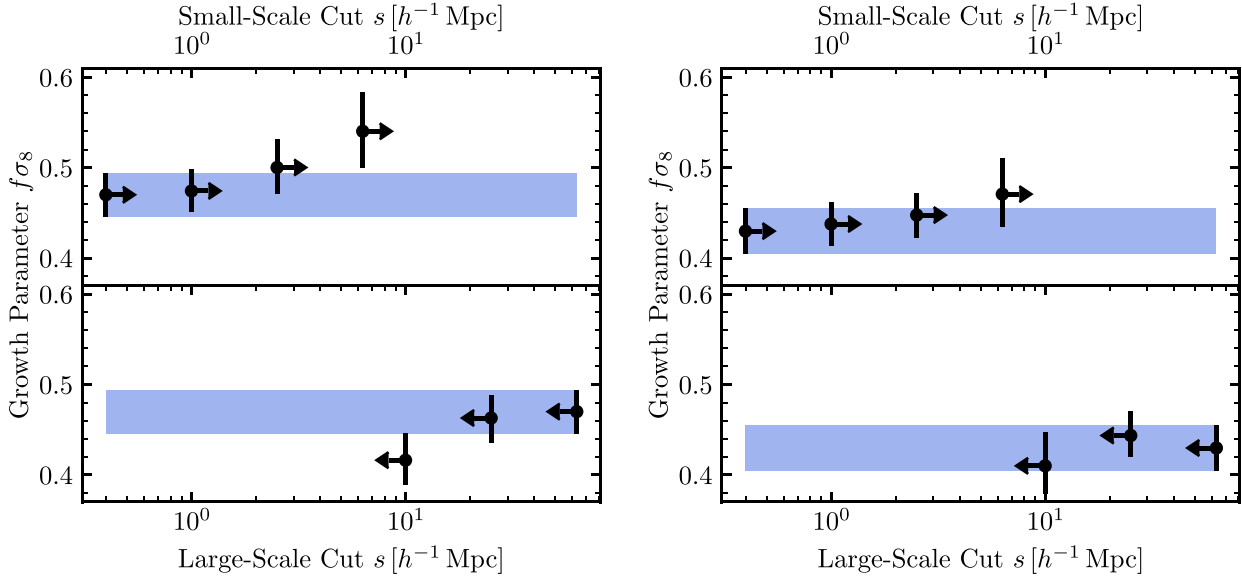
$0.430 \pm 0.027$ , respectively. Thus, our results are not significantly affected by the choice of covariance matrix. Finally, the best-fitting  $\chi^2$  values improve from 21.1 and 22.8 to 18.2 and 17.0 for the two samples, respectively.

### 5.3.4 Projected correlation function

Previous studies (e.g. Reid et al. 2014; Guo et al. 2015a; Zhai et al. 2019) of RSDs have used the so-called projected correlation function as an observable. The projected correlation function  $w_p$  is defined via

$$w_p(r_p) = \int_{-\pi_{\text{max}}}^{+\pi_{\text{max}}} \xi(r_p, \pi) d\pi, \quad (31)$$

Our rationale for not using this observables is that it mostly encodes data that are already contained within the multipoles. Additionally, building a covariance matrix becomes more difficult with increasing number of data points. None the less, we test here how the inclusion



**Figure 10.** The dependence of the cosmological constraints on the scales considered in the analysis. By default, we utilize scales within  $400 h^{-1} \text{ kpc} < s < 63 h^{-1} \text{ Mpc}$ . We show the results for the  $0.18 < z \leq 0.3$  (left) and the  $0.3 < z \leq 0.43$  (right) samples. Upper panels indicate how constraints on  $f\sigma_8$  change if the minimum scale considered is raised from its original  $s_{\min} = 400 h^{-1} \text{ kpc}$  value. Similarly, lower panels show results for reducing the maximum scale from  $63 h^{-1} \text{ Mpc}$  to a lower value. The blue rectangle in each panel indicates the default scale cuts and constraints.

of  $w_p$  affects our cosmological constraints. We use the same scales for  $r_p$  as we used for  $s$  thus far.<sup>5</sup>

For both samples, including  $w_p$  in the fit shifts the  $f\sigma_8$  constraint by less than  $0.5\sigma$ . Additionally, the inclusion of the 11 data points from the projected correlation function increases the best-fitting  $\chi^2$  by 16.6 and 12.1 for the low and the high-redshift sample, respectively. Overall, this indicates a good agreement between the redshift-space correlation as measured by the multipoles and the projected correlation function. Beyond this consistency check, we also compute  $f\sigma_8$  constraints if only  $w_p$  is used as an observable. Since  $w_p$  is designed to suppress velocity information, we expect a significant loss of sensitivity to  $f\sigma_8$ . Reassuringly, we find that uncertainties on  $f\sigma_8$  degrade by roughly a factor of 4.

#### 5.4 Galaxy–halo connection

In the following, we will present constraints on galaxy–halo connection parameters  $\mathcal{G}$  obtained after marginalizing over cosmology. Formally, the posterior constraint on  $\mathcal{G}$  can be computed via

$$P(\mathcal{G}|\mathbf{D}) = \int P(\mathcal{G}, \mathcal{C}|\mathbf{D})d\mathcal{C} = \int P(\mathcal{G}|\mathcal{C}, \mathbf{D})P(\mathcal{C}|\mathbf{D})d\mathcal{C}. \quad (32)$$

Since the simulations sample the prior space in  $\mathcal{C}$  and the posterior probability of  $\mathcal{C}$  is proportional to the cosmological evidence, an estimate for the posterior probability of  $\mathcal{G}$  marginalized over  $\mathcal{C}$  is

$$\widehat{P(\mathcal{G}|\mathbf{D})} = \sum_i \hat{\mathcal{Z}}(\mathcal{C}_i|\mathbf{D})P(\mathcal{G}|\mathcal{C}_i, \mathbf{D}), \quad (33)$$

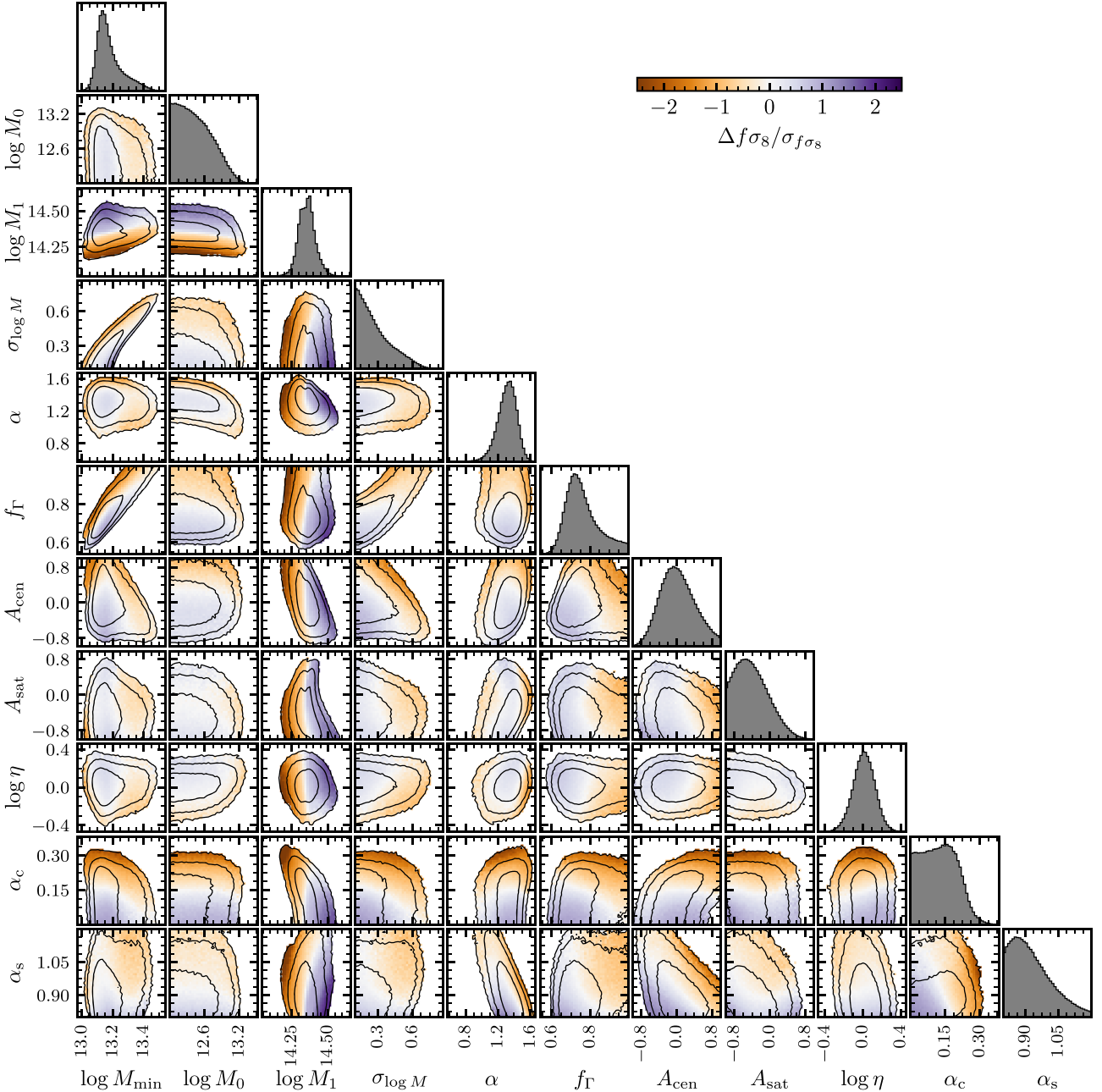
<sup>5</sup>In the following, we use the same smoothing technique described in Section 2.3 to estimate the covariance matrix. Note that for the covariance between  $w_p$  and any multipole moment, we do not exclude elements where  $r_p = s$  from the smoothing. The reason is that  $w_p(r_p)$  depends on the correlation function at all  $s$  with  $r_p \leq s \leq (r_p^2 + \pi_{\max}^2)^{1/2}$ , not just  $s = r_p$ .

**Table 3.** Posterior constraints on galaxy–halo connection parameters after marginalizing over cosmology.

Parameter	Posterior	
	$0.18 < z \leq 0.30$	$0.30 < z \leq 0.43$
$\log M_{\min}$	$13.158^{+0.105}_{-0.051}$	$13.45^{+0.17}_{-0.16}$
$\log M_0$	$12.41^{+0.36}_{-0.28}$	$12.40^{+0.38}_{-0.28}$
$\log M_1$	$14.356^{+0.054}_{-0.055}$	$14.548^{+0.176}_{-0.092}$
$\sigma_{\log M}$	$0.24^{+0.19}_{-0.10}$	$0.47^{+0.16}_{-0.21}$
$\alpha$	$1.32^{+0.10}_{-0.12}$	$1.36^{+0.35}_{-0.29}$
$f\Gamma$	$0.746^{+0.107}_{-0.064}$	$0.77^{+0.16}_{-0.18}$
$A_{\text{cen}}$	$0.00^{+0.43}_{-0.36}$	$0.39^{+0.34}_{-0.34}$
$A_{\text{sat}}$	$-0.44^{+0.43}_{-0.35}$	$-0.46^{+1.00}_{-0.41}$
$\log \eta$	$0.01^{+0.11}_{-0.11}$	$-0.21^{+0.18}_{-0.17}$
$\alpha_c$	$0.124^{+0.080}_{-0.083}$	$0.236^{+0.087}_{-0.081}$
$\alpha_s$	$0.914^{+0.113}_{-0.075}$	$0.97^{+0.14}_{-0.12}$

where the sum goes over all of the Aemulus simulations. The posterior probability of galaxy–halo connection parameters at a given cosmology,  $P(\mathcal{G}|\mathcal{C}_i, \mathbf{D})$ , is a natural by-product of the MULTINEST analysis. Note that we use the model evidence  $\hat{\mathcal{Z}}$  instead of the directly measured evidence  $\mathcal{Z}$  because the latter is more noisy, as can be seen in Fig. 9.

In Table 3 and Figs 11 and 12, we show our derived posterior constraints on the galaxy–halo connection parameters. In the following, we discuss some noteworthy findings. First, we do not find strong evidence for assembly bias, i.e.  $A_{\text{cen}} \neq 0$  or  $A_{\text{sat}} \neq 0$ . However, we only place very weak constraints, in part due to the degeneracy between assembly bias and  $f\sigma_8$  (Lange et al. 2019c). Furthermore, our findings are consistent with satellites following an NFW profile with the same concentration as dark matter, i.e.  $\eta = 1$ . Additionally, we find some evidence of central velocity bias  $\alpha_c > 0$  for the  $0.30 < z \leq 0.43$  sample but no evidence for central velocity bias for the



**Figure 11.** Posterior constraints on galaxy–halo connection parameters for the  $0.18 < z \leq 0.30$  sample after marginalization over cosmology. Panels along the diagonal show marginalized 1D posteriors. The off-diagonal panels show as contours the 68 per cent, 95 per cent, and 99 per cent posterior mass. Additionally, background colours indicate the difference between the local mean of the  $f\sigma_8$  posterior as a function of galaxy–halo connection parameter compared to the global mean of the  $f\sigma_8$  posterior, i.e. averaged over all galaxy–halo connection parameters, in units of the  $f\sigma_8$  uncertainty.

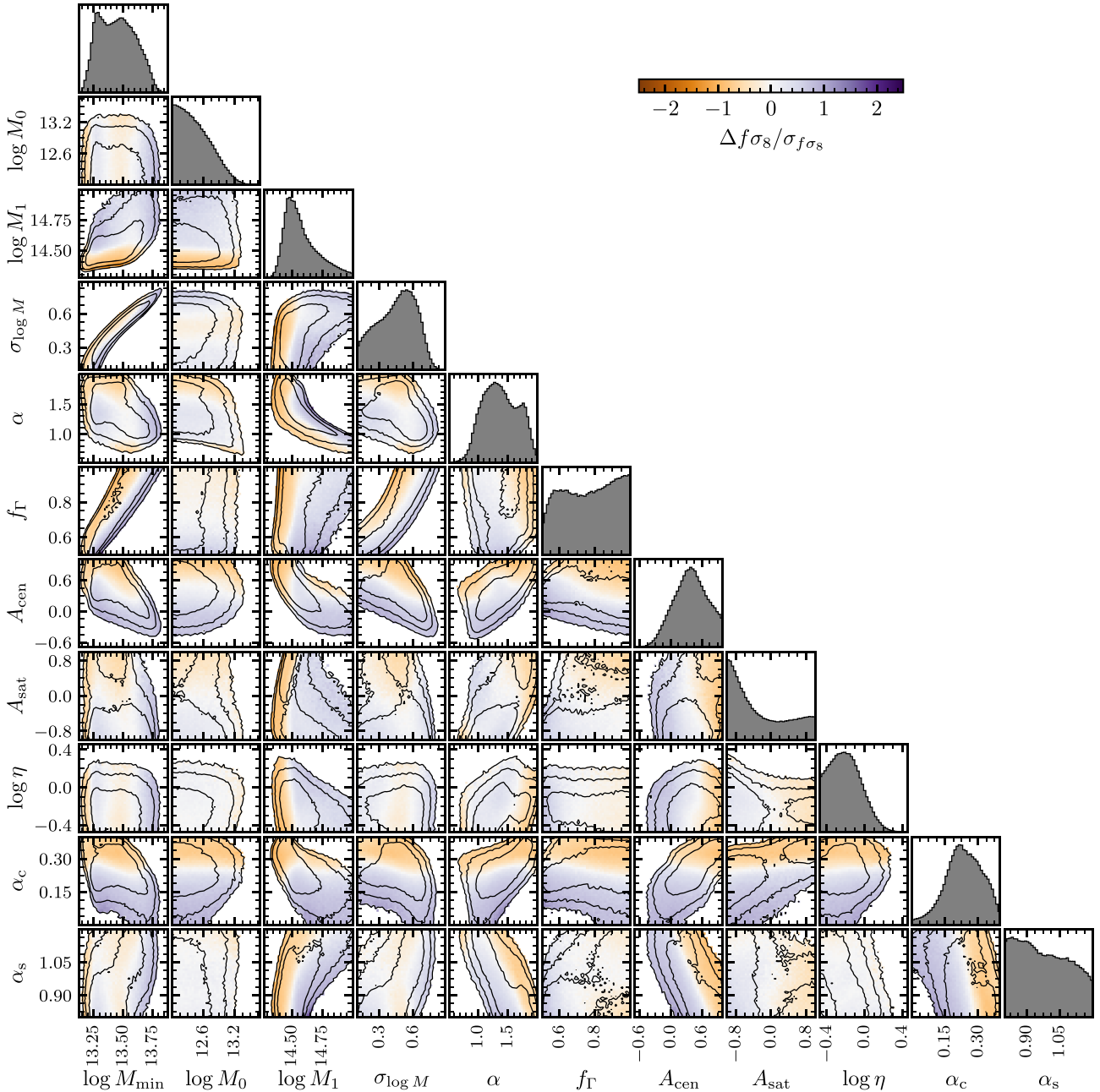
$0.18 < z \leq 0.30$  sample. Finally, our analysis is also consistent with no satellite velocity bias, i.e.  $\alpha_s = 1$ , for both samples.

Figs 11 and 12 also show the degeneracy between galaxy–halo connection parameters and  $f\sigma_8$ . At each point in the galaxy–halo connection parameter space, we calculate the average  $f\sigma_8$  of the simulation boxes contributing to the posterior mass and compare this against the  $f\sigma_8$  constraints we inferred. We find that  $\log M_1$ ,  $A_{\text{cen}}$ , and  $\alpha_c$  are most strongly degenerate with the  $f\sigma_8$  posterior. Finally, we note that some of our galaxy–halo connection parameters are limited by the prior. In many cases, the limiting priors are determined by the ranges in which the parameters can be defined, e.g.  $f_\Gamma \leq +1$ ,  $-1$

$\leq |A_{\text{cen}}| \leq +1$ ,  $-1 \leq |A_{\text{sat}}| \leq +1$  and  $\alpha_c > 0$ . Additionally,  $\log M_0$  has little impact on the HOD once  $\log M_0 < 12$  and  $\sigma_{\log M}$  describing the rate of transition from  $N_{\text{cen}} = 0$  to  $N_{\text{cen}} = 1$  is not expected to be very narrow, i.e.  $0.0 \leq \sigma_{\log M} < 0.1$ . Given that  $M_0$  and  $\sigma_{\log M}$  do not strongly correlate with  $f\sigma_8$ , it is unlikely that our prior choice impacts the  $f\sigma_8$  posterior substantially.

### 5.5 Comparison with galaxy–galaxy lensing

Under cosmological  $\Lambda$ CDM parameters compatible with the Planck Collaboration (2020) CMB analysis, Leauthaud et al. (2017) have



**Figure 12.** Similar to Fig. 11 but for the  $0.30 < z \leq 0.43$  sample.

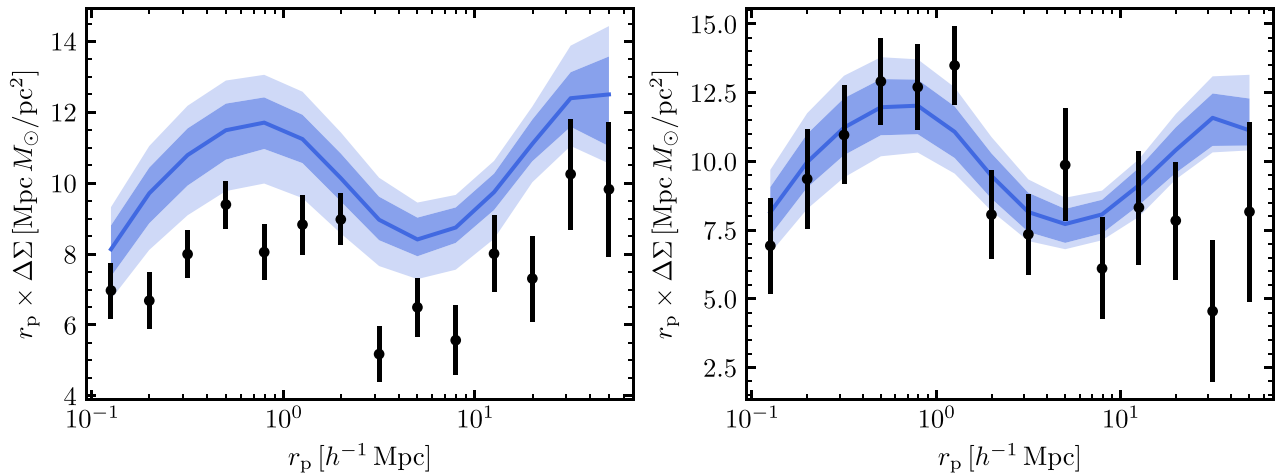
shown that several galaxy–halo connection models fitted to the clustering properties of galaxies in BOSS do not correctly predict their galaxy–galaxy lensing amplitude. The lensing amplitude, the so-called excess surface density  $\Delta\Sigma$ , is defined as

$$\Delta\Sigma(r_p) = \langle \Sigma(< r_p) \rangle - \Sigma(r_p). \quad (34)$$

In the above equation  $r_p$  is the projected comoving separation from the galaxy,  $\langle \Sigma(< r) \rangle$  the mean surface mass density (in comoving units) for separations less than  $r_p$  and  $\Sigma(r_p)$  the surface mass density at  $r_p$ . The combination of galaxy clustering and galaxy–galaxy lensing is particularly sensitive to the cosmological parameters  $\Omega_{m,0}$  and  $\sigma_8(z)$ ; see e.g. More et al. 2013; Wibking et al. 2019). Thus, a mismatch in the lensing prediction after fitting the clustering and assuming Planck Collaboration (2020) parameters could indicate a

tension in these two parameters with respect to the CMB under  $\Lambda$ CDM. At the same time, especially on small scales, care must be taken because other effects like baryonic feedback (Leauthaud et al. 2017; Lange et al. 2019b) or galaxy assembly bias (Lange et al. 2019b; Yuan & Eisenstein 2019; Yuan, Eisenstein & Leauthaud 2020) can affect the lensing prediction. Similarly, modifications to general relativity could also modify the expected lensing amplitude (Leauthaud et al. 2017).

We compute  $\Delta\Sigma$  for a given mock catalogue using a random 0.5 per cent subset of all simulation particles and the `mean_delta_sigma` routine in HALOTOOLS. In Fig. 13, we show the marginalized posterior predictions for the galaxy–galaxy lensing amplitudes of the two samples. The marginalization over cosmology is performed analogously to equation (33). We also show



**Figure 13.** Posterior predictions for the galaxy–galaxy lensing amplitude of galaxies in the  $0.18 < z \leq 0.30$  (left) and the  $0.30 < z \leq 0.43$  sample (right). Bands denote the 68 per cent and 95 per cent posterior mass. Measurements from cross-correlating BOSS LOWZ targets with shape catalogues from SDSS are shown by error bars. The lensing amplitude is over-predicted for the low-redshift sample. Our prediction agrees much better with the measurements for the high-redshift sample but in this redshift range the inferred  $f\sigma_8$  from RSDs is 10 per cent lower than the Planck Collaboration (2020)  $\Lambda$ CDM prediction at  $2\sigma$  significance.

measurements of the galaxy–galaxy lensing amplitude from cross-correlating BOSS LOWZ targets with sources from SDSS. We refer the reader to Singh et al. (2020) regarding details of the SDSS lensing measurements. Note that there is an overall 6 per cent uncertainty on the total normalization of the lensing amplitude, stemming from photometric redshift uncertainties in SDSS.

In agreement with previous studies (Leauthaud et al. 2017; Lange et al. 2019b, 2021; Yuan et al. 2020), we find that the lensing amplitude tends to be over-predicted. The over-prediction is significant for the low-redshift sample whilst our prediction agrees much better with the measurements for the high-redshift sample. However, for the higher redshift sample, the inferred  $f\sigma_8$  is 10 per cent from RSDs is lower than the Planck Collaboration (2020)  $\Lambda$ CDM prediction at  $2\sigma$  significance.

The novelty of the result here is that the lensing prediction is marginalized over cosmology, whereas the aforementioned studies showed the mismatch under  $\Lambda$ CDM parameters of the Planck Collaboration (2020) CMB analysis. We note, however, that our marginalization over cosmology implicitly includes the cosmological priors of the Aemulus simulations (DeRose et al. 2019). Despite this additional freedom in cosmological parameters, we still find that our model for the low-redshift sample fails to correctly predict the galaxy–galaxy lensing amplitude. Finally, it is worth remembering that our lensing predictions, which are based on dark matter-only simulations, do not account for the impact of baryons on the matter distribution (Leauthaud et al. 2017; Lange et al. 2019b; Amodeo et al. 2021). Furthermore, Yuan et al. (2021) show that overdensity-based assembly bias can further reduce this lensing tension, even when modelled in addition to concentration or  $V_{\text{max}}$ -based assembly bias. Both effects, baryonic feedback and more complex assembly bias models, could potentially reconcile or at least alleviate the tension on small scales,  $r_p \lesssim 3 h^{-1} \text{ Mpc}$ , but are unlikely to impact differences on larger scales.

## 6 DISCUSSION

### 6.1 Constraints on the growth rate

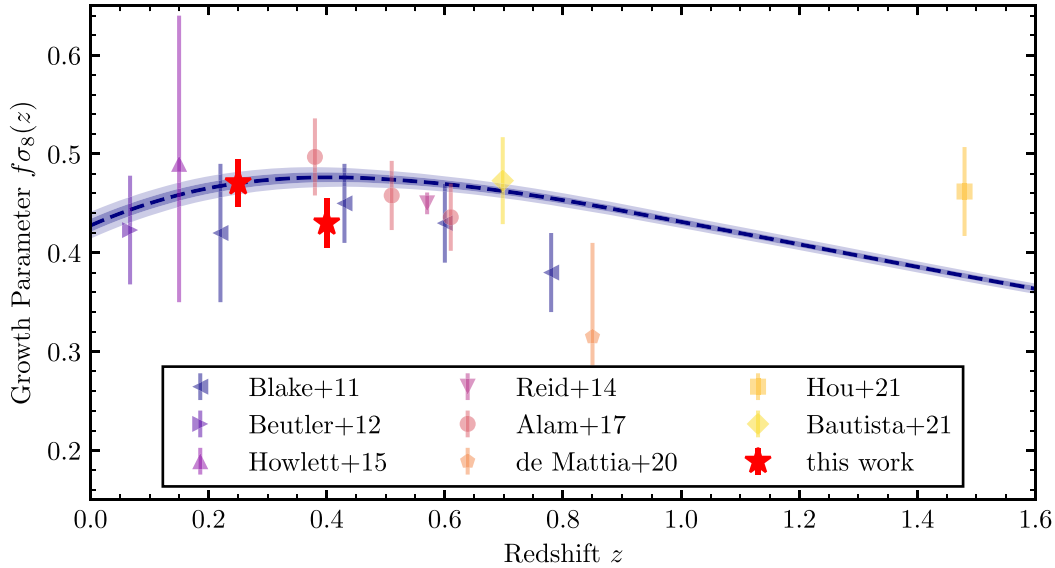
In Fig. 14, we compare our constraints on  $f\sigma_8(z)$  against other results from the literature and the  $\Lambda$ CDM Planck Collaboration (2020)

CMB prediction. The literature comparison is chosen to represent the leading results on large scales from the 6dF Galaxy Survey (Beutler et al. 2012), the WiggleZ survey (Blake et al. 2011), SDSS main galaxy sample (Howlett et al. 2015), BOSS (Alam et al. 2017), and eBOSS (Bautista et al. 2021; de Mattia et al. 2021; Hou et al. 2021). Our results are consistent with other results in the literature, as well as CMB  $\Lambda$ CDM predictions. At the same time, our results follow the trend of low-redshift measurements of  $f\sigma_8$  falling slightly below the  $\Lambda$ CDM + CMB prediction. Future work is needed to verify whether this difference is statistically significant. We also see that our constraints on  $f\sigma_8$  are roughly a factor of two more stringent than any other study on large scales, demonstrating the potentially large benefits of extending the analysis to the full range of scales accessible with observations.

In Fig. 14, we also include the full-scale BOSS RSD analysis of Reid et al. (2014). However, the analysis of Reid et al. (2014) is based on the assumption that changes in  $f\sigma_8$  are completely degenerate with a simple linear scaling of the velocity field on all scales. Zhai et al. (2019) argue that this approximation might lead to errors and cosmological parameter constraints that are artificially narrow. Our work is not based on this assumption and should therefore be more robust. We note, however, that the analysis of Reid et al. (2014) is based on  $\sim 500\,000$  galaxies in the BOSS CMASS sample. Given this roughly seven times larger galaxy sample, we estimate that if we applied our analysis framework to the same galaxy sample, we would obtain a roughly similar precision on  $f\sigma_8$ . Notwithstanding questions about the validity of the scaling approximation in Reid et al. (2014), it is encouraging to see that the non-linear analysis of Reid et al. (2014) shows results broadly comparable with ours:  $f\sigma_8$  is compatible but slightly lower than the Planck Collaboration (2020)  $\Lambda$ CDM prediction.

### 6.2 Consistency between small and large scales

In Fig. 10, we showed how constraints on  $f\sigma_8$  vary when excluding small or large scales  $s$  from the analysis. Generally, we find consistent results between small and large scales, but there are also  $1\text{--}2.5\sigma$  hints of small scales ( $< 10 h^{-1} \text{ Mpc}$ ) preferring lower values for  $f\sigma_8$  than larger scales, especially for the low-redshift sample. When



**Figure 14.** Comparison of different constraints on the growth rate of structure. We show results from the Planck Collaboration (2020) CMB analysis assuming a  $\Lambda$ CDM cosmology (blue). The bands denote the  $1\sigma$  and  $2\sigma$  ranges. Additionally, we show large-scale structure constraints from the WiggleZ survey (Blake et al. 2011), the 6dF Galaxy Survey (Beutler et al. 2012), and the SDSS main galaxy sample (Howlett et al. 2015), BOSS (Alam et al. 2017), and eBOSS (Bautista et al. 2021; de Mattia et al. 2021; Hou et al. 2021). We also compare our results (red) against the small-scale BOSS CMASS RSD analysis of Reid et al. (2014). When necessary, some results are slightly shifted in the  $x$ -direction for clarity.

analysing mock catalogues, as described in Section 4, we did not find comparable trends of a strong scale dependence. There are a number of possibilities for the findings in the data. First, it could be a real trend in the data driven by some aspect of galaxy formation and/or baryonic physics. Future mock tests using more complex galaxy models than SHAM, ideally based on hydrodynamical simulations, are needed to investigate this possibility further. Secondly, it could be related to an effect beyond the traditional  $\Lambda$ CDM and  $w$ CDM models. For example, whilst modifications to general relativity like  $f(R)$  gravity induce consistent changes to  $f\sigma_8$  on large, linear scales, the impact has a strong scale dependence in the non-linear regime (Fontanot et al. 2013; He et al. 2018; Alam et al. 2020). Thus, scale dependent  $f\sigma_8$  constraints under  $w$ CDM might, in fact, be a sign of modifications to general relativity. Thirdly, these trends could also be simply due to random statistical fluctuations. Upcoming surveys, especially the *Dark Energy Spectroscopic Instrument* (DESI) survey, will soon provide much larger samples of LRGs at similar redshifts. If the trends we see in the data are not due to random statistical fluctuations, higher signal-to-noise DESI data should easily confirm them with high significance.

### 6.3 Constraining power as a function of scale

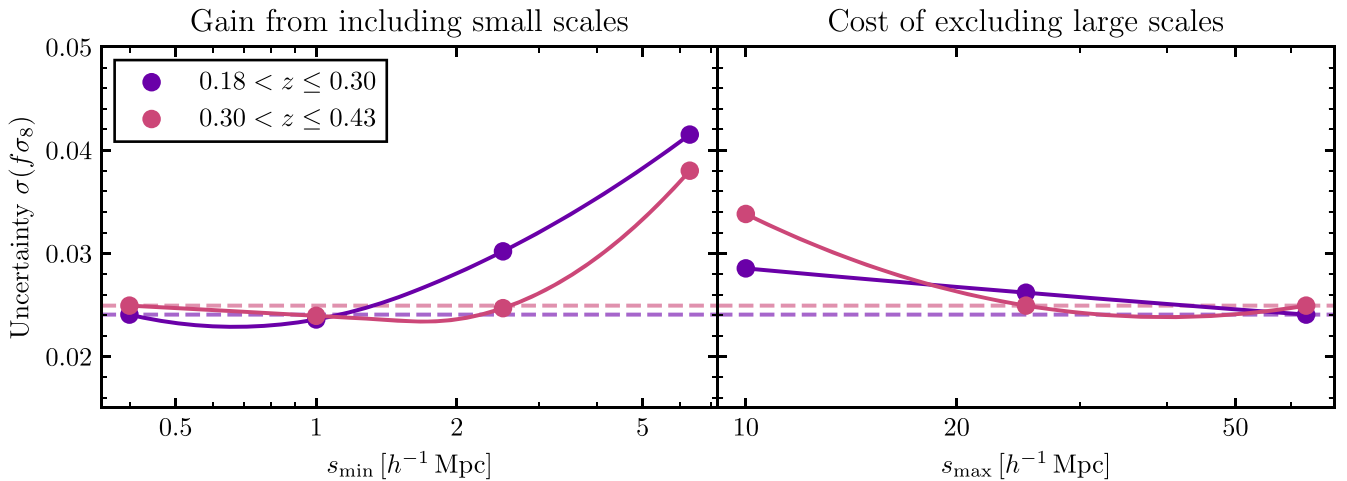
The ability to harness the cosmological constraining power of the nonlinear regime is the driving motivation behind our development of the CEM technique, as well as our simulation-based forward modelling methodology. In Fig. 15, we highlight the benefits of our analysis method, by showing how the precision of our cosmological constraints varies as a function of scales analyzed. In both panels, our fiducial analysis is conducted over scales  $s$  in the range  $400 h^{-1} \text{ kpc} < s < 63 h^{-1} \text{ Mpc}$ . In the left-hand panel of Fig. 15, the horizontal axis shows the minimum scale used in the analysis,  $s_{\min}$ ; each point in the figure shows the results of an RSD analysis that includes information restricted to  $s_{\min} < s < 63 h^{-1} \text{ Mpc}$ , with the  $y$ -axis showing the strength of the constraints on  $f\sigma_8$  derived in each case. Thus, larger values of  $s_{\min}$  correspond to analyses which

exclude more information from small scales, and so both curves in the left-hand panel naturally increase with  $s_{\min}$ .

There are two striking features about the left-hand panel of Fig. 15. First, the constraining power improves dramatically as deeply nonlinear scales  $s < 10 h^{-1} \text{ Mpc}$  are included in the analysis. For example, an analysis with  $s_{\min} = 2.5 h^{-1} \text{ Mpc}$  has a 50 per cent higher constraining power on  $f\sigma_8$  compared to an analysis with  $s_{\min} = 6 h^{-1} \text{ Mpc}$ . Secondly, the inclusion of scales smaller than  $s \lesssim 1 h^{-1} \text{ Mpc}$  provides negligible information on  $f\sigma_8$  that is not already contained in larger-scale modes. This is true even though the posteriors on some HOD parameters continue to improve as scales  $s_{\min} < 1 h^{-1} \text{ Mpc}$  are included in the analysis. This indicates that when trying to improve constraining power on  $f\sigma_8$  by including smaller scales in an RSD-multipole analysis of LOWZ galaxies, we reach a point of diminishing returns at  $s_{\min} \lesssim 1 h^{-1} \text{ Mpc}$ , since the bulk of the degeneracies between  $f\sigma_8$  and the galaxy–halo connection have already been broken.

Whereas the left-hand panel of Fig. 15 illustrates the gain in cosmological constraining power we have achieved by including small-scale information in our analysis, the right-hand panel shows the loss in constraining power that would be incurred by discarding measurements on larger spatial scales. The vertical axis in the right-hand panel shows the strength of the constraints on  $f\sigma_8$  derived by an analysis restricted to  $400 h^{-1} \text{ kpc} < s < s_{\max}$ , plotted as a function of  $s_{\max}$  on the horizontal axis. Thus larger values of the horizontal axis correspond to RSD analyses that include more information from large scales, and so each curve decreases monotonically with  $s_{\max}$ .

The salient feature of the right-hand panel of Fig. 15 is the shallow slope of each curve. For example, an RSD analysis over the range of scales  $s \lesssim 30 h^{-1} \text{ Mpc}$  has practically the same constraining power as an analysis with  $s \lesssim 60 h^{-1} \text{ Mpc}$ , and an analysis limited to  $s_{\max} \lesssim 10 h^{-1} \text{ Mpc}$  only suffers a 50 per cent loss in precision relative to our fiducial all-scale analysis. Considering the left- and right-hand panels of Fig. 15 together, we conclude that for RSD analyses of galaxies in the BOSS LOWZ sample, most of the multipole information about  $f\sigma_8$  is contained in the scales  $2 h^{-1} \text{ Mpc} \lesssim s \lesssim 20 h^{-1} \text{ Mpc}$  (for



**Figure 15.** Similar to Fig. 10 but showing how the constraints on  $f\sigma_8$  depend on scale cuts. In the left-hand panel, we show the dependence when the small-scale cut is increased compared to its default value of  $s_{\min} = 400 h^{-1} \text{ kpc}$ . Similarly, the right-hand panel shows the change in constraining power if the large-scale cut is reduced to values lower than  $s_{\max} = 63 h^{-1} \text{ Mpc}$ .

similar conclusions based on a forecasting analysis of cluster-galaxy cross-correlations, see Salcedo et al. 2020). Evidently, once the cosmological information content of the quasi-to-nonlinear regime has been harvested, large-scale modes contain only very modest additional information about structure growth.

This observation has important implications for the computational demands of RSD analyses. Covariance matrix estimation of cosmological observables requires a large number of independently simulated volumes, which even for present-day analyses typically number in the hundreds or thousands in order to sufficiently sample the variance of the largest-scale modes measured in the analysis (e.g. Hartlap et al. 2007). Discarding measurements on larger scales and restricting attention to  $s_{\max} \lesssim 30 h^{-1} \text{ Mpc}$  only mildly degrades the constraints, but reduces the cosmological volume of the simulations needed to generate independent realizations of the predicted data vector. This benefit may prove to be particularly important in the coming decade of cosmological analyses, since the number of independent realizations required by covariance matrix estimation increases sharply with the dimension of the predicted data vector (Taylor et al. 2013).

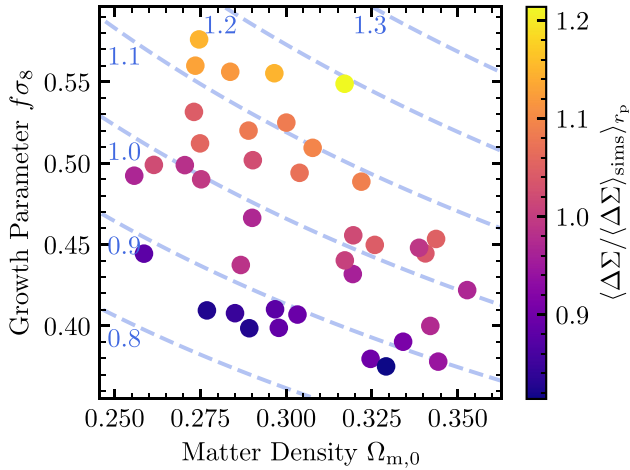
#### 6.4 Galaxy–halo connection

In addition to the cosmological growth rate, our analysis also puts constraints on the galaxy–halo connection, marginalized over cosmology. For example, within statistical uncertainties, we find no strong evidence for galaxy assembly bias, i.e. our results are compatible with  $A_{\text{cen}} = A_{\text{sat}} = 0$ . However, our constraints are very broad, partially due to the degeneracy between galaxy assembly bias parameters and  $f\sigma_8$  (Lange et al. 2019c). That we do not find evidence for assembly bias from galaxy clustering is in contrast to previous studies by Zentner et al. (2019) and Yuan et al. (2021). The former study finds evidence for galaxy assembly bias by fitting  $w_p$  in the SDSS main galaxy sample and the latter study by analysing the anisotropic two-point correlation function in BOSS CMASS. We note that both galaxy samples are different from the BOSS LOWZ sample. In principle, there can be a galaxy assembly bias in both the SDSS and BOSS CMASS samples whilst being absent in the BOSS LOWZ sample. Similarly, that both studies assume a fixed cosmology might also contribute to

their claimed detection of galaxy assembly bias. However, Yuan et al. (2021) analyse a total of  $\sim 600\,000$  galaxies, an order of magnitude larger than the two samples we analyse. Thus, it is well possible that due to the high precision clustering measurements their detection of galaxy assembly bias is robust to marginalization over cosmology.

In terms of central galaxy velocity bias, we find some evidence for a positive bias in the high-redshift sample,  $\alpha_c = 0.236^{+0.087}_{-0.081}$ . This corresponds to a scenario in which the central galaxy has an additional velocity scatter with respect to the halo core. However, our data cannot firmly rule out  $\alpha_c = 0$ . In contrast, the low-redshift sample prefers such a no-velocity-bias scenario. Given that the velocity offset is expressed with respect to the halo core, our results of a small to non-existent central velocity bias are in good agreement with other observational studies (Reid et al. 2014; Guo et al. 2016) and theoretical predictions by Ye et al. (2017). Unfortunately, a direct comparison with Guo et al. (2015a, b) is not possible because in these studies the central velocity bias is expressed with respect to the average velocity of the inner 25 per cent and 100 per cent of dark matter particles, respectively. Generally, this leads to higher inferred values for  $\alpha_c$  (Ye et al. 2017). When expressed with respect to the halo core, the results of the theoretical study of Ye et al. (2017) indicate  $\alpha_c \sim 0.05\text{--}0.10$  for galaxies with host halo masses  $10^{13}\text{--}10^{14} h^{-1} M_\odot$ , the halo mass range we are probing.

Additionally, our results in both the low- and high-redshift samples are compatible with satellites tracing dark matter spatially, i.e.  $\eta = 1$ , and having no velocity bias with respect to the prediction from Jeans equilibrium, i.e.  $\alpha_s = 1$ . The result on the spatial bias depends on the stellar mass of the satellite galaxy (Lange et al. 2021), making a direct comparison with other works difficult. In the low-redshift, low-stellar mass SDSS main galaxy sample, the data prefer satellites being less concentrated than dark matter (see e.g. Lange et al. 2019a). However, the situation is less clear for the more massive galaxies in the BOSS CMASS sample where previous studies have obtained good fits to the data without the need for  $\eta \neq 1$  (Reid et al. 2014; Guo et al. 2015a). Finally, we obtain no strong constraint on the satellite velocity bias parameter  $\alpha_s$ , contrary to Guo et al. (2015a). This can be largely attributed to a degeneracy with cosmology that was not accounted for in Guo et al. (2015a) and the smaller galaxy samples analysed here.



**Figure 16.** The dependence of the lensing amplitude on cosmological parameter. Results are shown for the  $0.18 < z \leq 0.30$  sample. Each dot corresponds to one of the Aemulus simulations. The colour indicates the scale-averaged ratio of the mean posterior lensing amplitude divided by the lensing prediction of marginalization over cosmology, i.e. the blue line in the left-hand panel of Fig. 13. Blue lines show a simple power-law fit to the dependence of this ratio on  $\Omega_{m,0}$  and  $f\sigma_8$ .

Both LOWZ samples target LRGs of comparable number densities and redshifts. Thus, their inferred galaxy–halo connection parameters should be similar but not necessarily identical. We can use this as another consistency check of our results. For example, the high-redshift sample has a lower number density than the low-redshift sample. Thus, we expect the average halo mass to host galaxies to increase, i.e. we expect  $M_{\min}$  and  $M_1$  to be larger for the high-redshift sample. Indeed, this is what we infer, as shown in Table 3. Besides these two parameters that are very directly determined by the number density of the sample, all other parameters should be similar between the redshift bins. Reassuringly, as shown in Table 3, to within statistical uncertainties these other parameters are identical between the two redshift bins. Finally, the high-redshift sample targets more luminous galaxies. Because centrals are most often the brightest galaxies in their respective haloes, we expect the satellite fraction  $f_{\text{sat}}$  to decrease with the luminosity of the sample. Our results of  $f_{\text{sat}} = 14.8 \pm 1.5$  per cent for  $z \sim 0.25$  and  $f_{\text{sat}} = 11.5 \pm 2.9$  per cent for  $z \sim 0.4$  are consistent with this expectation.

Finally, we also find that certain galaxy–halo connection parameters are degenerate with the  $f\sigma_8$  constraints. The strongest correlations are found for  $M_1$  that regulates the number of satellites per halo,  $\alpha_c$  the amount of central velocity bias and  $A_{\text{cen}}$  the central velocity bias parameter. This implies that observations providing independent constraints on these parameters could tighten our cosmological constraints in the future.

### 6.5 Lensing is low

Given Fig. 13, it is interesting to ask what cosmological parameters would alleviate the lensing tension for the low-redshift sample. The galaxy–galaxy lensing amplitude is well-known to be a function of  $\Omega_{m,0}$  and  $\sigma_8(z)$ . Thus, in Fig. 16, we show how the lensing amplitudes from the cosmology-dependent best-fitting galaxy–halo models scale with  $f\sigma_8$  and  $\Omega_{m,0}$ . A simple linear fit reveals that the scale-average lensing amplitude roughly scales with  $\Omega_{m,0}^{0.6} (f\sigma_8)^{0.9}$ . This relation helps explain why the high-redshift sample that predicts a 10 per cent lower  $f\sigma_8$  than the CMB  $\Lambda$ CDM analysis does not

overpredict the lensing amplitude, whilst the low-redshift sample with an inferred  $f\sigma_8$  matching the  $\Lambda$ CDM CMB prediction does. In order to fit the lensing amplitude of the  $0.18 < z \leq 0.30$  sample, one would likely have to assume a cosmology with a lower matter density or less growth of structure than the best-fitting Planck Collaboration (2020)  $\Lambda$ CDM values. Note that whilst both lensing and RSDs are sensitive to  $\sigma_8(z)$ , lensing additionally depends on  $\Omega_{m,0}$  whilst RSDs has a extra dependence on the growth rate  $f$  which is determined by both  $\Omega_{m,0}$  and the dark energy equation of state parameter  $w_0$ . Thus, a combined lensing and clustering analysis will be sensitive to deviations from  $\Lambda$ CDM for which  $w_0 = -1$ . Similarly, gravitational lensing is sensitive to the impact of gravity on relativistic particles whereas RSDs measure its impact on non-relativistic particles. Thus, a lensing plus redshift-space clustering study would also be sensitive to deviations from general relativity. We leave a careful combined cosmological analysis of RSDs and galaxy–galaxy lensing to future work. Such a combined RSD plus lensing analysis would also have to marginalize over complex galaxy assembly bias models and the effect of baryonic feedback.

Finally, Zu (2020) recently proposed that the lensing amplitude mismatch can be solved on small scales,  $r_p < 1 h^{-1}$  Mpc, by more complex galaxy–halo models. Amongst others, the author suggests to assume sub-Poisson distributions for the number of satellites per dark matter halo. This works by suppressing the one-halo galaxy–galaxy clustering terms, particularly the *Fingers of God* effect at fixed HOD. In turn, this allows more satellites for a given clustering amplitude. Since satellites have a lower lensing amplitude at fixed host halo mass than centrals, assuming a sub-Poisson satellite distribution leads to lower predicted lensing amplitudes. However, Zu (2020) only considered galaxy–halo models constrained by the projected correlation function  $w_p$ , not the redshift–space correlation function which tightly constrains BOSS satellite fractions. We show in Appendix C that the change in the posterior satellite fraction is only at the level of a few per cent, even when assuming a maximally sub-Poisson satellite distribution. In turn, this would only lower the lensing amplitude by the same level. Thus, non-Poisson numbers for satellites alone are unlikely to solve the ‘lensing is low’ tension on small scales.

## 7 SUMMARY AND CONCLUSION

In this work, we perform the first cosmological RSD analysis of BOSS data using simulations with varying cosmology. The simulation-based modelling approach based on the Aemulus suite (DeRose et al. 2019) allows us to perform the analysis on scales from  $0.4 h^{-1}$  Mpc to  $63 h^{-1}$  Mpc, significantly extending what is possible with a purely analytic modelling framework. Our analysis uses a sophisticated HOD modelling approach that accounts for galaxy assembly bias and velocity bias. The large range in scales also allows us to place some of the tightest constraints on the growth parameter  $f\sigma_8$  recorded to date, even after marginalizing over all galaxy–halo connection parameters.

We first test the recovery of  $f\sigma_8$  on mock catalogues constructed from the SHAM model of Lehmann et al. (2017). These mocks contain many of the complexities one would expect from realistic galaxy populations and can alter galaxy clustering substantially, including galaxy assembly bias (Zentner, Hearin & van den Bosch 2014; McCarthy, Zheng & Guo 2019; Padilla et al. 2019), ellipsoidal satellite populations and halo alignment (van Daalen, Angulo & White 2012), satellite velocity bias (Guo et al. 2015a; Ye et al. 2017) as well as non-Poisson satellite numbers (Jiang & van den Bosch 2017; Jiménez et al. 2019). The mock catalogues are constructed

from a  $4 (h^{-1} \text{ Gpc})^3$  volume and projected on to the three simulation axes. Thus, the mock measurements have a substantially higher signal-to-noise ratio than the observations which come from a single projection of a volume of  $< 1 (h^{-1} \text{ Gpc})^3$ . Despite the complexity in the mock catalogues and the substantially higher precision of the mock measurements compared to observations, we find that the growth rate parameters are accurately recovered to within statistical uncertainties.

We then apply our analysis framework to two roughly volume-limited samples of LRGs at redshifts  $z \sim 0.25$  and  $z \sim 0.4$ . We infer  $f\sigma_8 = 0.471 \pm 0.024$  and  $0.430 \pm 0.025$  for the low- and high-redshift sample, respectively. The corresponding predictions from the Planck Collaboration (2020) CMB analysis under  $\Lambda$ CDM cosmology are  $0.470 \pm 0.006$  and  $0.476 \pm 0.005$ , respectively. Thus, the low-redshift sample agrees very well with the CMB predictions, whereas the high-redshift sample falls  $\sim 10$  per cent lower at  $\sim 2\sigma$  significance. Overall, whilst our RSD results follow the trend of low-redshift observations suggesting less structure growth than  $\Lambda$ CDM CMB predictions, when considering both samples together, our RSD results are still in good agreement with the Planck Collaboration (2020)  $\Lambda$ CDM forecasts. We show that most of the cosmological constraining power of the analysis comes from scales  $2 h^{-1} \text{ Mpc} \lesssim s \lesssim 20 h^{-1} \text{ Mpc}$ , i.e. from substantially smaller scales than those commonly analysed in cosmological large-scale structure studies. Thus, despite only analysing a small fraction of the BOSS galaxy sample, our constraints on  $f\sigma_8$  are nominally more stringent than those of any analysis focusing on large scales, i.e.  $s \gtrsim 30 h^{-1} \text{ Mpc}$  only. Overall, in agreement with previous works (Cacciato et al. 2013; Reid et al. 2014; Lange et al. 2019c; Wibking et al. 2019; Zhai et al. 2019), our study highlights the potential benefits of extending the cosmological analysis into the non-linear regime.

We also study the consistency of small- and large-scale constraints. We find  $1-2.5\sigma$  hints that our analysis prefers smaller  $f\sigma_8$  when analysing small scales only ( $s \lesssim 10 h^{-1} \text{ Mpc}$ ) compared to large scales only ( $s \gtrsim 6 h^{-1} \text{ Mpc}$ ). The difference is the strongest for the low-redshift sample, reaching roughly  $2.5\sigma$  significance. However, our model is able to fit observations on all scales well with  $\chi^2_\nu \approx 1$  and we do not observe a similar scale dependence in the much more accurate mock observations. New data from DESI will yield more constraining measurements and will be able to determine if the difference in  $f\sigma_8$  constraints is a statistical fluke, a sign of modelling systematics, or some form of new physics. We stress the importance of conducting more mock tests in the future to test the veracity of cosmological constraints from all-scale modelling analyses, such as the one performed here. Ideally, the mock tests would be based on hydrodynamical simulations or semi-analytic models of galaxy formation. Such tests will be critical for upcoming studies with higher precision measurements. Another potential source of systematic error is the correlation between galaxy orientation, the local tidal field, and the large-scale velocity field (Hirata 2009; Obuljen, Percival & Dalal 2020; Singh, Yu & Seljak 2021). If target selection was directly or indirectly based on galaxy orientation, this can lead to biases in the derived values of  $f\sigma_8$ . We do not expect this to be a significant issue in the present analysis. However, this effect will become more important for the DESI survey where target selection is based on fibre magnitude as such a selection indirectly depends on orientation.

In the future, we plan to include galaxy–galaxy lensing in our all-scale cosmological analysis. In this work, we have shown how our predictions for the galaxy–galaxy lensing amplitude compares against measurements from SDSS (Singh et al. 2020). We find that the lensing amplitude is underpredicted for the low-redshift

sample and in good agreement for the high-redshift sample. As expected, the lensing amplitude at fixed clustering is positively correlated  $\Omega_{m,0}$  and  $\sigma_8(z)$ . This likely contributes to the high-redshift sample with lower implied  $f\sigma_8$  showing a better agreement with the measured lensing amplitude. Overall, a combined lensing and redshift-space clustering analysis would likely find evidence for less growth of structure or lower matter density  $\Omega_{m,0}$  than implied by the CMB plus  $\Lambda$ CDM forecast. However, a full cosmological interpretation of the lensing amplitude would involve marginalizing over baryonic feedback (Leauthaud et al. 2017; Lange et al. 2019c), as well as more tests on mock catalogues. Additionally, one could combine galaxy–galaxy lensing measurements with observations of the Sunyaev-Zeldovich effect to directly constraint the strength of baryonic feedback (Schaan et al. 2021; Amodeo et al. 2021). We leave such a detailed study to future work but note that it could provide further interesting measurements on the growth of structure in the late Universe, possible violations to general relativity and the validity of the  $\Lambda$ CDM model.

## ACKNOWLEDGEMENTS

We thank the Aemulus collaboration for making their simulations publicly available and for the letting us take part in the Aemulus mock challenge that motivated the mock tests presented in this study. We also thank Sukhdeep Singh for providing the SDSS galaxy–galaxy lensing measurements, Sihang Yuan for useful conversations on RSD modelling, as well as Kuan Wang and Risa Wechsler for commenting on earlier versions of this manuscript. Finally, we also thank the anonymous referee for the numerous insightful comments which improved the presentation of this paper.

We acknowledge use of the lux supercomputer at UC Santa Cruz, funded by NSF MRI grant AST 1828315. This material is based on work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0019301. AL acknowledges support from the David and Lucille Packard foundation, and from the Alfred P. Sloan foundation. HG acknowledges the support from the National Natural Science Foundation of China (Nos. 11833005, 11922305). Work done by APH was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under contract DE-AC02-06CH11357.

This work made use of the following software packages: MATPLOTLIB (Hunter 2007), SCIPY, NUMPY (van der Walt, Colbert & Varoquaux 2011), ASTROPY (Astropy Collaboration 2013), COLUSSUS (Diemer 2015), HALOTOOLS (Hearin et al. 2017), MULTINEST (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2019), PYMULTINEST (Buchner et al. 2014), SCIKIT-LEARN (Pedregosa et al. 2012), SPYDER and GNOME LATEX.

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German

Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## DATA AVAILABILITY

The Aemulus and UNIT simulations used in this article are publicly available at <https://aemulusproject.github.io/> and <http://www.unitsims.org/>, respectively. The SDSS data analysed is available at <https://www.sdss.org/>. All derived data generated in this research, as well as code used will be shared on reasonable request to the corresponding author.

## REFERENCES

- Abbott T. M. C. et al., 2018, *Phys. Rev. D*, 98, 043526
- Ahumada R. et al., 2020, *ApJS*, 249, 3
- Alam S. et al., 2017, *MNRAS*, 470, 2617
- Alam S. et al., 2020, preprint ([arXiv:2011.05771](https://arxiv.org/abs/2011.05771))
- Alcock C., Paczynski B., 1979, *Nature*, 281, 358
- Amodeo S. et al., 2021, *Phys. Rev. D*, 103, 063514
- Anderson L. et al., 2014, *MNRAS*, 441, 24
- Asgari M. et al., 2021, *A&A*, 645, A104
- Astropy Collaboration, 2013, *A&A*, 558, A33
- Azzalini A., Valle A. D., 1996, *Biometrika*, 83, 715
- Bautista J. E. et al., 2021, *MNRAS*, 500, 736
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
- Beltz-Mohrmann G. D., Berlind A. A., Szewciw A. O., 2020, *MNRAS*, 491, 5771
- Berlind A. A., Weinberg D. H., 2002, *ApJ*, 575, 587
- Beutler F. et al., 2012, *MNRAS*, 423, 3430
- Blake C. et al., 2011, *MNRAS*, 415, 2876
- Blake C. et al., 2013, *MNRAS*, 436, 3089
- Blinov N., Kelly K. J., Krnjaic G., McDermott S. D., 2019, *Phys. Rev. Lett.*, 123, 191102
- Buchner J. et al., 2014, *A&A*, 564, A125
- Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, *MNRAS*, 430, 767
- Carlson J., Reid B., White M., 2013, *MNRAS*, 429, 1674
- Chisari N. E. et al., 2018, *MNRAS*, 480, 3962
- Chuang C.-H. et al., 2019, *MNRAS*, 487, 48
- Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, 647, 201
- D'Amico G., Senatore L., Zhang P., Zheng H., 2021, *J. Cosmol. Astropart. Phys.*, 2021, 072
- de Mattia A. et al., 2021, *MNRAS*, 501, 5616
- DeRose J. et al., 2019, *ApJ*, 875, 69
- Desjacques V., Jeong D., Schmidt F., 2018, *Phys. Rep.*, 733, 1
- Di Valentino E., Melchiorri A., Silk J., 2020, *Nat. Astron.*, 4, 196
- Diemer B., 2015, Astrophysics Source Code Library, record ascl:1501.016
- Eisenstein D. J. et al., 2001, *AJ*, 122, 2267
- Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
- Feroz F., Hobson M. P., Trotta R., 2010, preprint ([arXiv:1001.0719](https://arxiv.org/abs/1001.0719))
- Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2019, *Open J. Astrophys.*, 2, 10
- Fontanot F., Puchwein E., Springel V., Bianchi D., 2013, *MNRAS*, 436, 2672
- Guo H., Zehavi I., Zheng Z., 2012, *ApJ*, 756, 127
- Guo H. et al., 2015a, *MNRAS*, 446, 578
- Guo H. et al., 2015b, *MNRAS*, 453, 4368
- Guo H. et al., 2016, *MNRAS*, 459, 3040
- Guo H., Yang X., Lu Y., 2018, *ApJ*, 858, 30
- Hadzhiyska B., Bose S., Eisenstein D., Hernquist L., Spergel D. N., 2020, *MNRAS*, 493, 5506
- Hamilton A. J. S., 1992, *ApJ*, 385, L5
- Handley W., 2021, *Phys. Rev. D*, 103, L041301
- Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399
- He J.-h., Guzzo L., Li B., Baugh C. M., 2018, *Nat. Astron.*, 2, 967
- Hearin A. P., Zentner A. R., van den Bosch F. C., Campbell D., Tollerud E., 2016, *MNRAS*, 460, 2552
- Hearin A. P. et al., 2017, *AJ*, 154, 190
- Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, *ApJ*, 715, 104
- Hellwing W. A., Schaller M., Frenk C. S., Theuns T., Schaye J., Bower R. G., Crain R. A., 2016, *MNRAS*, 461, L11
- Hikage C. et al., 2019, *PASJ*, 71, 43
- Hill J. C., McDonough E., Toomey M. W., Alexander S., 2020, *Phys. Rev. D*, 102, 043507
- Hirata C. M., 2009, *MNRAS*, 399, 1074
- Hoshino H. et al., 2015, *MNRAS*, 452, 998
- Hou J. et al., 2021, *MNRAS*, 500, 1201
- Howlett C., Ross A. J., Samushia L., Percival W. J., Manera M., 2015, *MNRAS*, 449, 848
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Ivanov M. M., McDonough E., Hill J. C., Simonović M., Toomey M. W., Alexander S., Zaldarriaga M., 2020, *Phys. Rev. D*, 102, 103502
- Jiang F., van den Bosch F. C., 2017, *MNRAS*, 472, 657
- Jiménez E., Contreras S., Padilla N., Zehavi I., Baugh C. M., Gonzalez-Perez V., 2019, *MNRAS*, 490, 3532
- Jing Y. P., Zhang P., Lin W. P., Gao L., Springel V., 2006, *ApJ*, 640, L119
- Kaiser N., 1987, *MNRAS*, 227, 1
- Keeley R. E., Joudaki S., Kaplinghat M., Kirkby D., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 035
- Kreisch C. D., Cyr-Racine F.-Y., Doré O., 2020, *Phys. Rev. D*, 101, 123505
- Kwan J., Heitmann K., Habib S., Padmanabhan N., Lawrence E., Finkel H., Frontiere N., Pope A., 2015, *ApJ*, 810, 35
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Lange J. U., van den Bosch F. C., Zentner A. R., Wang K., Villarreal A. S., 2019a, *MNRAS*, 487, 3112
- Lange J. U., Yang X., Guo H., Luo W., van den Bosch F. C., 2019b, *MNRAS*, 488, 5771
- Lange J. U., van den Bosch F. C., Zentner A. R., Wang K., Hearin A. P., Guo H., 2019c, *MNRAS*, 490, 1870
- Lange J. U., Leauthaud A., Singh S., Guo H., Zhou R., Smith T. L., Cyr-Racine F.-Y., 2021, *MNRAS*, 502, 2074
- Leauthaud A. et al., 2016, *MNRAS*, 457, 4021
- Leauthaud A. 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
- Lewandowski M., Perko A., Senatore L., 2015, *J. Cosmol. Astropart. Phys.*, 2015, 019
- Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, *MNRAS*, 432, 1544
- Matsubara T., 2014, *Phys. Rev. D*, 90, 043537
- McCarthy K. S., Zheng Z., Guo H., 2019, *MNRAS*, 487, 2424
- More S., van den Bosch F. C., Cacciato M., More A., Mo H., Yang X., 2013, *MNRAS*, 430, 747
- Muir J. et al., 2021, *Phys. Rev. D*, 103, 023528
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Neistein E., Khochfar S., 2012, preprint ([arXiv:1209.0463](https://arxiv.org/abs/1209.0463))
- Nishimichi T. et al., 2019, *ApJ*, 884, 29
- Obuljen A., Percival W. J., Dalal N., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 058
- Okumura T., Seljak U., Desjacques V., 2012, *J. Cosmol. Astropart. Phys.*, 2012, 014
- Padilla N., Contreras S., Zehavi I., Baugh C. M., Norberg P., 2019, *MNRAS*, 486, 582

- Padmanabhan N., White M., 2008, *Phys. Rev. D*, 77, 123540
- Parejko J. K. et al., 2013, *MNRAS*, 429, 98
- Peacock J. A., Dodds S. J., 1994, *MNRAS*, 267, 1020
- Pedregosa F. et al., 2012, preprint ([arXiv:1201.0490](https://arxiv.org/abs/1201.0490))
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Planck Collaboration VI, 2020, *A&A*, 641, A6
- Reid B. A., White M., 2011, *MNRAS*, 417, 1913
- Reid B. A., Seo H.-J., Leauthaud A., Tinker J. L., White M., 2014, *MNRAS*, 444, 476
- Ross A. J. et al., 2017, *MNRAS*, 464, 1168
- Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, *ApJ*, 672, 19
- Saito S. et al., 2016, *MNRAS*, 460, 1457
- Salcedo A. N., Wibking B. D., Weinberg D. H., Wu H.-Y., Ferrer D., Eisenstein D., Pinto P., 2020, *MNRAS*, 491, 3061
- Samushia L., Percival W. J., Raccanelli A., 2012, *MNRAS*, 420, 2102
- Sargent W. L. W., Turner E. L., 1977, *ApJ*, 212, L3
- Schaan E. et al., 2021, *Phys. Rev. D*, 103, 063513
- Scoccimarro R., 2004, *Phys. Rev. D*, 70, 083007
- Seljak U., 2000, *MNRAS*, 318, 203
- Seljak U., McDonald P., 2011, *J. Cosmol. Astropart. Phys.*, 2011, 039
- Singh S., Mandelbaum R., Seljak U., Rodríguez-Torres S., Slosar A., 2020, *MNRAS*, 491, 51
- Singh S., Yu B., Seljak U., 2021, *MNRAS*, 501, 4167
- Sinha M., Berlind A. A., McBride C. K., Scoccimarro R., Piscionere J. A., Wibking B. D., 2018, *MNRAS*, 478, 1042
- Smith A., de Mattia A., Burtin E., Chuang C.-H., Zhao C., 2021, *MNRAS*, 500, 259
- Springel V. et al., 2018, *MNRAS*, 475, 676
- Taruya A., Nishimichi T., Saito S., 2010, *Phys. Rev. D*, 82, 063522
- Taylor A., Joachimi B., Kitching T., 2013, *MNRAS*, 432, 1928
- van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, *MNRAS*, 415, 3649
- van Daalen M. P., Angulo R. E., White S. D. M., 2012, *MNRAS*, 424, 2954
- van den Bosch F. C., Norberg P., Mo H. J., Yang X., 2004, *MNRAS*, 352, 1302
- van den Bosch F. C., Weinmann S. M., Yang X., Mo H. J., Li C., Jing Y. P., 2005, *MNRAS*, 361, 1203
- van den Bosch F. C. et al., 2007, *MNRAS*, 376, 841
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- Vattis K., Koushiappas S. M., Loeb A., 2019, *Phys. Rev. D*, 99, 121302
- Verde L., Treu T., Riess A. G., 2019, *Nat. Astron.*, 3, 891
- Villaescusa-Navarro F. et al., 2021, *ApJ*, 915, 71
- Wang L., Reid B., White M., 2014, *MNRAS*, 437, 588
- Wechsler R. H., Tinker J. L., 2018, *ARA&A*, 56, 435
- Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, *MNRAS*, 366, 2
- White M., 2014, *MNRAS*, 439, 3630
- White M., Reid B., Chuang C.-H., Tinker J. L., McBride C. K., Prada F., Samushia L., 2015, *MNRAS*, 447, 234
- Wibking B. D. et al., 2019, *MNRAS*, 484, 989
- Wibking B. D., Weinberg D. H., Salcedo A. N., Wu H.-Y., Singh S., Rodríguez-Torres S., Garrison L. H., Eisenstein D. J., 2020, *MNRAS*, 492, 2872
- Xu X., Zehavi I., Contreras S., 2021, *MNRAS*, 502, 3242
- Yang X., Mo H. J., van den Bosch F. C., 2003, *MNRAS*, 339, 1057
- Ye J.-N., Guo H., Zheng Z., Zehavi I., 2017, *ApJ*, 841, 45
- Yuan S., Eisenstein D. J., 2019, *MNRAS*, 486, 708
- Yuan S., Eisenstein D. J., Leauthaud A., 2020, *MNRAS*, 493, 5551
- Yuan S., Hadzhiyska B., Bose S., Eisenstein D. J., Guo H., 2021, *MNRAS*, 502, 3582
- Zentner A. R., Berlind A. A., Bullock J. S., Kravtsov A. V., Wechsler R. H., 2005, *ApJ*, 624, 505
- Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, *MNRAS*, 443, 3044
- Zentner A. R., Hearin A., van den Bosch F. C., Lange J. U., Villarreal A., 2019, *MNRAS*, 485, 1196
- Zhai Z. et al., 2019, *ApJ*, 874, 95
- Zheng Z., Guo H., 2016, *MNRAS*, 458, 4015
- Zheng Z., Coil A. L., Zehavi I., 2007, *ApJ*, 667, 760
- Zu Y., 2020, preprint ([arXiv:2010.01143](https://arxiv.org/abs/2010.01143))

## APPENDIX A: SIMULATION STATISTICS

Tables A1 and A2 show the cosmological parameters, goodness-of-fit values, and evidence values with respect to the two BOSS LOWZ samples assuming the default analysis choices.

**Table A1.** Overview of the simulations used in this work and their corresponding evidence and goodness-of-fit values with respect to the observations.

Name	$H_0$ [km s <sup>-1</sup> Mpc <sup>-1</sup> ]	$w_0$	$\sigma_8$	$\Omega_{m,0}$	$\Omega_{b,0}$	$n_s$	$N_{\text{eff}}$	$f\sigma_8$	$z = 0.25$ $\Delta \ln \mathcal{Z}$	$\chi^2$	$f\sigma_8$	$z = 0.40$ $\Delta \ln \mathcal{Z}$	$\chi^2$
B00	63.4	-0.82	0.77	0.34	0.056	0.98	2.92	0.445	-5.0	34.0	0.444	-3.3	30.1
B01	73.1	-1.13	0.90	0.26	0.042	0.98	3.17	0.499	-0.8	23.9	0.517	-6.3	31.7
B02	63.7	-0.68	0.69	0.32	0.057	1.00	3.26	0.380	-11.6	43.6	0.379	-5.3	29.4
B03	64.0	-0.74	0.67	0.33	0.055	0.95	3.56	0.375	-7.5	37.3	0.375	-8.4	37.3
B04	65.0	-0.77	0.75	0.30	0.052	0.97	2.66	0.407	-2.6	27.7	0.410	-3.4	26.0
B05	72.8	-1.33	0.93	0.28	0.039	0.93	2.96	0.556	-6.5	30.9	0.573	-23.0	61.0
B06	62.7	-0.71	0.71	0.34	0.058	0.97	2.71	0.400	-10.7	43.1	0.398	-3.9	29.0
B07	64.4	-0.87	0.78	0.34	0.055	0.97	3.94	0.453	-2.3	27.9	0.454	-2.8	27.5
B08	69.4	-1.16	0.89	0.30	0.043	0.95	3.60	0.525	-5.0	27.5	0.536	-12.2	41.4
B09	62.4	-0.83	0.72	0.35	0.055	0.95	3.90	0.422	-2.6	28.8	0.421	-2.8	25.6
B10	72.1	-1.24	0.85	0.29	0.042	0.96	4.24	0.502	-2.7	26.7	0.515	-11.1	42.2
B11	67.7	-0.86	0.81	0.29	0.049	1.00	2.83	0.437	-1.6	24.9	0.445	-3.6	28.9
B12	65.4	-0.88	0.79	0.33	0.053	0.95	2.88	0.450	-1.9	25.7	0.452	-1.1	22.8
B13	71.1	-1.12	0.87	0.28	0.043	0.98	3.00	0.490	-0.7	23.4	0.505	-6.8	35.4
B14	68.7	-1.12	0.92	0.32	0.048	0.97	2.75	0.549	-9.8	36.5	0.557	-14.3	43.5
B15	74.1	-1.30	0.91	0.27	0.039	0.93	3.73	0.532	-3.3	26.0	0.549	-17.4	52.9
B16	70.1	-1.13	0.81	0.29	0.044	0.97	3.77	0.466	-1.7	26.5	0.478	-4.7	30.3
B17	74.4	-1.25	0.87	0.26	0.040	0.95	3.22	0.492	-1.0	25.0	0.512	-6.9	34.6
B18	70.8	-1.03	0.74	0.28	0.046	0.95	4.28	0.410	-2.1	27.9	0.420	-6.1	31.1
B19	72.4	-1.09	0.81	0.26	0.043	0.97	3.68	0.444	-0.1	23.1	0.460	-3.8	29.3
B20	67.1	-0.99	0.84	0.32	0.049	0.95	3.39	0.489	-0.6	22.7	0.494	-5.5	31.3
B21	66.4	-0.87	0.76	0.32	0.054	0.98	3.85	0.432	-3.3	28.6	0.435	-2.3	28.0
B22	68.1	-1.03	0.88	0.31	0.046	0.96	2.62	0.509	-1.9	24.7	0.518	-10.7	40.9
B23	62.0	-0.57	0.57	0.32	0.059	0.97	3.47	0.305	-33.7	77.8	0.303	-30.5	72.4
B24	63.0	-0.76	0.69	0.33	0.057	0.96	4.15	0.390	-4.3	30.7	0.390	-3.4	27.5
B25	65.7	-0.95	0.79	0.32	0.048	0.93	3.09	0.456	-1.4	25.5	0.460	-3.7	28.2
B26	71.8	-1.13	0.90	0.27	0.043	0.94	2.79	0.512	-0.6	21.1	0.527	-11.9	42.4
B27	67.4	-0.96	0.73	0.30	0.047	0.97	4.02	0.410	-2.0	27.4	0.417	0.0	22.9
B28	74.8	-1.40	0.96	0.27	0.039	0.96	3.81	0.576	-13.4	43.8	0.596	-22.9	58.6
B29	71.4	-1.24	0.93	0.30	0.044	0.94	3.43	0.555	-6.9	30.4	0.568	-18.0	50.1
B30	73.4	-1.22	0.87	0.27	0.041	0.96	4.07	0.499	0.0	23.5	0.516	-10.1	42.6
B31	73.8	-1.38	0.94	0.27	0.039	0.96	3.34	0.560	-8.4	34.9	0.579	-15.2	45.7
B32	68.4	-0.93	0.72	0.29	0.048	0.95	3.98	0.399	-2.1	26.3	0.406	-3.4	23.1
B33	66.1	-0.88	0.78	0.32	0.054	0.99	3.64	0.440	-1.7	26.4	0.444	-2.4	26.4
B34	69.1	-1.03	0.86	0.30	0.048	0.95	3.13	0.494	-1.9	26.3	0.503	-7.1	32.6
B35	61.7	-0.61	0.68	0.34	0.062	1.00	3.05	0.378	-14.0	50.1	0.374	-5.2	28.5
B36	70.4	-1.11	0.91	0.29	0.044	0.97	3.30	0.520	-2.1	24.4	0.533	-14.7	47.3
B37	66.7	-0.85	0.73	0.30	0.051	0.98	3.51	0.399	-4.1	31.1	0.404	-6.4	34.0
B38	69.7	-0.96	0.74	0.29	0.049	0.98	4.11	0.408	-4.9	33.7	0.416	-3.0	28.7
B39	64.7	-0.94	0.76	0.34	0.052	0.96	4.19	0.448	-3.4	30.2	0.450	-3.2	29.3

**Table A2.** Similar to Table A1 but for the test simulations of the Aemulus simulation suite.

Name	$H_0$ [km s <sup>-1</sup> Mpc <sup>-1</sup> ]	$w_0$	$\sigma_8$	$\Omega_{m,0}$	$\Omega_{b,0}$	$n_s$	$N_{\text{eff}}$	$f\sigma_8$	$z = 0.25$ $\Delta \ln \mathcal{Z}$	$\chi^2$	$f\sigma_8$	$z = 0.40$ $\Delta \ln \mathcal{Z}$	$\chi^2$
T00-0	63.2	-0.73	0.69	0.33	0.058	0.98	2.95	0.386	-6.8	36.2	0.386	-1.9	25.8
T00-1	63.2	-0.73	0.69	0.33	0.058	0.98	2.95	0.386	-7.6	37.2	0.386	-5.1	28.0
T00-2	63.2	-0.73	0.69	0.33	0.058	0.98	2.95	0.386	-9.9	43.4	0.386	-4.1	30.1
T00-3	63.2	-0.73	0.69	0.33	0.058	0.98	2.95	0.386	-12.3	47.8	0.386	-6.4	33.2
T00-4	63.2	-0.73	0.69	0.33	0.058	0.98	2.95	0.386	-9.5	41.4	0.386	-5.6	33.9
T01-0	65.7	-0.86	0.75	0.31	0.053	0.97	3.20	0.424	-2.8	29.6	0.428	-2.8	25.9
T01-1	65.7	-0.86	0.75	0.31	0.053	0.97	3.20	0.424	-3.3	28.3	0.428	-2.5	26.1
T01-2	65.7	-0.86	0.75	0.31	0.053	0.97	3.20	0.424	-3.2	29.4	0.428	-2.1	25.5
T01-3	65.7	-0.86	0.75	0.31	0.053	0.97	3.20	0.424	-3.5	28.8	0.428	-1.7	26.3
T01-4	65.7	-0.86	0.75	0.31	0.053	0.97	3.20	0.424	-2.9	28.8	0.428	-4.7	27.0
T02-0	68.2	-1.00	0.81	0.30	0.048	0.96	3.45	0.461	-1.3	25.6	0.469	-3.0	27.2
T02-1	68.2	-1.00	0.81	0.30	0.048	0.96	3.45	0.461	-1.1	25.7	0.469	-3.2	28.4
T02-2	68.2	-1.00	0.81	0.30	0.048	0.96	3.45	0.461	-0.8	25.3	0.469	-4.9	30.1
T02-3	68.2	-1.00	0.81	0.30	0.048	0.96	3.45	0.461	-3.2	28.6	0.469	-1.7	24.0
T02-4	68.2	-1.00	0.81	0.30	0.048	0.96	3.45	0.461	-0.9	24.6	0.469	-5.3	32.4
T03-0	70.7	-1.13	0.86	0.29	0.043	0.95	3.70	0.496	-0.2	22.7	0.508	-8.5	38.6
T03-1	70.7	-1.13	0.86	0.29	0.043	0.95	3.70	0.496	-0.2	23.4	0.508	-5.0	32.2
T03-2	70.7	-1.13	0.86	0.29	0.043	0.95	3.70	0.496	0.5	21.6	0.508	-5.5	30.1
T03-3	70.7	-1.13	0.86	0.29	0.043	0.95	3.70	0.496	-1.4	24.9	0.508	-4.4	30.0
T03-4	70.7	-1.13	0.86	0.29	0.043	0.95	3.70	0.496	-1.5	24.2	0.508	-10.5	40.4
T04-0	73.2	-1.27	0.91	0.28	0.040	0.94	3.95	0.531	-5.1	27.2	0.548	-18.2	49.8
T04-1	73.2	-1.27	0.91	0.28	0.040	0.94	3.95	0.531	-2.7	25.1	0.548	-11.6	39.7
T04-2	73.2	-1.27	0.91	0.28	0.040	0.94	3.95	0.531	-6.3	31.7	0.548	-16.8	50.3
T04-3	73.2	-1.27	0.91	0.28	0.040	0.94	3.95	0.531	-1.7	24.3	0.548	-9.5	38.2
T04-4	73.2	-1.27	0.91	0.28	0.040	0.94	3.95	0.531	-2.8	25.7	0.548	-14.4	47.5
T05-0	69.7	-1.09	0.82	0.28	0.045	0.95	3.70	0.464	1.6	20.7	0.476	-3.9	28.2
T05-1	69.7	-1.09	0.82	0.28	0.045	0.95	3.70	0.464	-0.7	23.4	0.476	-8.7	38.2
T05-2	69.7	-1.09	0.82	0.28	0.045	0.95	3.70	0.464	-0.1	22.1	0.476	-6.2	33.8
T05-3	69.7	-1.09	0.82	0.28	0.045	0.95	3.70	0.464	-1.3	26.2	0.476	-4.0	29.1
T05-4	69.7	-1.09	0.82	0.28	0.045	0.95	3.70	0.464	0.0	23.1	0.476	-4.4	33.6
T06-0	66.7	-0.90	0.80	0.32	0.051	0.97	3.20	0.457	-2.1	27.0	0.461	-8.2	38.4
T06-1	66.7	-0.90	0.80	0.32	0.051	0.97	3.20	0.457	-4.8	32.9	0.461	-5.9	34.1
T06-2	66.7	-0.90	0.80	0.32	0.051	0.97	3.20	0.457	-3.3	30.7	0.461	-6.1	33.8
T06-3	66.7	-0.90	0.80	0.32	0.051	0.97	3.20	0.457	-4.5	30.3	0.461	-1.5	22.1
T06-4	66.7	-0.90	0.80	0.32	0.051	0.97	3.20	0.457	-3.4	29.8	0.461	-2.1	26.6

## APPENDIX B: SCATTER IN COSMOLOGICAL EVIDENCE

When building a model for the cosmological evidence  $\mathcal{Z}(\mathcal{C}|\mathbf{D})$ , we make the simplifying assumption that it depends solely on  $f\sigma_8$  and, at most, one additional cosmological parameter. At the same time, it is evident from Figs 5, 6, and 9 that the models for  $\mathcal{Z}(\mathcal{C}|\mathbf{D})$  does not fit the data perfectly. A concern is that scatter between calculated and model evidence values is due to unmodelled cosmological parameter dependencies and that neglecting such dependencies biases our marginal constraints on  $f\sigma_8$ .

First, not all unmodelled dependencies necessarily bias the constraints on  $f\sigma_8$ . For example, if the dependence on a second cosmological parameter  $\gamma$  is perfectly separable, i.e.  $\mathcal{Z}(f\sigma_8, \gamma|\mathbf{D}) = f(f\sigma_8)g(\gamma)$ , neglecting the dependence on  $\gamma$  should not significantly bias the constraint on  $f\sigma_8$ . Additionally, we have performed systematic checks in Section 4 that show that our results are robust with respect to various modelling choices.

However, we can use the test simulations of the Aemulus simulation suite to investigate more directly what the source of the scatter is. The test simulations of the Aemulus suite are seven groups of five simulations, each. All simulations of the same group share the exact same cosmological parameters but different random seeds for the initial conditions. Thus, any scatter in the evidence in a group

of simulations with the same cosmological parameters is entirely due to random fluctuations. In Table A2, we show the cosmological evidence for all test simulations with respect to our two observational data sets. For each set of five simulations, we estimate the variance in the scatter,

$$\hat{S}_{\ln \mathcal{Z}}^2 = \frac{1}{5-1} \sum_{i=1}^5 [\ln \mathcal{Z}_i(\mathcal{C}|\mathbf{D}) - \text{Mean}(\ln \mathcal{Z}(\mathcal{C}|\mathbf{D}))]^2. \quad (\text{B1})$$

For the  $7 \times 2 = 14$  combinations of simulation sets and observations, the average variance in the scatter is  $\langle \hat{S}_{\ln \mathcal{Z}}^2 \rangle = 3.6$ . At the same time, when fitting a model to  $\mathcal{Z}(\mathcal{C}|\mathbf{D})$ , we also obtain an estimate for the scatter  $S_{\ln \mathcal{Z}}^2$  that is due to both statistical fluctuations and possibly unmodelled parameters. If the scatter in the evidence values around the best fit were primarily due to unmodelled cosmological parameter dependencies, we would overestimate the scatter found for the test simulations where the only source of scatter is statistical fluctuations. However, our model predicts  $\langle \hat{S}_{\ln \mathcal{Z}}^2 \rangle$  to be in the range 2.1–6.1 (95 per cent certainty), agreeing well with our finding of 3.6. Thus, the observed scatter in the evidence values of the test simulations are compatible with cosmic variance being the major source of scatter in Figs 5, 6, and 9.

### APPENDIX C: SUB-POISSON SCATTER FOR SATELLITES

In our default analysis, we assume that the satellite occupation numbers at fixed halo mass and concentration follow a Poisson distribution. This is motivated by the finding that subhaloes that host satellite galaxies follow a distribution that is close to Poisson (Jiang & van den Bosch 2017). Recently, Zu (2020) has advocated for maximally sub-Poisson satellite numbers in order to alleviate the so-called ‘lensing is low’ tension. This integer distribution with the smallest possible scatter assigns  $i$  satellites with a probability of  $p = 1 - \langle N_s \rangle + i$  and  $i + 1$  satellites with a probability  $1 - p$ , where  $i$  is an integer with  $i \leq \langle N_s \rangle < i + 1$ . For this distribution, one can show  $\langle N_s(N_s - 1) \rangle = 2\langle N_s \rangle i - i(i + 1)$ .

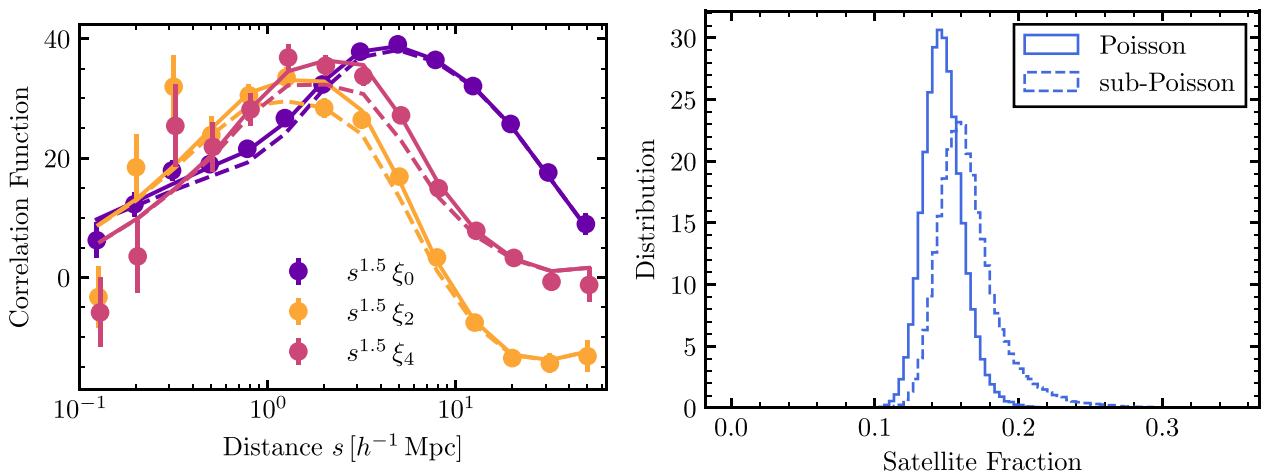
We note that subhalo occupation numbers, whilst slightly sub-Poisson for low occupation numbers, are far from maximally sub-Poisson (Jiang & van den Bosch 2017). Additionally, whether each subhalo host a satellite is determined by galaxy formation physics, an additional stochastic process. If galaxies hosted by subhaloes evolve independently of each other, the distribution will be driven towards a Poisson distribution. Similarly, if satellite properties inside each halo are positively correlated with each other, an effect called one-halo galactic conformity (Weinmann et al. 2006), this will drive the distribution of satellites further towards a super-Poisson distribution. In both cases, one would expect the satellite number distribution to be even further away from a maximally sub-Poisson distribution than subhaloes.

None the less, we have performed the analysis for the  $0.18 \leq z < 0.30$  LOWZ sample assuming a maximally sub-Poisson distribution for satellites. In Fig. C1, we show how a change in the satellite number distribution affects the observables. The galaxy–halo model for both the Poisson and sub-Poisson model is the best-fitting model for the  $0.18 \leq z < 0.30$  LOWZ sample when Poisson numbers are assumed. Thus, all differences arise entirely from changes to

the second moment of the satellite distribution. For the sub-Poisson distribution, we observe that all the multipoles of the redshift–space correlation function are slightly suppressed on small scales. This makes sense since assuming sub-Poisson satellite number suppresses the one-halo satellite–satellite terms, which scales with  $\langle N_s(N_s - 1) \rangle$ . Although  $\langle N_s(N_s - 1) \rangle$  is strongly suppressed compared to the Poisson assumption, the effect is weak because the one-halo central–satellite term dominates over the one-halo satellite–satellite terms (see e.g. Zheng & Guo 2016) and is unaffected by the second moment of the satellite distribution.

Given the significant changes in the multipoles at fixed cosmology and galaxy–halo connection parameters, it is evident that assumption on the satellite numbers will affect posterior constraints. For example, the right-hand panel of Fig. C1 demonstrates how cosmology-averaged inferences on the satellite fraction  $f_{\text{sat}}$  are affected by the choice of satellite distribution. Assuming a maximally sub-Poisson distribution implies slightly higher satellite fractions. This makes sense since a sub-Poisson distribution reduces the strength of the *Fingers of God* effect and the latter is positively correlated with  $f_{\text{sat}}$ . However, even in the extreme case of a maximally sub-Poisson distribution, the shift in the posterior constraint on  $f_{\text{sat}}$  compared to the Poisson scenario is very small,  $\Delta f_{\text{sat}} \sim 0.02$ . Thus, by itself, assuming a strongly sub-Poisson satellite distribution is likely not enough to make the large satellite fractions advocated for in Zu (2020) compatible with RSD constraints in this work or other results from the literature (Reid et al. 2014; Guo et al. 2015a; Saito et al. 2016).

Finally, we do not find that the choice of the second moment for the satellite distribution impacts the cosmological constraint significantly: Assuming a maximally sub-Poisson distribution for satellites reduces the inferred  $f\sigma_8$  posterior by only 0.012 or around  $0.5\sigma$ . More modest variations from Poisson distributions would likely result in even smaller shifts of  $f\sigma_8$ .



**Figure C1.** Impact of the second moment of the satellite distribution on observables (left) and posterior constraints on the satellite fraction (right). In both panels, we compare models assuming a Poisson distribution for satellites (solid) against a maximally sub-Poisson distribution (dashed).

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.