Building a digital twin of a luminous red galaxy spectroscopic survey: galaxy properties and clustering covariance

César Hernández-Aguayo [1,2,3]★ Francisco Prada,[4]★ Carlton M. Baugh[1] and Anatoly Klypin[5]

[1]*Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, D-85748 Garching, Germany*
[3]*Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany*
[4]*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, E-18080 Granada, Spain*
[5]*Astronomy Department, New Mexico State University, Las Cruces, NM 88001, USA*

ABSTRACT

Upcoming surveys will use a variety of galaxy selections to map the large-scale structure of the Universe. It is important to make accurate predictions for the properties and clustering of such galaxies, including the errors on these statistics. Here, we describe a novel technique which uses the semi-analytical model of galaxy formation GALFORM, embedded in the high-resolution *N*-body Planck-Millennium simulation, to populate a thousand halo catalogues generated using the Parallel-PM *N*-body GLAM code. Our hybrid scheme allows us to make clustering predictions on scales that cannot be modelled in the original *N*-body simulation. We focus on luminous red galaxies (LRGs) selected in the redshift range $z = 0.6 - 1$ from the GALFORM output using *similar* colour-magnitude cuts in the r, z, and $W1$ bands to those that will be applied in the Dark Energy Spectroscopic Instrument (DESI) survey, and call this illustrative sample 'DESI-like' LRGs. We find that the LRG-halo connection is non-trivial, leading to the prediction of a non-standard halo occupation distribution; in particular, the occupation of central galaxies does not reach unity for the most massive haloes, and drops with increasing mass. The GLAM catalogues reproduce the abundance and clustering of the LRGs predicted by GALFORM. We use the GLAM mocks to compute the covariance matrices for the two-point correlation function and power spectrum of the LRGs and their background dark matter density field, revealing important differences. We also make predictions for the linear-growth rate and the baryon acoustic oscillations distances at $z = 0.6$, 0.74, and 0.93. All 'DESI-like' LRG catalogues are made publicly available.

Key words: methods: data analysis – methods: statistical – galaxies: formation – galaxies: haloes – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Luminous red galaxies (LRGs) display a strong clustering signal, thanks to their bright intrinsic luminosity and large stellar masses, that makes them an ideal tracer of the large-scale structure of the Universe (Eisenstein et al. 2005a; Zehavi et al. 2005). LRGs were used to extract the scale of the baryon acoustic oscillations (BAO) in the local large-scale structure using the redshift-space correlation function measured from the Sloan Digital Sky Survey (SDSS; Eisenstein et al. 2005b). LRGs have also been used to study the impact of redshift-space distortions (RSDs) on their small and large scale clustering (see e.g. Zehavi et al. 2005; Wake et al. 2008; Cabre & Gaztanaga 2009a,b; Crocce et al. 2011; Samushia, Percival & Raccanelli 2012). Additionally, the large-scale clustering of LRGs has also been used to constrain the cosmological parameters (Eisenstein et al. 2005b; Tegmark et al. 2006; Sanchez et al. 2009; Aubourg et al. 2015; Tröster et al. 2020), and to test modified gravity models (see e.g. Barreira, Sánchez & Schmidt 2016; Hernández-Aguayo et al. 2019).

LRGs are the main targets of the SDSS-III Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), in the redshift range $0.2 < z < 0.75$. This survey has provided the most precise measurements to date of cosmological distances using BAO and the growth rate using RSDs at effective redshifts $z = 0.38$, 0.51, and 0.61 (see Alam et al. 2017 and references therein). Recently, the SDSS-IV extended-BOSS survey (eBOSS; Dawson et al. 2016; Prakash et al. 2016) has presented the first clustering measurements of LRGs at $z \sim 0.7$ (Zhai et al. 2017; Bautista et al. 2018; Icaza-Lizaola et al. 2020).

The Dark Energy Spectroscopic Instrument[1] (DESI) survey aims to measure BAO scales and the growth of structure through RSDs at an unprecedented level of precision (DESI Collaboration et al. 2016). This imminent survey will target luminous red galaxies in the redshift range from $z = 0.4$ to 1, [OII] emission-line galaxies (ELGs) in the range $0.6 < z < 1.6$, QSOs (tracers) up to $z = 2.1$, and QSOs (Ly-α) at higher redshifts ($2.1 < z < 3.5$). In addition to a bright galaxy sample at low redshifts $z < 0.4$, DESI will provide a total of $\sim$35 million biased tracers of the large-scale structure of

★ E-mail: cesarhdz@MPA-Garching.MPG.DE (CHA); f.prada@csic.es (FP)

[1] https://www.desi.lbl.gov

the Universe over 14 000 deg^2 (see Kitanidis et al. 2019 for details). The LRG target selection at $z < 0.6$ will be complementary to that performed in the SDSS-IV/eBOSS (Prakash et al. 2016); hence, we will focus here on the DESI LRGs at $z \geq 0.6$.

Our aim here is to provide a qualitative study of the properties and clustering of LRGs which meet the selection requirements of a real survey such as DESI. We select DESI-like LRGs from the output of the semi-analytic model (SAM) of galaxy formation GALFORM (Cole et al. 2000) run on the Planck-Millennium *N*-body simulation (Baugh et al. 2019), and provide estimates of the large-scale galaxy clustering using the GLAM code, which allows us to generate a substantial number of large galaxy mock catalogues (Klypin & Prada 2018). This hybrid approach takes the SAM calculations made using a high-resolution, moderate volume *N*-body simulation and uses the results to populate a large number ($\mathcal{O}(10^3)$) of larger volume low-resolution simulations run with GLAM. This allows us to make predictions for the large-scale clustering of LRGs on scales, such as the BAO scale, that were inaccessible in the simulation used to run the SAM. Furthermore, by being able to generate a large number of independent realization of the density field at relatively low computational cost, we can estimate the covariance on two-point statistics of the large-scale structure. Note that we describe the LRG samples drawn from the models as being 'DESI-like' because the colour-magnitude selection of DESI LRGs remains to be finalized (see Section 3). The key point of using LRGs is that a colour-magnitude selection is applied to the models, resulting in a non-canonical prediction for the halo occupation distribution (HOD), illustrating the need to use a physically motivated galaxy formation model, such as the GALFORM SAM.

The use of SAMs to study the properties and clustering of LRGs is not new. Almeida, Baugh & Lacey (2007) and Almeida et al. (2008) presented predictions for the abundance, structural, and photometric properties of LRGs using two earlier versions of GALFORM (Baugh et al. 2005; Bower et al. 2006). The authors found that their predictions were in good agreement with different observations from the SDSS (Bernardi et al. 2003, 2005; Wake et al. 2006). More recently, Stoppacher et al. (2019) used the GALACTICUS SAM (Benson 2012) run on the MultiDark Planck 2 simulation (Klypin et al. 2016) to study the galaxy-halo connection and clustering of the BOSS-CMASS DR12 sample (Alam et al. 2015), finding good agreement between predictions and observations.

Recently, Zhou et al. (2020a) presented small-scale ($r \lesssim 20\,h^{-1}$Mpc) clustering measurements of DESI-like LRGs selected from the DESI Legacy Imaging Surveys[2] (Dey et al. 2019) and fitted their results using the HOD framework. Since spectroscopic redshifts are not yet available for these targets, these authors estimated photometric redshifts (photo-z) using the Dark Energy Camera Legacy Survey (DECaLS) imaging. There are a number of differences between the work carried out by Zhou et al. (2020a) and our paper: first, we are interested in providing a study of the impact of the DESI-LRG target selection on galaxy properties and the galaxy-halo connection using a physical model of galaxy formation, GALFORM; and secondly, we focus on the large-scale galaxy clustering and in the generation of a large number of mock catalogues to provide an accurate estimate of the covariance of the clustering measurements. Both of these objectives are beyond the reach of the original simulation used to run the SAM and mark a key advantage of our hybrid approach.

In order to extract the cosmological information from our GLAM mock catalogues for the DESI LRG tracers, it is necessary to meet the requirements of the expected error budget for DESI. Hence, it is imperative to construct covariance matrices for our clustering measurements (see e.g. Baumgarten & Chuang 2018; Blot et al. 2019; Colavincenzo et al. 2019; Lippich et al. 2019 and references therein). Here, we make predictions of the linear-growth rate through a linear theory description of RSDs (Kaiser 1987; Hamilton 1992); and an isotropic analysis of the BAO scale (see e.g. Anderson et al. 2014) in configuration and Fourier space using the covariance matrices constructed from our GLAM catalogues.

The outline of the paper is as follows. In Section 2, we present the simulations used in our analysis. Section 3 describes the selection of DESI-like LRGs from GALFORM. In Section 4, we provide a detailed study of the galaxy-halo connection of DESI-like LRGs. Our results for the galaxy clustering and covariance errors are presented in Section 5. Finally, in Section 6, we give our summary and conclusions. We stress that is it not our intention to produce definitive predictions for the clustering and covariance of DESI LRGs, in part, because the detailed LRG selection is still to be finalized. Instead, our aim is to illustrate our approach using a sample that approximates to DESI LRGs.

2 SIMULATIONS AND GALAXY FORMATION IN SEMI-ANALYTICAL MODELS

Here, we introduce the Planck Millennium *N*-body simulation and the galaxy formation model (Section 2.1). The GLAM simulations are described in Section 2.2. In Section 2.3, we show the halo mass function and halo clustering of our simulations.

2.1 Galaxy formation in the Planck Millennium simulation

The Planck Millennium *N*-body simulation (hereafter the PMILL simulation; Baugh et al. 2019) follows the evolution of 5040^3 dark matter particles in a cosmological volume of $542.16^3\,h^{-3}$Mpc^3 (800^3Mpc^3). The simulation was run using a reduced memory version of the GADGET-2 *N*-body code (Springel 2005), employing the cosmological parameters corresponding to the 2014 results from the Planck collaboration (Ade et al. 2014)

$$\{\Omega_b, \Omega_m, h, n_s, \sigma_8\} = \{0.04825, 0.307, 0.6777, 0.9611, 0.8288\}.$$

The large number of dark matter particles used in the PMILL simulation gives a mass resolution of $1.06 \times 10^8\,h^{-1}$M$_\odot$ and a halo mass limit, corresponding to 20 particles, of $2.12 \times 10^9\,h^{-1}$M$_\odot$. The simulation starts at $z = 127$, with initial conditions generated using second-order Lagrangian perturbation theory (Jenkins 2010) and the PANPHASIA code (Jenkins 2013). The halo properties and selected particle information are saved in 271 snapshots. Haloes and sub-haloes were identified with SUBFIND (Springel et al. 2001). SUBFIND first identifies haloes using a friend-of-friends (FoF) algorithm with a linking length of $b = 0.2$ times the mean interparticle separation. Then, these FoF groups (main or distinct haloes) are split into subhaloes of bound particles. SUBFIND uses several definitions of halo mass; we use M_{200m} which is the mass enclosed within a radius where the average overdensity is 200 times the mean density of the Universe. The subhalo mass is just the sum of the mass of the particles that are gravitationally bound to that subhalo. The haloes and subhaloes are used to build halo merger trees using the DHALO code (Jiang et al. 2014).

Here, we use the GALFORM semi-analytical model of galaxy formation (Cole et al. 2000; Baugh 2006; Gonzalez-Perez et al.

2014; Lacey et al. 2016) to populate the dark matter haloes in the PMILL simulation with galaxies. We use the recalibration of the Gonzalez-Perez et al. (2014) model presented by Baugh et al. (2019) to identify LRGs and study their clustering. In order to match local observations of galaxies, just two of the parameters describing the physical processes modelled in GALFORM were changed slightly by Baugh et al., from the values adopted by Gonzalez-Perez et al., to take into account the change in cosmology and mass resolution in the PMILL compared with the original *N*-body simulation used by Gonzalez-Perez et al., and an improvement to the treatment of galaxy mergers (see Baugh et al. 2019 for further details of these changes; we note that Gonzalez-Perez et al. 2018 used an updated version of their model, which also included the new galaxy merger scheme first implemented by Campbell et al. 2015 and explained in full by Simha & Cole 2017).

2.2 GLAM simulations

GLAM is a new *N*-body Parallel Particle-Mesh (PM) code developed for the massive production of large volume mock galaxy catalogues (Klypin & Prada 2018). GLAM first generates the density field at an early epoch, including peculiar velocities, for a particular cosmological model and initial conditions. The code uses a regularly spaced three-dimensional (3D) mesh of size N_g^3 that covers the cubic domain L^3 of a simulation box using N_p^3 particles. The size of a cell, $\Delta x = L/N_g$, and the mass of each particle, m_p, define the force and mass resolutions, respectively (see appendix A of Klypin & Prada 2018 for details).

We generate 1000 GLAM simulations using the same cosmology and linear perturbation theory power spectrum as used in the PMILL simulation. Because our goal is to study the clustering of LRGs, the GLAM simulations follow the evolution of 2000^3 particles of mass $1.06 \times 10^{10}\,h^{-1}\mathrm{M}_\odot$ in a cubic box of size $1\,h^{-1}\mathrm{Gpc}$ with $N_s = 136$ time-steps, and mesh of $N_g = 4000$. This numerical set-up yields a spatial resolution of $\Delta x = 0.25\,h^{-1}\mathrm{Mpc}$. The initial conditions were generated using the Zeldovich approximation starting at $z_{\mathrm{ini}} = 100$.

Haloes in GLAM are identified with the bound density maximum (BDM) halo finder (Klypin & Holtzman 1997). Only distinct haloes are saved in our catalogues. In BDM the virial mass, M_{vir}, is adopted as the definition of halo mass. The virial mass of a halo corresponds to the mass enclosed within a spherical overdensity of radius R_{vir}, such that the mean overdensity within this radius is $\Delta_{\mathrm{vir}} \approx 330$ times the mean matter density of the Universe at the present time. The virial overdensity, $\Delta_{\mathrm{vir}}(z)$, is computed using the approximation of Bryan & Norman (1998). Only halo catalogues are saved in 21 snapshots between $0 < z < 1.2$ for each realization.

2.3 Halo mass function and halo clustering

To check the performance of our GLAM simulations, we compare the halo mass function and the halo two-point correlation function measured from them with those obtained from the PMILL simulation. Since we are interested in LRGs at $z \geq 0.6$, we use halo catalogues corresponding to snapshots at $z = 0.6, 0.74$, and 0.93, where $z = 0.74$ corresponds to the median redshift of the expected $n(z)$ distribution of LRGs in DESI (DESI Collaboration et al. 2016; Zhou et al. 2020a). In a future work, we plan to build proper light-cones using all the GLAM halo catalogues available in the relevant redshift range.

The upper panels in Fig. 1 show the differential halo mass function measured at $z = 0.6, 0.74$, and 0.93 from the PMILL run (black solid lines) and the GLAM simulations (blue dots with errorbars) using M_{vir} as the halo mass definition. We use the mass conversion algorithm

of Hu & Kravtsov (2003) to convert $M_{200\mathrm{m}}$ into M_{vir} for the PMILL measurements. We find good agreement between the GLAM and PMILL results, with a difference of less than 10 per cent for haloes with mass $\log_{10}(M_{\mathrm{vir}}/h^{-1}\mathrm{M}_\odot) > 12.5$ at all redshifts. This mass value is well below the typical LRG host halo mass (see below). The differences seen between the results from GLAM and PMILL for lower mass haloes are due to the lower resolution in the GLAM simulations. The differences seen at the high-mass end are due to the much smaller volume of the PMILL simulation compared with that used in the GLAM simulations.

The real-space clustering of haloes of mass $\log_{10}(M_{\mathrm{vir}}/h^{-1}M_\odot) > 12.5$ is shown in the lower panels of Fig. 1 at different redshifts. We find good agreement in the clustering measured on scales $r > 2\,h^{-1}\mathrm{Mpc}$ between the two types of simulations. There is a 10 per cent difference over the separation range $2 < r/h^{-1}\mathrm{Mpc} < 40$. Nevertheless, GLAM predicts a higher clustering amplitude for $r \sim 1\,h^{-1}\mathrm{Mpc}$ with respect to that measured in the PMILL simulation. This effect is due to the different algorithms used to find dark matter haloes, i.e. BDM predicts more halo pairs at small separations, hence resulting in a higher clustering amplitude on small scales. As we will see in Section 5, the difference in the halo clustering does not affect the clustering of LRGs when an appropriate HOD is applied to the GLAM catalogues.

3 SELECTION OF LUMINOUS RED GALAXIES

The DESI team plan to use the 3.4 μm band (*W1*) from the space-based *Wide-Field Infrared Survey Explorer* (*WISE*), in combination with the *r* and *z* bands from the DESI Legacy Imaging Surveys (Dey et al. 2019), to select LRGs efficiently in the redshift range $0.6 < z < 1.0$ (DESI Collaboration et al. 2016). The DESI LRG selection is currently being developed and will doubtless be revised further once the DESI Survey Validation is completed and the performance of the spectrograph on the telescope has been quantified. As we stated in the introduction, it is not our intention to make definitive predictions for DESI LRGs, but rather to present a 'proof-of-concept' study that illustrates our methodology. To meet this purpose, we adopt the LRG selection described by Zhou et al. (2020a). This is an updated version of the DESI LRG target selection from DESI Collaboration et al. (2016), which we hereafter refer to as 'DESI-like', since it uses the same filter set that will be applied in the final DESI-LRG selection, if not the precise numerical values of the cuts

$$z < 20.41 \tag{1}$$

$$-0.6 < (z - W_1) - 0.8(r - z), \tag{2}$$

$$r - z > 0.9, \tag{3}$$

$$r - z > (z - 17.18)/2. \tag{4}$$

We note that, for reference, at the time of writing, the DESI LRG selection based on Legacy Surveys Data Release 8 is described in Zhou et al. (2020b).

GALFORM outputs observer frame absolute magnitudes with dust attenuation, M_{AB}, so we need to convert these into apparent magnitudes, m_{AB}, in order to apply the above cuts

$$m_{\mathrm{AB}} = M_{\mathrm{AB}} + 5\log_{10}(d_L(z)/h^{-1}\mathrm{Mpc}) + 25 - 2.5\log_{10}(1 + z), \tag{5}$$

where the magnitudes are on the AB-magnitude system, $d_L(z)$ is the cosmological luminosity distance in units of $h^{-1}\mathrm{Mpc}$, and the factor $-2.5\log_{10}(1 + z)$ is from the band shifting of the filter width.

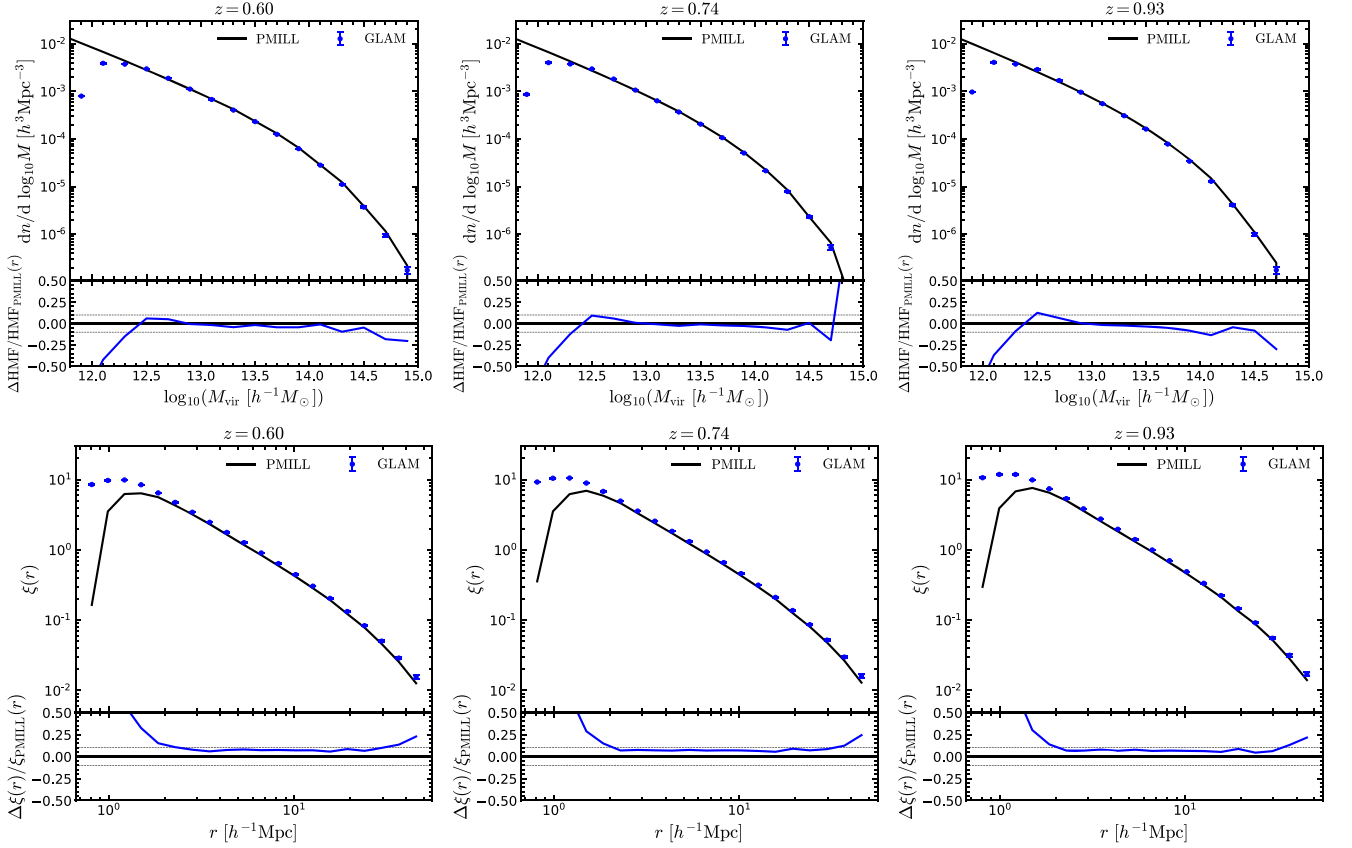The left-hand panel of Fig. 2 shows GALFORM galaxies in the

Figure 1. Top row: differential halo mass function in the PMILL (black solid lines) and the mean of 1000 GLAM simulations (blue dots) as a function of $M_{\rm vir}$. Bottom row: Real-space halo two-point correlation function measured from the PMILL (black solid lines) and the mean of 1000 GLAM simulations (blue dots) for haloes with mass $M_{\rm vir} > 10^{12.5}\,h^{-1}{\rm M}_\odot$. We show measurements at $z = 0.60$ (left-hand column), $z = 0.74$ (middle column), and $z = 0.93$ (right-hand column). Errorbars correspond to the 1σ standard deviation over 1000 GLAM realization. The lower subpanels show the relative difference between the GLAM measurements and that from the PMILL. The horizontal dashed lines indicate a 10 per cent difference.

redshift range $0.6 < z < 1$ in the $(r - z) - (z - W_1)$ colour–colour plane. The black contours show the locus of galaxies with stellar mass in excess of $10^9\,h^{-1}{\rm M}_\odot$ and the red contours show the galaxies that meet the DESI LRG selection criteria set out in equations (2) and (3). The right-hand panel of Fig. 2 shows the distribution of galaxies in the $z - (r - z)$ colour–magnitude plane, again showing all galaxies with stellar mass above $10^9\,h^{-1}{\rm M}_\odot$ (black contours) along with those which satisfy the LRG selection (red contours). The stellar mass cut of $10^9\,h^{-1}{\rm M}_\odot$ is much lower than we expect for the stellar mass of LRGs (see below), but is applied for illustrative purposes, to allow us to see the locus of the GALFORM galaxies in the colour–magnitude planes, before the photometric LRG selection is applied. Note that in these panels, we simply show all of the galaxies that pass the stellar mass cut or LRG selection from *each* of the nine PMILL snapshots that fall within the redshift interval. As such, we are mainly interested in the locus of the GALFORM galaxies in these colour–magnitude planes, rather than the detailed changes in the density of points.

Reassuringly, the red contours in the $(r - z) - (z - W_1)$ colour–colour plane are well within the black polygons denoting the selection boundaries; the blue colour boundary of the $r - z$ versus z selection box is a key component in setting the space density of LRGs, as the red contours touch this cut. At $z = 0.6$, GALFORM predicts that around 6.2 million galaxies in the PMILL volume have stellar mass $M_* > 10^9\,h^{-1}{\rm M}_\odot$ but only a small fraction (0.84 per cent) of these galaxies are selected as LRGs.

Fig. 3 shows the space density, $n(z)$, of DESI LRGs predicted using GALFORM. We have applied the colour–magnitude cuts (equation 1 to 4) to nine PMILL snapshots in the redshift range $0.6 < z < 1$ to obtain the abundance of LRGs – the redshift of the snapshots is indicated by the points in Fig. 3. In the same figure, we show the number density of DESI-like LRGs inferred from observations using photometric redshifts from Zhou et al. (2020a) (black dashed line). We note that GALFORM underpredicts the abundance of LRGs at all redshifts, with the discrepancy reaching a factor of ≈ 1.7 at $z \sim 0.66$. The predicted space densities could be reconciled with those inferred observationally using photometric redshifts by perturbing, for example, the $r - z$ selection to a bluer colour in Fig. 2. This would result in a modest shift in the predicted clustering, a reduction of around 10 per cent over the scales of interest, and without any material change to the HODs predicted by GALFORM. Here, we are interested in showing the theoretical predictions from the GALFORM model and the applications on the large-scale clustering of our GLAM catalogues, rather than making a detailed prediction for LRG clustering, so we proceed without changing the colour selection in this way.

To further investigate the impact of the LRG colour–magnitude selection on the galaxy population predicted by GALFORM, we present in Fig. 4, the stellar mass and luminosity functions for all galaxies and for those selected as DESI LRGs. The top panels of Fig. 4 show the evolution with redshift of the stellar mass function (sMF) for
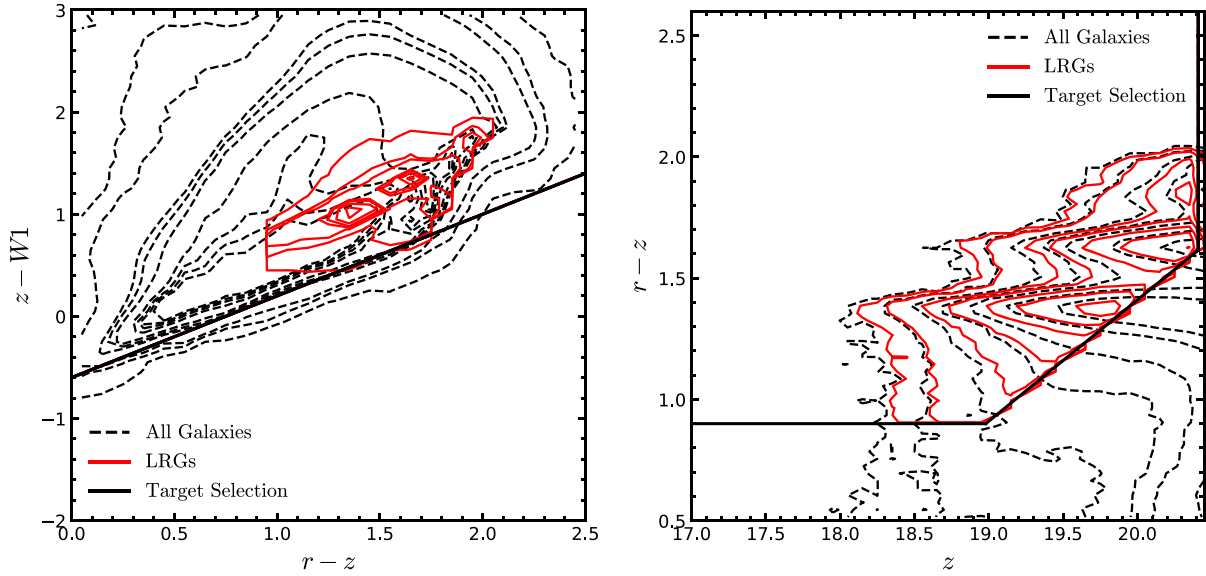
Figure 2. Colour–colour (left-hand) and colour–magnitude (right-hand) diagrams predicted using the GALFORM snapshots at $z = 0.6$ to 1 and using the r, z, and $W1$ bands. Dashed black lines represent the distribution of all galaxies with stellar mass $M_* > 10^9 \, h^{-1} M_\odot$ from the GALFORM output. Red solid lines show the locus of GALFORM galaxies which remain after applying the DESI LRG selection cuts. The solid black polygons indicate the DESI LRG photometric selection given by equations (1)–(4), the same used by Zhou et al. (2020a). For the $(r - z) - (z - W_1)$ panel, the contours display the abundance of objects with values: $\log_{10} n = -8.20, -7.20, -6.69, -5.65, -5.46, -5.20, -4.72, -4.50, -4.35, -4.20$ (M_*-selected galaxies; black dashed lines), and $\log_{10} n = -8.00, -7.40, -6.49, -5.35, -5.16, -5.00, -4.77$ (LRGs; solid red lines). Whereas, the contours in the $z - (r - z)$ panel show values of the number density $\log_{10} n = -8.20, -7.50, -6.90, -6.50, -6.20, -5.79, -5.49, -5.19, -4.89$ (M_*-selected galaxies; black dashed lines), and $\log_{10} n = -7.30, -6.70, -6.30, -6.10, -5.69, -5.39$ (LRGs; solid red lines).
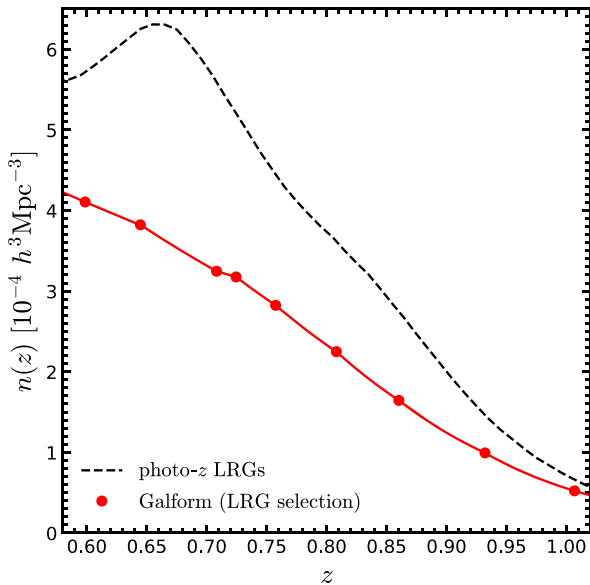

Figure 3. The space density of LRGs meeting the DESI selection criteria, as predicted using GALFORM, as a function of redshift. We show the nine PMILL snapshots between $0.6 < z < 1$ (red dots). The red solid line simply connects the points. The dashed black line shows the space density of DESI-LRGs estimated observationally using photometric redshifts by Zhou et al. (2020a).

all galaxies and for LRGs, for $z = 0.6$, 0.74, and 0.93. Given the halo mass resolution of the PMILL, robust predictions can be made using GALFORM for galaxies with stellar masses $M_* > 10^7 \, h^{-1} M_\odot$ (Baugh et al. 2019). As expected, the LRG sample is dominated

by massive galaxies, although not *all* massive galaxies are LRGs. These massive galaxies are predicted to be in massive dark matter haloes above the mass at which heating by active galactic nuclei suppresses gas cooling (Contreras et al. 2015; Mitchell et al. 2016). Some massive galaxies, however, have recent star formation driven by the cold gas accreted in galaxy mergers, making their $r - z$ colour too blue to be selected as LRGs. The predicted sMF of LRGs drops sharply below $\log_{10}(M_*/h^{-1} M_\odot) = 11.1$, but is similar to the overall SMF for larger stellar masses. The amplitude of the LRG SMF is similar at $z = 0.6$ and 0.74, which reflects the lack of evolution seen in the overall SMF. As we can see from Fig. 3, the number density of LRGs drops from $4.11 \times 10^{-4} \, h^3 \mathrm{Mpc}^{-3}$ at $z = 0.6$ to $3.02 \times 10^{-4} \, h^3 \mathrm{Mpc}^{-3}$ at $z = 0.74$, while at $z = 0.93$ the abundance of LRGs is $0.99 \times 10^{-4} \, h^3 \mathrm{Mpc}^{-3}$.

Similar to the plots showing the galaxy sMF, in the lower panels of Fig. 4, we show, at the same redshifts as used in the top row, the luminosity functions for the r, z, and $W1$ bands for all galaxies and for LRGs. We find a similar trend as that discussed for the sMFs. The fraction of bright galaxies that are selected as LRGs increases with the wavelength of the band: above a threshold luminosity, all galaxies in the W_1-band are LRGS, whereas only a fraction, around a half, of galaxies that are bright in the r-band are LRGs. Below the threshold luminosity, the fraction of galaxies that are LRGs plunges dramatically.

4 THE GALAXY−(SUB)HALO CONNECTION OF DESI LUMINOUS RED GALAXIES

To explore the galaxy–(sub)halo connection of the DESI-like LRGs predicted by GALFORM, we first examine their HOD. The HOD is an useful tool to understand the galaxy–halo connection, clustering and evolution of galaxies in general (see the review by Wechsler & Tinker
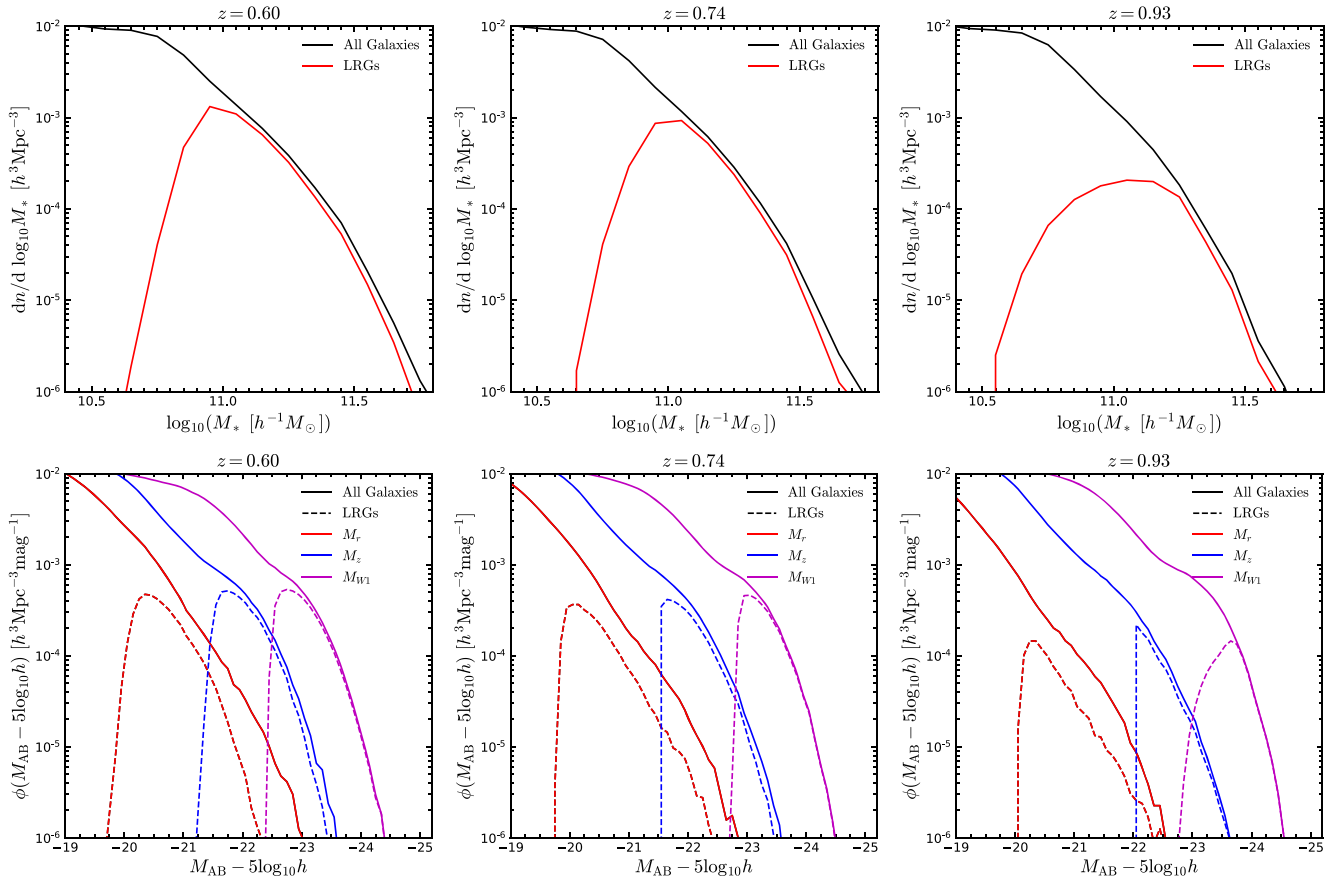
Figure 4. Stellar mass (upper panels) and luminosity (lower panels) functions predicted by GALFORM at $z = 0.6$ (left-hand), $z = 0.74$ (middle), and $z = 0.93$ (right-hand) for all galaxies from the GALFORM output and LRGs. Different colours and line styles indicate different properties and selections as indicated in the legend.

2018). The HOD specifies the average number of galaxies (centrals and satellites) hosted by a dark matter halo. Previous observational studies have described the HOD of LRGs using a functional form that distinguishes between central and satellite galaxies (see e.g. Blake, Collister & Lahav 2008; Brown et al. 2008; Padmanabhan et al. 2009; Zheng et al. 2009). In a traditional HOD, there is a transition in the mean number of central galaxies from $\langle N_c \rangle = 0$ to $\langle N_c \rangle = 1$ with increasing halo mass and the occupation by satellites ($\langle N_s \rangle$) follows a power law in halo mass (Zheng et al. 2005).

Fig. 5 shows the evolution of the HOD of DESI LRGs as predicted by GALFORM in the redshift range $z = 0.6 - 1$. We show the predicted HOD for the nine redshifts we used to measure the evolution of the LRG number density distribution in Fig. 3. At first glance, we see that the occupancy of central galaxies ($\langle N_c \rangle$) does not reach the canonical value of unity at high halo masses, and even begins to decline after a peak at intermediate halo masses. This behaviour is typically seen in the models when galaxies are selected by their star formation rate instead of a property that correlates more closely with stellar mass (Contreras et al. 2013; Cowley et al. 2016; Jiménez et al. 2019). More recently Gonzalez-Perez et al. (2018) found similar behaviour for the HOD of emission-line galaxies selected by the colour–magnitude cuts that will be used by the DESI emission-line galaxy survey (see also Merson et al. 2019 and Gonzalez-Perez et al. 2020). The LRG population is dominated by central galaxies and contains a satellite fraction of $f_{\rm sat} \sim 0.10$–0.04 in the redshift range $= 0.6 - 1$, where the mean number of satellites $\langle N_s \rangle$ is close to a power law.

Fig. 5 shows that there is a clear turnover in the HOD predicted by GALFORM for central galaxies at intermediate redshifts ($z = 0.74$, 0.80). At higher redshifts than this, the trend is less clear due to the evolution in the halo mass function and the resulting lack of high mass haloes. One might have expected that the mean number of centrals would reach unity in massive haloes, due to the suppression of gas cooling through the heating of the hot gas halo by active galactic nuclei. However, some central galaxies in massive haloes can become too blue to be selected as LRGs due to star formation triggered by mergers, which use the cold gas brought in by the merging galaxy.

To develop a deeper understanding of the galaxy–(sub)halo connection, we now explore which *subhaloes* are able to host an LRG. To do so, we consider the number of subhaloes in haloes of different mass and the subhalo mass function, including a version that shows only those subhaloes that host an LRG. We also define a new galaxy sample for comparison purposes by ranking galaxies in order of decreasing stellar mass, and choosing a stellar mass cut to match the number density of the LRG sample. This comparison sample allows us to understand the impact of the selection cuts on the haloes and subhaloes that host LRGs; we call this the stellar mass selected sample.

The upper panels of Fig. 6 show the HOD for the LRG and stellar mass selected galaxy samples, which we compare to the total number of subhaloes available to host an LRG (see below for how this is defined). Focusing on the galaxy HODs first, the black, blue, and
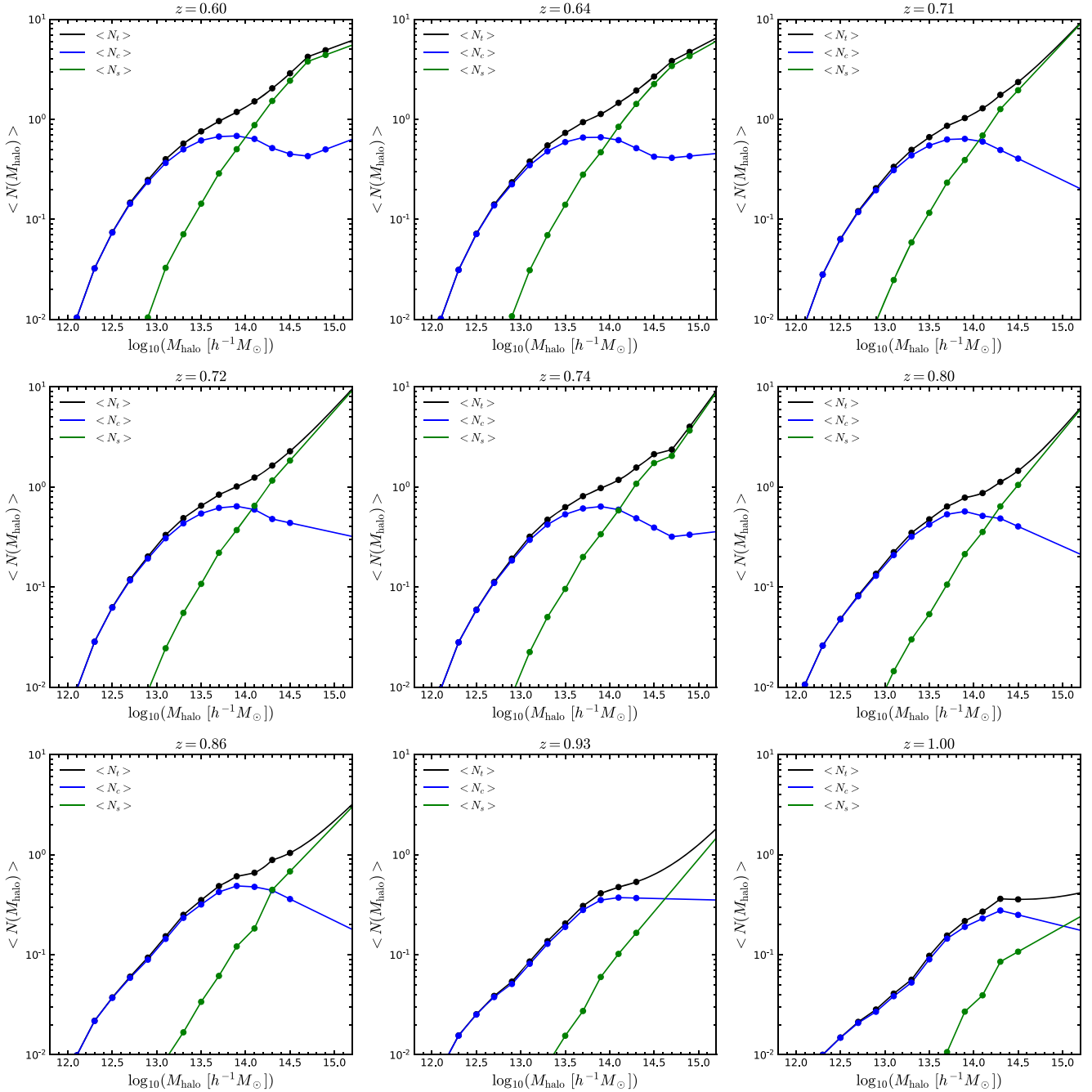
Figure 5. Halo occupation distribution of DESI-like LRGs predicted by GALFORM (symbols) as a function of their host halo masses. Each panel shows a different redshift between $z = 0.6$ and 1 as labelled. The solid lines connect the symbols. Outside the mass range for which model predictions are available, the solid lines show a power-law extrapolation of the HOD for centrals and satellites, based on the last measured points. The total, central, and satellite galaxy occupancy is shown in black, blue, and green, as labelled.

green lines in Fig. 6 show the number, respectively, of all galaxies, central galaxies and satellites galaxies as a function of halo mass; solid lines show the model predictions for the LRG sample and the dashed lines for the stellar mass selected sample. The light blue dashed lines show the number of subhaloes more massive than $M_{\rm subhalo} > 10^{12.5}\,h^{-1}\,{\rm M_\odot}$ as a function of the mass of their main host halo. This mass cut is arbitrary but was chosen because the HODs for the galaxy samples are significant for halo masses above this value. In an illustrative sense, a subhalo mass of $M_{\rm subhalo} \approx 10^{12.5}\,h^{-1}\,{\rm M_\odot}$,

based on the mass coverage of the galaxy sample HODs, could be loosely thought of as the minimum subhalo mass needed to host an LRG or a galaxy in the comparator stellar mass selected sample.

Fig. 6 shows us that only a small fraction of subhaloes with masses above $M_{\rm subhalo} > 10^{12.5}\,h^{-1}\,{\rm M_\odot}$ host an LRG: this fraction reduces from 22 to 8 per cent as the redshift increases from $z = 0.6$ to 0.93. The shape of the total (centrals+satellites) and satellite-only HOD is similar for LRGs and the stellar mass selected sample at $z = 0.6$ and 0.74. However, at $z = 0.93$, the DESI LRG selection cuts
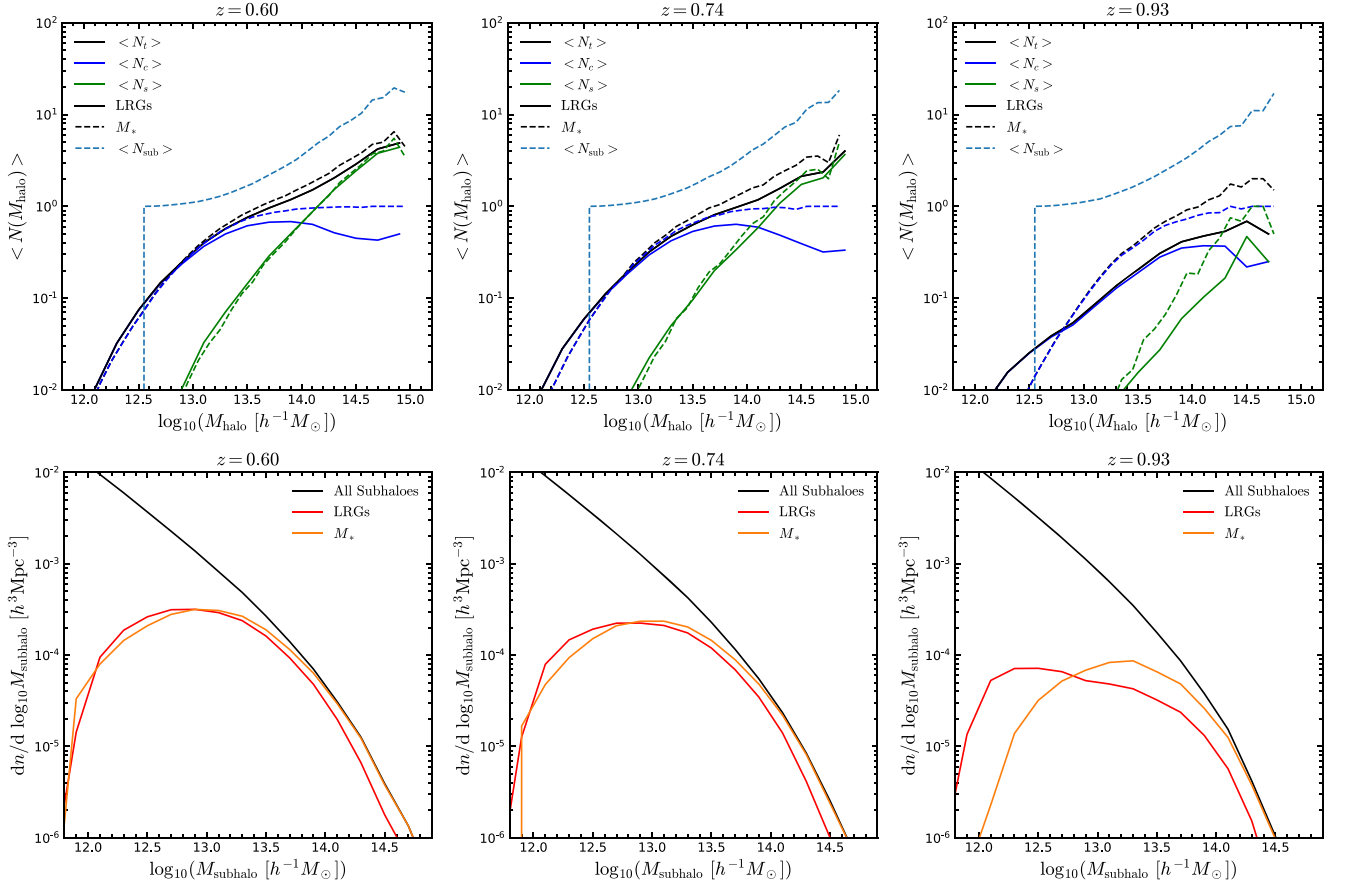
Figure 6. Upper panels: Halo occupation distribution of subhaloes (dashed cyan line), LRGs (solid lines), and galaxies ranked by stellar mass (dashed lines) at $z = 0.6$ (left-hand panel), $z = 0.74$ (middle panel), and $z = 0.93$ (right-hand panel). For galaxies, the occupation of total, centrals, and satellites are specified as black, blue, and green lines, respectively. Lower panels: Subhalo mass functions measured using all galaxies (black solid lines), LRGs (red solid lines), and galaxies ranked by stellar mass (orange solid lines) at $z = 0.6$ (left-hand panel), $z = 0.74$ (middle panel), and $z = 0.93$ (right-hand panel).

modify the form of the LRG HOD away from that of the stellar mass selected sample. The HODs of central galaxies in the two samples are markedly different at all redshifts shown in Fig. 6. The HOD of stellar mass selected central rises to unity with increasing halo mass, but for the LRGs it turns over after reaching a maximum below unity. This behaviour is swamped by the satellite HOD so that the overall HODs for the LRG and stellar mass samples differ less than the central HODs. At the highest redshift shown in Fig. 6, the transition from zero to peak occupancy fraction for centrals is slower for the LRGs than for the stellar mass sample. As centrals dominate the overall sample at lower halo masses, this produces a significant difference in the HOD for LRGs and the stellar mass selected sample.

To gain further insight into the LRG subhalo population, we show the subhalo mass function in the lower panels of Fig. 6. Two versions of the subhalo mass function are shown: one is the 'dark matter view', in which we include all subhaloes and the other is the 'galaxy view', in which case, a subhalo is only included if it contains a galaxy in the sample. If the 'galaxy view' version of the subhalo mass function coincides with the 'dark matter view', then all subhaloes at that mass that could host a galaxy do so. In the case of the stellar mass selected samples shown in the bottom row of Fig. 6, we see that the most massive subhaloes all host a galaxy. As we move to lower masses, the galaxy-view subhalo mass function falls below the dark-matter view version; for these masses only a fraction of

the available subhaloes host a galaxy. Eventually, as we continue to move towards even lower subhalo masses, there is a dramatic downturn in the galaxy-view subhalo mass function, with only a tiny fraction, less than one in a thousand subhaloes hosting a galaxy. Qualitatively, the galaxy-view subhalo mass functions for the LRGs are similar to those for the stellar mass selected sample, with one exception: at the massive end, not all subhaloes host an LRG. This difference becomes more pronounced with increasing redshift. The conclusion of this comparison is that it is essential to perform the full colour–magnitude selection to define the LRG sample. Applying a stellar mass cut to attain a target number density of objects is a fair approximation to performing the full photometric selection at low redshifts, but results in a fundamentally different set of subhaloes being chosen with increasing redshift.

5 GALAXY CLUSTERING

In previous sections, we explored the impact of the DESI LRG colour–magnitude selection on galaxy statistics such as the sMF and the luminosity functions at different wavelengths. We also presented predictions for which haloes and subhaloes contain LRGs. Here, we take this a step further by investigating the evolution of the clustering in configuration and Fourier space, in both real- and redshift-space. We measure the clustering from the simulations with the NBODYKIT toolkit (Hand et al. 2018).

5.1 Galaxy clustering in the PMILL and GLAM simulations

First, we present in Fig. 7 a comparison of the predicted real-space galaxy two-point correlation function for pair separations in the range $0.7 < r/[h^{-1}\text{Mpc}] < 50$ at redshifts 0.6, 0.74, and 0.93 for LRGs and the stellar mass selected sample. Since LRGs do not populate all of the most massive (sub)haloes, as seen in the lower panels of Fig. 6, the LRG sample is less biased than the stellar mass selected one, leading to a smaller clustering amplitude on all scales. We find a constant offset in the clustering amplitude of around 10 per cent between the samples at $z = 0.6$ and 0.74 on all scales. At higher redshifts, where the DESI-LRG colour–magnitude cuts have a bigger impact on which subhaloes host LRGs, we find that the difference in clustering amplitude increases to 50 per cent on large scales, rising to ∼150 per cent on small scales. The larger difference on small scales at $z = 0.93$ is due to the abundance of satellite galaxies in the different galaxy samples; as seen in the upper right-hand panel of Fig. 6, the stellar mass selected sample has a larger satellite fraction than the DESI-LRG sample. This comparison shows that selecting LRGs using stellar mass as a proxy for the full colour–magnitude selection leads to a significant change in the predicted clustering signal.

As we mentioned before, one of our aims is to produce a large number of mock DESI LRG catalogues using the GLAM code to give an accurate estimate of the galaxy clustering signal and its full covariance matrix of errors. For this reason, we populate our 1000 GLAM simulations with LRGs using the tabulated HOD predicted by GALFORM (see Fig. 5), as explained below.

Since GALFORM predicts an HOD for DESI-LRGs that does not appear to follow any of the popular parametric forms in the literature (see appendix A of Contreras et al. 2013), we bypass carrying out a fit altogether and instead use the tabulated model predictions for the HOD directly to populate GLAM haloes with LRGs. Hence, in order to populate a given GLAM halo, we interpolate between the HOD values predicted by GALFORM to the GLAM halo mass (see below for further details). In the case of the most massive haloes, we extrapolate beyond the halo mass range of the HOD values; we do not have robust predictions for these haloes from the PMILL simulation due to its smaller volume compared to the GLAM boxes. This method was used recently by Merson et al. (2019), where the authors extracted the HOD of Hα galaxies from the GALACTICUS SAM catalogue (Benson 2012; Merson et al. 2018), and used this to populate the Millennium-MXXL halo light-cone from Smith et al. (2017).

In detail our HOD method is as follows. We assign a central galaxy to a GLAM halo if $\langle N_c \rangle > \mathcal{U}(0, 1)$, where $\langle N_c \rangle$ is the mean number of central galaxies that could be found in a GLAM halo and $\mathcal{U}(0, 1)$ is a uniform random number between 0 and 1. Recall that the GALFORM predictions for the HOD of central galaxy LRGs never reach unity. We place the central galaxy at the centre of mass of the host halo, and give it the velocity of the centre of mass. The number of satellite galaxies is drawn from a Poisson distribution with mean equal to $\langle N_s \rangle$, as derived from the tabulated HOD predicted using GALFORM. Satellite galaxies are radially distributed within the virial radius, $(0 < r < R_{\text{vir}})$, following a Navarro–Frenk–White (NFW) density profile (Navarro, Frenk & White 1996, 1997), with a uniform angular distribution. The satellite is assigned a velocity that is made up of the halo velocity plus a perturbation along the x, y, and z coordinates drawn from a Gaussian distribution with variance equal to the 1D velocity dispersion of the host halo.

We measure the real- and redshift-space clustering in configuration and Fourier space from the GLAM-HOD catalogues and compare

these with their PMILL counterparts to corroborate the precision of our method. In addition, the real-space clustering measurements provide us a relation between the distribution of galaxies and the underlying dark-matter density field via the galaxy bias (Peebles 1980). The galaxy bias is directly measured from our GLAM LRG mocks as

$$b(k, z) = \sqrt{\frac{P_g(k, z)}{P_m(k, z)}} \quad \text{or} \quad b(r, z) = \sqrt{\frac{\xi_g(r, z)}{\xi_m(r, z)}}, \tag{6}$$

where $P_g(k, z)(\xi_g(r, z))$ and $P_m(k, z)(\xi_g(r, z))$ are the real-space galaxy and dark matter power spectra (correlation functions) at a given redshift, respectively. We tried both approaches, using the power spectrum and correlation function to estimate the bias. We determine the galaxy bias on scales $k < 0.1\,h\,\text{Mpc}^{-1}$ for the power spectrum and $r > 3\,h^{-1}\text{Mpc}$ for the correlation function and found consistent answers, $b(z) = 1.84, 1.96, 2.06$ at $z = 0.6$, 0.74, and 0.93, respectively. The DESI-like LRG bias has been estimated from the measured angular power spectrum and from the halo model of the photo-z LRGs giving the following relations, $b(z) = 1.6/D(z)$ (Kitanidis et al. 2019) and $b(z) = 1.5/D(z)$ (Zhou et al. 2020a). Note these relations are slightly different to the value of $b(z) = 1.7/D(z)$ reported in DESI Collaboration et al. (2016), where $D(z)$ is the linear growth factor at redshift z, with $D(z = 0) = 1$. For the cosmological parameters used in the PMILL simulation, the linear growth factor is $D(z = 0.6) = 0.73$, $D(z = 0.74) = 0.69$, and $D(z = 0.93) = 0.63$, which means that the values we recover for the bias are slightly lower than those inferred from the observations, more similar to $b(z) = (1.3 - 1.4)/D(z)$. We also note that Zhai et al. (2017) estimated the bias for a slightly different LRG sample, finding $b = 2.3$ for $z \sim 0.7$.

We use the distant-observer approximation to shift the positions of galaxies from real- to redshift-space, treating the z-axis as the line of sight

$$\mathbf{s} = \mathbf{r} + \frac{(1 + z)v_z}{H(z)}\hat{e}_z, \tag{7}$$

where $\mathbf{r}$ is the coordinate vector in real space, $\mathbf{s}$ is the equivalent of this in redshift-space, and z is the redshift of the simulation snapshot used to generate the galaxy catalogue. $H(z)$ is the Hubble parameter, v_z and $\hat{e}_z$ are the components of the velocity and the unit vector along the z-direction.

We measure the monopole and quadrupole moments of the redshift-space correlation function, $\xi_l(s)$, and power spectrum, $P_l(k)$, using

$$\xi_l(s) = (2l + 1) \int_0^1 \xi(s, \mu)\mathcal{L}_l(\mu)\,d\mu, \tag{8}$$

$$P_l(k) = (2l + 1) \int_0^1 P(k, \mu)\mathcal{L}_l(\mu)\,d\mu, \tag{9}$$

where $\xi(s, \mu)$ and $P(k, \mu)$ are the full 2D correlation function and power spectrum, μ is the cosine of the angle between the separation vector, $\mathbf{s}$ or $\mathbf{k}$, and the line of sight in configuration or Fourier space, respectively. The $\mathcal{L}_l(\mu)$ are the Legendre polynomials where $l = 0$ is the monopole and $l = 2$ is the quadrupole. We use 20 bins logarithmically spaced over the separation range $0.7 < s/[h^{-1}\text{Mpc}] < 50$ in which to measure the correlation function. The power spectrum is measured in the range $0 < k/[h\,\text{Mpc}^{-1}] < k_{\text{Nyq}}$ using linear bins in k with separation $\Delta k = 0.006\,h\,\text{Mpc}^{-1}$, where $k_{\text{Nyq}} = \pi N_{\text{mesh}}/L_{\text{box}}$ is the 1D Nyquist frequency, $N_{\text{mesh}} = 512$ and L_{box} is the box size of the PMILL or GLAM simulations. In all cases, we adopt 30 linearly spaced bins between 0 and 1 for μ.
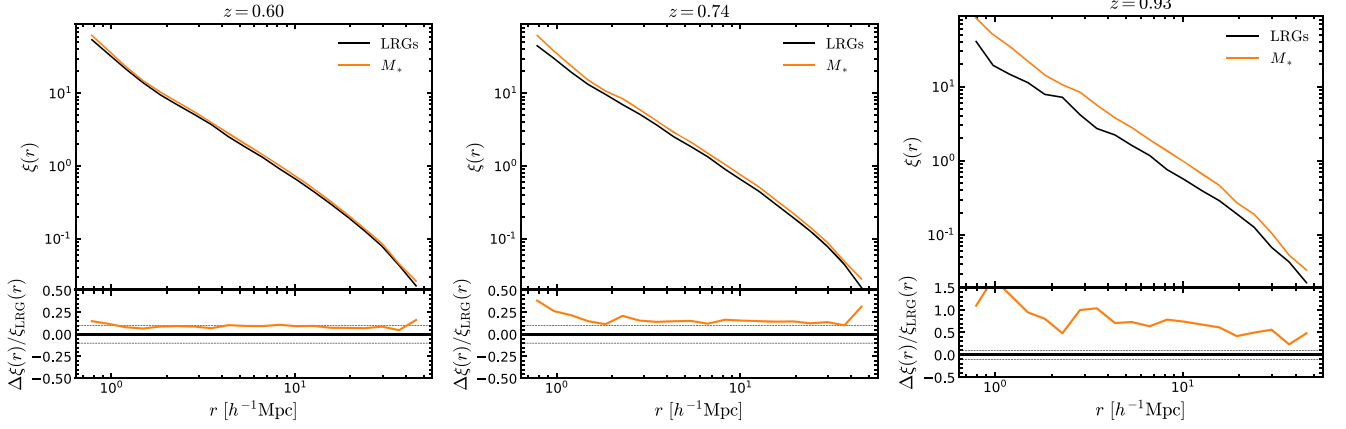
Figure 7. The real-space galaxy correlation functions predicted by GALFORM for LRGs (red lines) and the stellar mass selected sample (orange line) at $z = 0.6$ (left-hand panel), $z = 0.74$ (middle panel), and $z = 0.93$ (right-hand panel). The lower subpanels show the relative difference between the LRG and the stellar mass selected samples. The horizontal dashed lines indicate a 10 per cent difference.
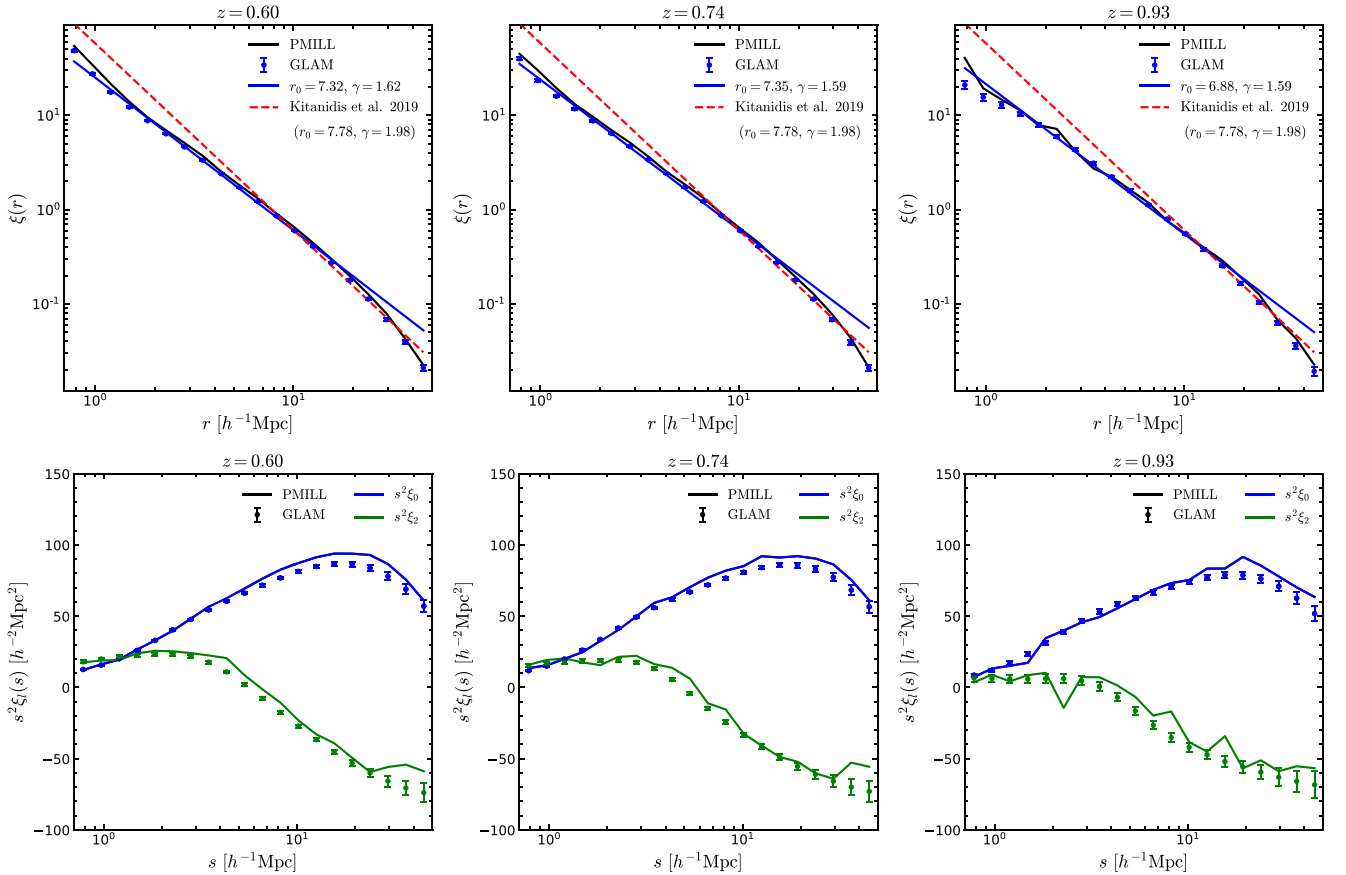

Figure 8. Upper panels: Real-space galaxy correlation function of the GALFORM-PMILL LRGs (black lines) and the HOD-GLAM LRGs (blue symbols with error bar). We also show the best-fitting power-law form, $\xi(r) = (r/r_0)^{-\gamma}$, to the DESI-LRG measurements reported by Kitanidis et al. (2019, red dashed lines) and to our measurements (blue solid lines). Lower panels: Redshift-space monopole and quadrupole moments of the correlation function for GALFORM-PMILL LRGs (solid lines) and GLAM-HOD LRGs (symbols with error bar). In the case of the GLAM-HOD LRGs measurements, we show the mean and standard deviation over 1000 realization. The measurements are made at $z = 0.6$, 0.74, and 0.93, as labelled at the top of each panel.

In the upper panels of Fig. 8, we display the real-space clustering measured from the GALFORM output (black line) and the GLAM LRG mock catalogues (blue symbols with error-bars). Additionally, we show the best-fitting power-law fit to the correlation function reported by Kitanidis et al. (2019) ($r_0 = 7.78\,h^{-1}\mathrm{Mpc}$, $\gamma = 1.98$) which agrees well with our measurements, especially on scales $r \geq r_0$. Note that Kitanidis et al. fitted the angular correlation function in the range $0.001° <$

$\theta < 1°$ which translates to comoving separation $\theta_{\min} D_A(z) < r/[h^{-1}\text{Mpc}] < \theta_{\max} D_A(z)$, where D_A is the angular-diametre distance. We also show results when fitting our GLAM measurements with a power law using the range mentioned above, $\xi(r) = (r/r_0)^{-\gamma}$, finding $r_0/[h^{-1}\text{Mpc}] = (7.316\pm0.022, 7.346\pm0.024, 6.883\pm0.04)$ and $\gamma = (1.623 \pm 0.006, 1.592 \pm 0.007, 1.589 \pm 0.012)$ at $z = 0.6$, 0.74, and 0.93.

The lower panels of Fig. 8 shows the predicted multipoles of the redshift-space correlation function of the GLAM-HOD LRGs (symbols with errorbars), plotted in comparison with their GALFORM counterparts (solid lines). We find excellent agreement between the clustering measured in both real- and redshift-space for the GLAM and GALFORM LRGs at all scales and all redshifts.

In Fig. 9, we display the clustering measurements in Fourier space. First, we note the good agreement between the GALFORM and GLAM measurements on all scales. In the upper panels of Fig. 9, we also show the measured dark-matter power spectrum scaled by the galaxy bias squared relations of Kitanidis et al. (2019, cyan lines), Zhou et al. (2020a, magenta lines), and from our simulations, equation (6). We find that our measurements slightly underpredict the bias value compare to the measured relations estimated by Kitanidis et al. (2019) and Zhou et al. (2020a). In the lower panels of Fig. 9, we show the multipole moments of the redshift space power spectrum, finding almost perfect agreement between the GALFORM and GLAM measurements on scales $k > 0.1\,h\,\text{Mpc}^{-1}$. Nevertheless, there is a noisy signal for the GALFORM quadrupole of the redshift-space power spectrum, due to the smaller box size of the PMILL. Nevertheless, this signal is in good agreement with the predictions from GLAM over the range $0.1 < k/[h\,\text{Mpc}^{-1}] < 0.3$.

We conclude that populating GLAM haloes using our interpolated-HOD method reproduces accurately the clustering of LRGs predicted directly by GALFORM on all scales of interest.

5.2 Large-scale galaxy clustering and covariance matrices

In general, it is not possible to measure the 3D clustering of galaxies in real-space from observations. Some compromise involving projection is usually required to obtained a real-space statistic, such as the angular correlation function or the projected correlation function. The most direct 3D clustering measurements from surveys provide statistics in redshift-space, which are affected by peculiar velocities. Moreover, future surveys like DESI aim to measure galaxy clustering on scales up to $\sim 200\,h^{-1}\text{Mpc}$. Hence, taking advantage of our GLAM-HOD machinery, here, we present predictions for the large-scale galaxy clustering and covariance matrices of DESI-like LRGs for the correlation function and power spectrum. These quantities are fundamental for error estimates on the measurements of BAO and RSD (see e.g. Alam et al. 2017).

In the following, we focus our attention on the large-scale clustering of DESI-like LRGs for pair separations in the range $0 < s/[h^{-1}\text{Mpc}] < 150$ for the correlation function. For the power spectrum, we show results in the wavenumber range $0.01 < k/[h\,\text{Mpc}^{-1}] < 0.3$.

The upper panels of Figs 10 and 12 display the mean and standard deviation of the multipoles of the correlation function and the power spectrum calculated over 1000 GLAM DESI-like LRGs realization at $z = 0.6$, 0.74, and 0.93. We also measure the covariance matrix, **C**, of each estimator **E**, as follows

$$C_{ij} = \frac{1}{N_s - 1} \sum_{k=1}^{N_s} \left(E_i^k - \bar{E}_i\right)\left(E_j^k - \bar{E}_j\right), \qquad (10)$$

where $N_s = 1000$ is the number of mocks, $\bar{E}_i = 1/N_s \sum_k E_i^k$ is the mean value of the estimator in the i-th separation bin, and E_i^k is the corresponding measurement from the k-th mock. The standard deviation is estimated from the diagonal elements of the covariance matrix,

$$\sigma_i = \sqrt{C_{ii}}. \qquad (11)$$

We show the diagonal error contribution, σ_{E_i}/E_i, of the moments of the correlation function and power spectrum in the lower subpanels of the upper row of Figs 10 and 12. We observe an increase in the size of the error contribution at large scales, especially for the monopole and quadrupole in configuration space.

We display the correlation matrix

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}, \qquad (12)$$

in the middle (monopole) and bottom (quadrupole) panels of Figs 10 and 12 for the correlation function and power spectrum, respectively. The diagonal and non-diagonal components have different magnitudes and evolve differently with redshift. Figs 11 and 13 show cuts through the correlation matrices corresponding to our measurements in configuration and Fourier space, respectively. These diagrams help us to better display the level of correlation and the structure of the matrices. In the case of the moments of the correlation function (Fig. 11), we show the cuts at four different separation bins, $s_i = (37.5, 72.5, 107.5, 142.5)\,h^{-1}\text{Mpc}$, while in Fourier space (Fig. 13), we use $k_i = (0.081, 0.154, 0.222, 0.289)\,h\,\text{Mpc}^{-1}$. We see a strong correlation between the bins close to the diagonal elements in the monopole and quadrupole of the correlation function at $z = 0.60$ and 0.74; this correlation becomes weaker at $z = 0.93$ (Fig. 11). In the case of the multipoles of the power spectrum, the off-diagonal elements are much less correlated than the diagonal components, with values close to zero (Fig. 13). This trend is strongest for the quadrupole of the power spectrum.

Klypin & Prada (2018) carried out an extensive study of the covariance and correlation matrix associated with the dark-matter power spectrum of GLAM simulations. Our results for the estimation of errors from the GLAM-HOD catalogues extends the work of Klypin & Prada to galaxies and to the correlation function. In detail, Fig. 14 shows the covariance analysis of the real-space DM and LRG power spectra. We summarize our findings as follows. First, in the upper panels, we display the measurements from our simulations, we observe that the size of the error is similar for both DM and LRGs at large-scales ($k < 0.05\,h\,\text{Mpc}^{-1}$) but on smaller scales the amplitude of the error of the galaxy power spectrum becomes larger with increasing redshift. We also show the DM power spectrum and its errors scaled by the LRG bias squared (see Section 5.1 for details) as a blue solid line (with a shaded region showing the 1σ error) in the upper panel of the first row of Fig. 14. Secondly, the correlation matrices are shown in the middle panels (upper middle panels for DM and lower middle panels for LRGs), we find that the amplitude of the DM correlation matrices are consistent with those reported by Klypin & Prada (2018). On the other hand, the correlation amplitude of the LRG power spectrum is similar to its analogue in redshift space (see middle panels of Fig. 12). Lastly, the evolution of the non-diagonal terms of the correlation matrices are displayed in the bottom panels of Fig. 14. We compare the level of correlation at four values of the separation bin, $k_i = (0.081, 0.154, 0.222, 0.289)\,h\,\text{Mpc}^{-1}$, finding a more complex behaviour from the LRGs correlation matrices with an increase amplitude at small scales, this behaviour is also consistent with our findings in redshift space (see Fig. 12). Moreover, the amplitude of the non-diagonal elements are similar for both DM
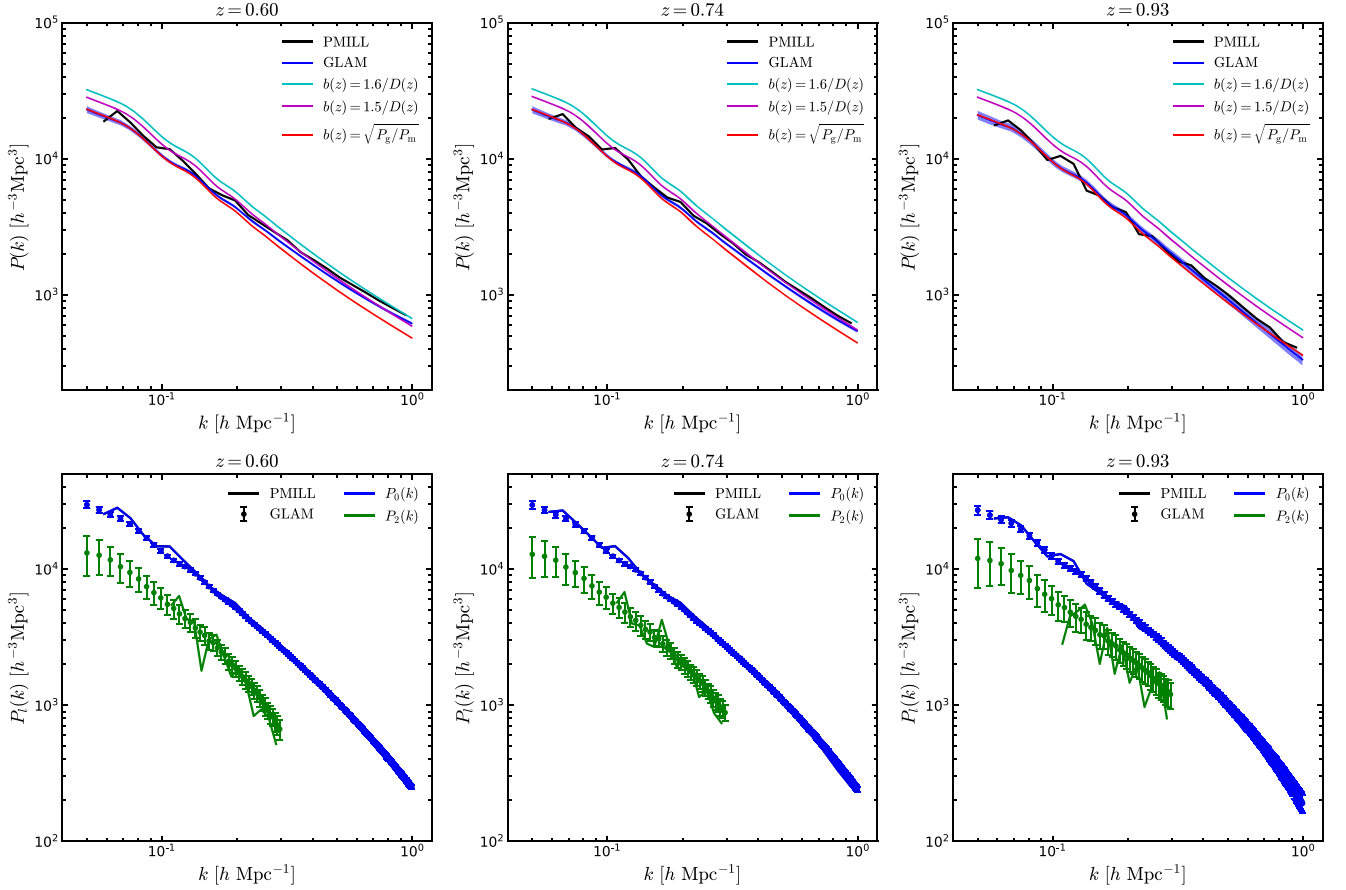
Figure 9. Upper panels: Real-space galaxy power spectrum of the GALFORM-PMILL LRGs (black lines) and the GLAM-HOD LRGs (blue lines, the shaded region represents the 1σ error over 1000 realization), we also show the dark-matter power spectrum multiplied by the galaxy bias squared estimations of Kitanidis et al. (2019, cyan solid lines), Zhou et al. (2020a, magenta solid lines), and from equation (6) (red solid line). Lower panels: Redshift-space monopole and quadropole moments of the power spectrum for GALFORM LRGs (solid lines) and GLAM LRGs (symbols with error bar). For the GLAM-HOD LRGs measurements, we show the mean and standard deviation over 1000 realization. The measurements are made at $z = 0.6$, 0.74, and 0.93, as labelled at the top of each panel.

and LRGs at $z = 0.93$ (bottom right-hand panel of Fig. 14). For all redshifts, we find that the off-diagonal terms of the LRGs correlation matrices are sizeable compared to their DM counterparts.

Finally, we can use the covariance matrix of each estimator to define a chi-squared to find the best-fitting cosmological parameters as follows

$$\chi^2 = \sum_{i,j=1}^{N_s} \left(E_i^{\text{th}} - E_i^{\text{obs}} \right) C_{ij}^{-1} \left(E_j^{\text{th}} - E_j^{\text{obs}} \right), \tag{13}$$

where C_{ij}^{-1} is the inverse of the covariance matrix, equation (10), E^{th} is the theoretical expectation of the estimator that depends on the cosmological parameters, and E^{obs} is the measured estimator from our GLAM-HOD catalogues. This definition is used in Sections 5.3 and 5.4.

It is instructive to compare the errors we obtain in the GLAM simulation boxes with the errors expected in the DESI measurements. DESI will measure the clustering of LRGs in a series of redshift shells over a solid angle of $14\,000\,\text{deg}^2$. We anticipate that DESI will sample a comoving volume of $V/[\text{h}^{-3}\text{Gpc}^3] = 2.63$, 3.15, and 4.10, respectively on bins centred at redshifts of $z = 0.65$, 0.75, and 0.95 (DESI Collaboration et al. 2016). Hence, to get a rough impression of how our error estimates (on the cosmological parameters) will scale to those expected for DESI, we can scale the

GLAM errors by the square root of the inverse volume ratio (e.g. Feldman, Kaiser & Peacock 1994): $\sigma'(z) = \sigma(z)\sqrt{V_{\text{GLAM}}/V_{\text{DESI}}(z)}$, where $V_{\text{GLAM}} = 1\,\text{h}^{-3}\text{Gpc}^3$. Note that we refrain from carrying out a more detailed comparison with the errors reported in DESI Collaboration et al. (2016), as there are a number of differences in the estimation of the errors in that study and ours. For example, the DESI Collaboration study uses a Fisher matrix method, which assumes Gaussian errors and neglects off-diagonal terms. They also employ multiple tracers, considering ELGs and quasars, as well as LRGs, and apply reconstruction techniques to sharpen the BAO signal.

5.3 Linear redshift-space distortions

In large volume galaxy surveys, we can extract information about the growth of structure through the linear growth rate, f, which is defined as the logarithmic derivative of the linear growth function of density perturbations, D, with respect to the scale factor, a

$$f \equiv \frac{\text{d}\ln D}{\text{d}\ln a}. \tag{14}$$

In linear perturbation theory, the relation between the redshift-space galaxy power spectrum, P_s, and its real-space counterpart, P_r, is given by Kaiser (1987)

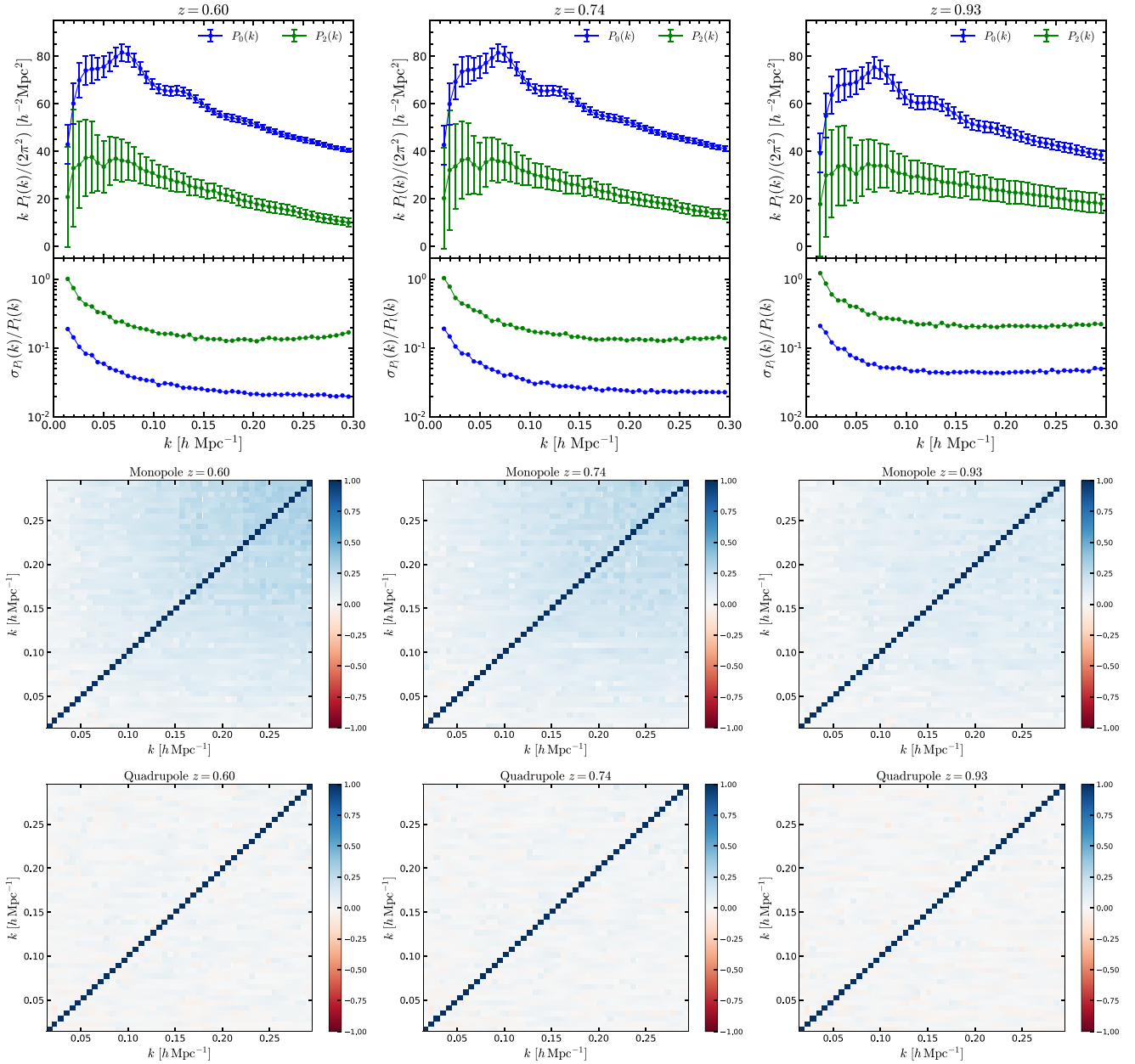$$P_s(k, \mu) = (1 + \beta\mu^2)^2 P_r(k). \tag{15}$$

Figure 10. Upper panels: Measured monopole (blue lines) and quadrupole (green lines) of the redshift-space two-point correlation function of DESI-like LRGs from our GLAM-HOD catalogues, the lower subpanels show the fractional error on the monopole and the quadrupole. The error is calculated using equation (11). Middle panels: Correlation matrix, equation (12), of the monopole. Bottom panels: Correlation matrix, equation (12), of the quadrupole. The colour bars in the correlation matrices display values from $-1 \le R(s_i, s_j) \le 1$. The measurements are made at $z = 0.6$, 0.74, and 0.93, as labelled at the top of each panel.

From equation (15), we can see that the amplitude of the RSD is related to the distortion parameter β, defined as

$$\beta(z) \equiv \frac{f(z)}{b(z)}, \tag{16}$$

where f is the linear growth rate (equation 14) and b is the linear galaxy bias both of which vary with redshift, equation (6).

The monopole and quadrupole moments of the power spectrum can be estimated from equations (9) and (15)

$$P_0(k) = \left(1 + \frac{2\beta}{3} + \frac{\beta^2}{5}\right) P_{\rm r}(k), \tag{17}$$

$$P_2(k) = \left(\frac{4\beta}{3} + \frac{4\beta^2}{7}\right) P_{\rm r}(k), \tag{18}$$

where $P_{\rm r}(k)$ is galaxy power spectrum in real-space.

On the other hand, the redshift-space correlation function can be expressed as follows (Hamilton 1992, 1998)

$$\xi(s, \mu) = [1 + \beta(\partial/\partial z)^2(\nabla^2)^{-1}]^2 \xi(r), \tag{19}$$

In linear theory, the monopole and quadrupole of the correlation function can be estimated using (Hamilton 1992), i. e.

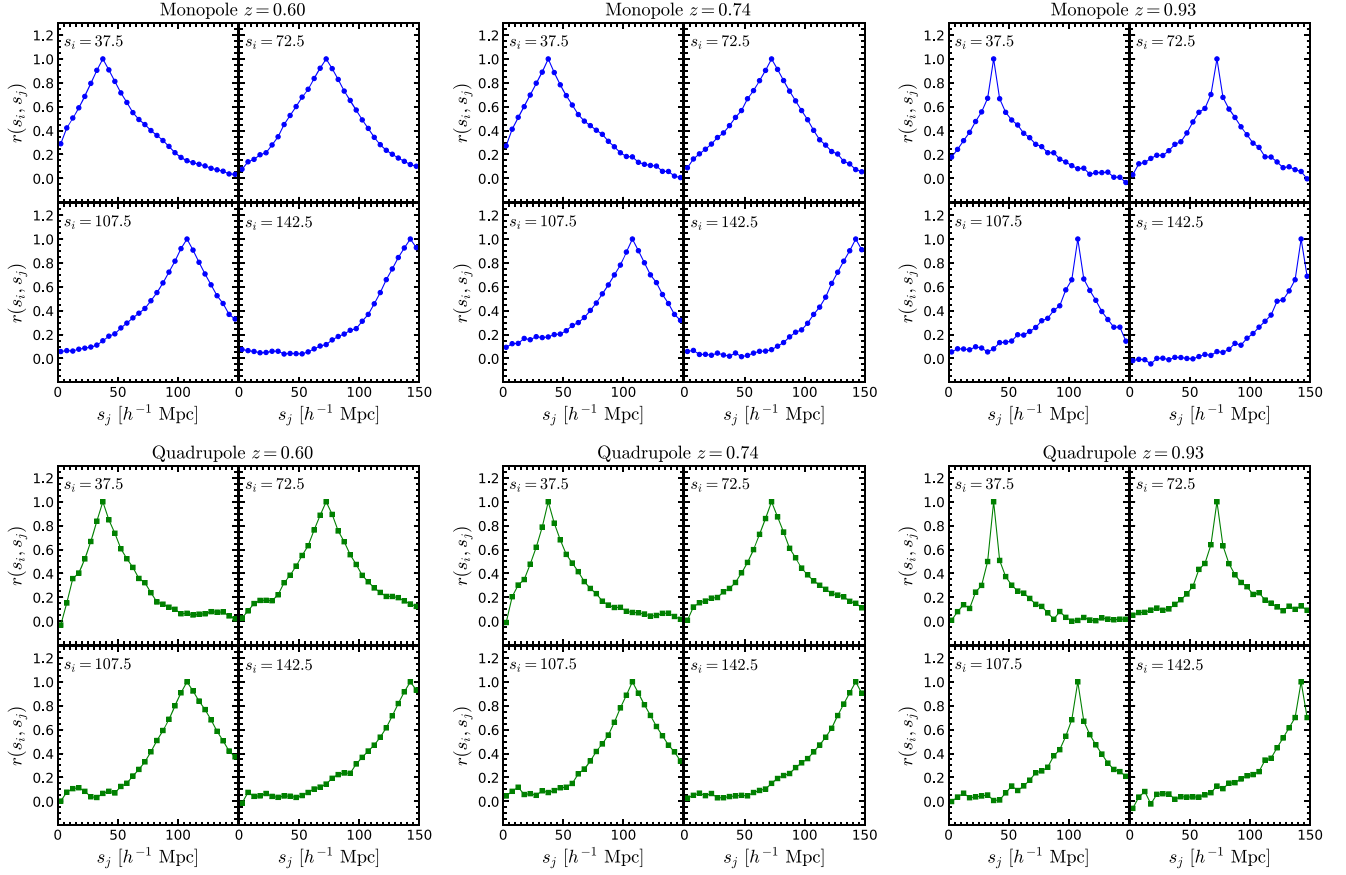$$\xi_0(s) = \left(1 + \frac{2\beta}{3} + \frac{\beta^2}{5}\right) \xi(r), \tag{20}$$

Figure 11. Cuts through the correlation matrix of the monopole (upper panels) and the quadrupole (bottom panels) of the redshift-space correlation function at four different values of s_i in units of $[h^{-1}\mathrm{Mpc}]$ as indicated in the panels. The measurements are made at $z = 0.6$, 0.74, and 0.93, as labelled at the top of each panel.

$$\xi_2(s) = \left(\frac{4\beta}{3} + \frac{4\beta^2}{7}\right)[\xi(r) - \bar{\xi}(r)], \tag{21}$$

where $\xi(r)$ is the galaxy correlation function in real-space and $\bar{\xi}$ is its volume integral out to pair separation r

$$\bar{\xi}(r) = \frac{3}{r^3}\int_0^r \xi(r')r'^2\,\mathrm{d}r'. \tag{22}$$

From equations (17)−(18) and equations (20)−(21), we can define two estimators to obtain the distortion parameter, β, or the linear growth rate, f, (Cole, Fisher & Weinberg 1994; Hawkins et al. 2003),

$$R(k/s) = \frac{P_0(k)}{P_r(k)} = \frac{\xi_0(s)}{\xi(r)} = 1 + \frac{2\beta}{3} + \frac{\beta^2}{5}, \tag{23}$$

and

$$Q(k/s) = \frac{P_2(k)}{P_0(k)} = \frac{\xi_2(s)}{\xi_0(s) - \bar{\xi}_0(s)} = \frac{(4/3)\beta + (4/7)\beta^2}{1 + (2/3)\beta + (1/5)\beta^2}, \tag{24}$$

where $\mathcal{F}(k/s)$ indicates that the quantity $\mathcal{F}$ can be a function of k or s and

$$\bar{\xi}_0(s) = \frac{3}{s^3}\int_0^s \xi_0(s')s'^2\,\mathrm{d}s', \tag{25}$$

is the volume average of the monopole in redshift space, the analogue of equation (22).

Fig. 15 shows our measurements of the $R(k/s)$, and the $Q(k/s)$ estimators from our DESI-GLAM LRG mock catalogues at the

median redshift $z = 0.74$. The black dashed line in each panel corresponds to the linear theory predictions. From the measurements in Fourier space (left-hand panels of Fig. 15), we can see that both estimators become closer to the linear theory predictions at scales $k \lesssim 0.1\,h\,\mathrm{Mpc}^{-1}$, this means that linear theory is only valid on sufficiently large scales. On small scales, where the non-linear motions of galaxies dominate, we observe a downturn in the signal of each estimator. The trend is similar in configuration space (right-hand panel of Fig. 15), where we observe that the linear theory limit is reached on scales $s > 20\,h^{-1}\mathrm{Mpc}$. All panels in Fig. 15 show the same range of values on the vertical axis, allow us to see that the errors are slightly different in Fourier and configuration space, especially in R.

To extract the linear growth rate, f, from our measurements, we perform a likelihood analysis by minimizing χ^2 defined by equation (13) by fitting the measurements of $R(k/s)$, equation (23), and $Q(k/s)$, equation (24) over the range of scales $k < 0.1\,h\,\mathrm{Mpc}^{-1}$ in Fourier space and $s > 20\,h^{-1}\mathrm{Mpc}$ in configuration space. We fix the galaxy bias and just allow the linear growth rate to vary. To do so, we employ the Monte Carlo Markov Chain (MCMC) technique implemented in the EMCEE PYTHON package (Foreman-Mackey et al. 2013).

In Fig. 16, we compare the predictions for the linear growth rate $f(z)$ obtained from our DESI-GLAM LRG mocks at $z = 0.6$, 0.74 and 0.93 with the current observational measurements from large galaxy surveys, including 6dFGRS at $z = 0.067$ (Beutler et al. 2012),
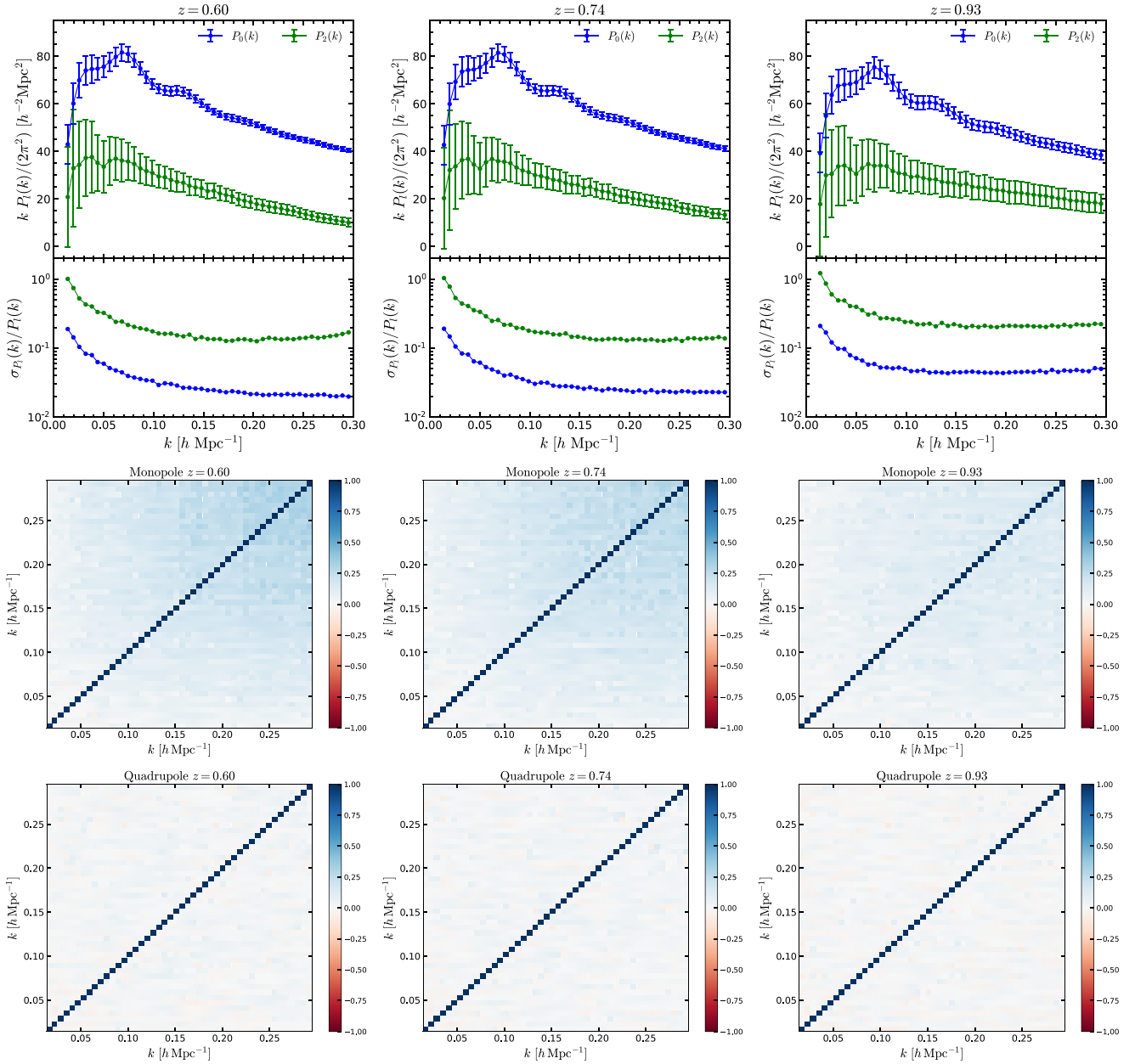
Figure 12. Same as Fig. 10, but for the multipoles of the power spectrum.

SDSS MGS at $z = 0.15$ (Howlett et al. 2015), 2dFGRS at $z = 0.17$ (Percival et al. 2004), GAMA at $z = 0.18$ and 0.38 (Blake et al. 2013), WiggleZ at $z = 0.22, 0.41, 0.6,$ and 0.78 (Blake et al. 2011a), BOSS DR12 at $z = 0.32, 0.51$ and $z = 0.61$ (Alam et al. 2017), eBOSS LRGs at $z = 0.698$ (Bautista et al. 2021; Gil-Marín et al. 2020), FastSound at $z = 1.4$ (Okumura et al. 2016) and the eBOSS DR14 QSO sample at $z = 1.52$ (Zarrouk et al. 2018). The black errorbars over the ΛCDM predictions indicate the estimated error from the DESI forecast (see table 2.3 of DESI Collaboration et al. 2016). Note that in this work, we do not include the light-cone and survey geometry effects on our mocks. These will be considered in a forthcoming paper.

Table 1 summarizes the best-fitting values of the linear growth rate, f, at $z = 0.60, 0.74,$ and 0.93 obtained from the estimators R (equation 23) and Q (equation 24) in configuration and Fourier

space. We also show the values from the fidiciual cosmology. We find very good agreement between our estimations and the theoretical predictions. The largest errors come from the R estimator in configuration space and Q in Fourier space, this might be due to the size of the error contribution of our measurements (see Section 5.2 for details). The best case is R in Fourier case, which estimates the linear growth rate with a precision better than 4 per cent.

It is expected that DESI will provide a means to distinguish between gravity models. For this reason, in Fig. 16, we also show the theoretical expectations from two representative modified gravity models: the $f(R)$ Hu–Sawicki model (Hu & Sawicki 2007) and the normal branch of the DGP model (nDGP; Dvali, Gabadadze & Porrati 2000). Previously, Hernández-Aguayo et al. (2019) presented predictions for the linear and non-linear RSDs in configuration space
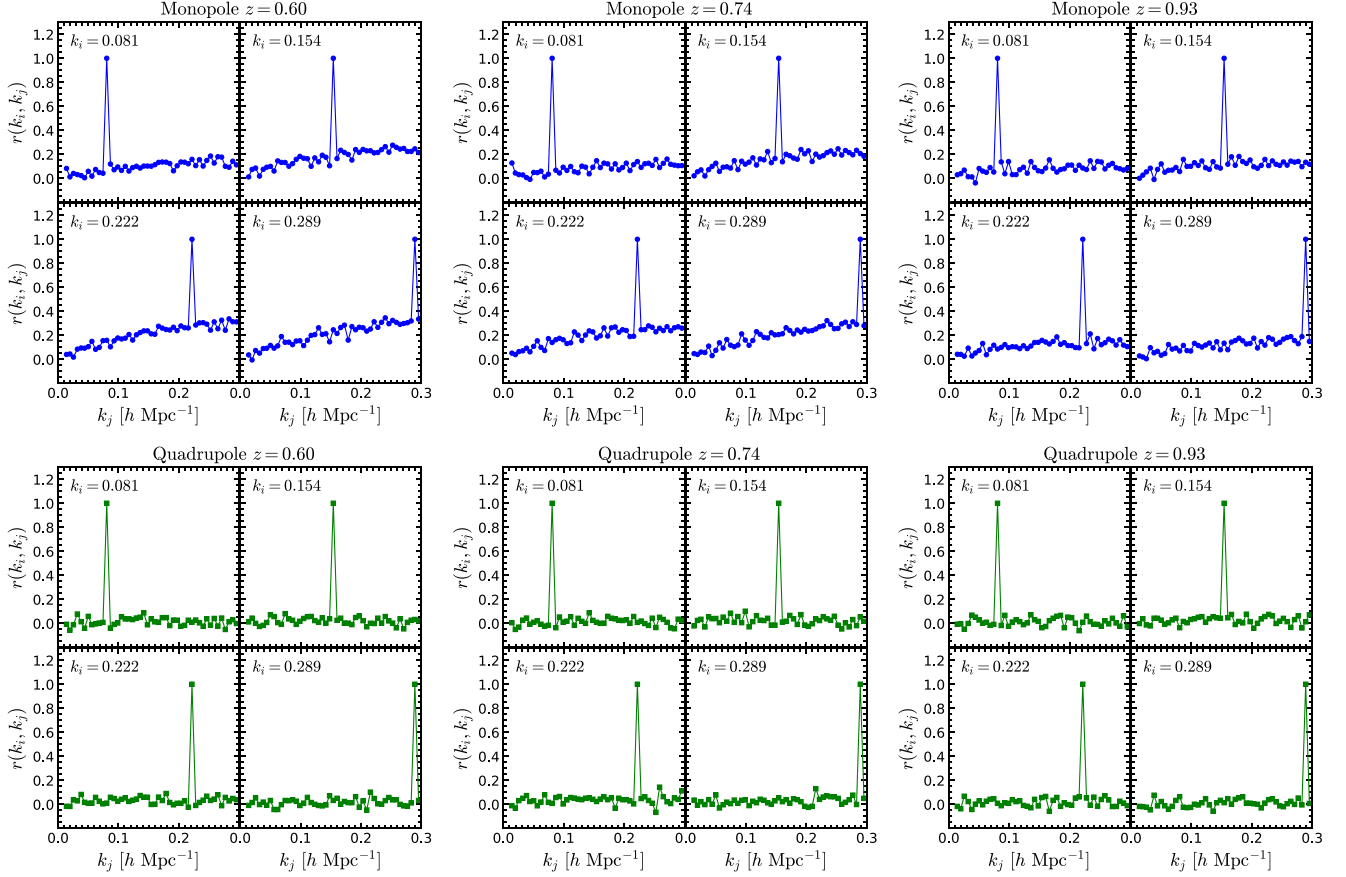
Figure 13. Same as Fig. 11, but for the multipoles of the power spectrum. In this case, we show the cuts through the correlation matrices at four different values of k_i in units of [h Mpc^{-1}].

for these models but for the BOSS-CMASS sample at $z \leq 0.5$ (Manera et al. 2012).

The linear growth for the matter fluctuations in these gravity models can be obtained by solving the equation of the linear growth factor, D,

$$D'' + \left(2 - \frac{3}{2}\Omega_{\mathrm{m}}(a)\right) D' - \frac{3}{2}\frac{G_{\mathrm{eff}}}{G}\Omega_{\mathrm{m}}(a)D = 0, \tag{26}$$

where $'$ denotes a derivative with respect to $\ln a$ and G_{eff} takes values of

$$\frac{G_{\mathrm{eff}}}{G} = \begin{cases} 1 + k^2/[3(k^2 + a^2 m_{f_R}^2(a))] & f(R), \\ 1 + 1/[3\beta_{\mathrm{DGP}}(a)] & \mathrm{nDGP}, \end{cases} \tag{27}$$

where

$$m_{f_R}^2(a) = \frac{H_0^2(\Omega_{\mathrm{m}} + 4\Omega_\Lambda)}{2|f_{R0}|} \left(\frac{\Omega_{\mathrm{m}}a^{-3} + 4\Omega_\Lambda}{\Omega_{\mathrm{m}} + 4\Omega_\Lambda}\right)^3, \tag{28}$$

and

$$\beta_{\mathrm{DGP}}(a) = 1 + \frac{\Omega_{\mathrm{m}}a^{-3} + 2\Omega_\Lambda}{2\sqrt{\Omega_{\mathrm{rc}}(\Omega_{\mathrm{m}}a^{-3} + \Omega_\Lambda)}}, \tag{29}$$

where H_0 is the present-day value of the Hubble parameter, Ω_{m} and Ω_Λ are the current matter and dark energy density parameters, respectively. f_{R0} and Ω_{rc} are free parameters of each model that affect the deviation from the ΛCDM model. Note that $G_{\mathrm{eff}}^{f(R)}$ is a function of time and scale, which means that the linear growth of structure for $f(R)$ gravity is scale dependent, while for nDGP it is

scale independent. In Fig. 16, we show the theoretical values of the linear growth rate, equation (14), of these models for the cases: $f_{R0} = -10^{-5}$ and the range of scales $0.01 \leq k/[h \, \mathrm{Mpc}^{-1}] \leq 0.1$ for $f(R)$-gravity and $\Omega_{\mathrm{rc}} = 0.25$ for the nDGP model.

We see that the size of the errors from the DESI forecast is small enough to distinguish between the ΛCDM and the nDGP model. However, it is still unclear if we will be able to rule out $f(R)$ gravity models using RSDs.

5.4 Isotropic measurements of the baryon acoustic oscillations scale

Another direct application of our GLAM-HOD catalogues is the prediction of the BAO feature for DESI-like LRGs at different redshifts. This scale was not accessible in the PMILL run due to its volume. We extract the BAO scale through the dilation parameter, α, which is related to physical distances via (Eisenstein et al. 2005b)

$$\alpha \equiv \frac{D_{\mathrm{V}}(z)r_{\mathrm{d}}^{\mathrm{fid}}}{D_{\mathrm{V}}^{\mathrm{fid}}(z)r_{\mathrm{d}}}, \tag{30}$$

where

$$D_{\mathrm{V}}(z) = \left[cz(1+z)^2 D_{\mathrm{A}}^2(z)H^{-1}(z)\right]^{1/3}, \tag{31}$$

$D_{\mathrm{A}}(z)$ is the angular-diametre distance, r_{d} is the sound horizon at the baryon drag epoch ($z_{\mathrm{d}} \sim 1020$), and the superscript 'fid' indicates the value of the distances in our fiducial cosmology, i.e. the PMILL
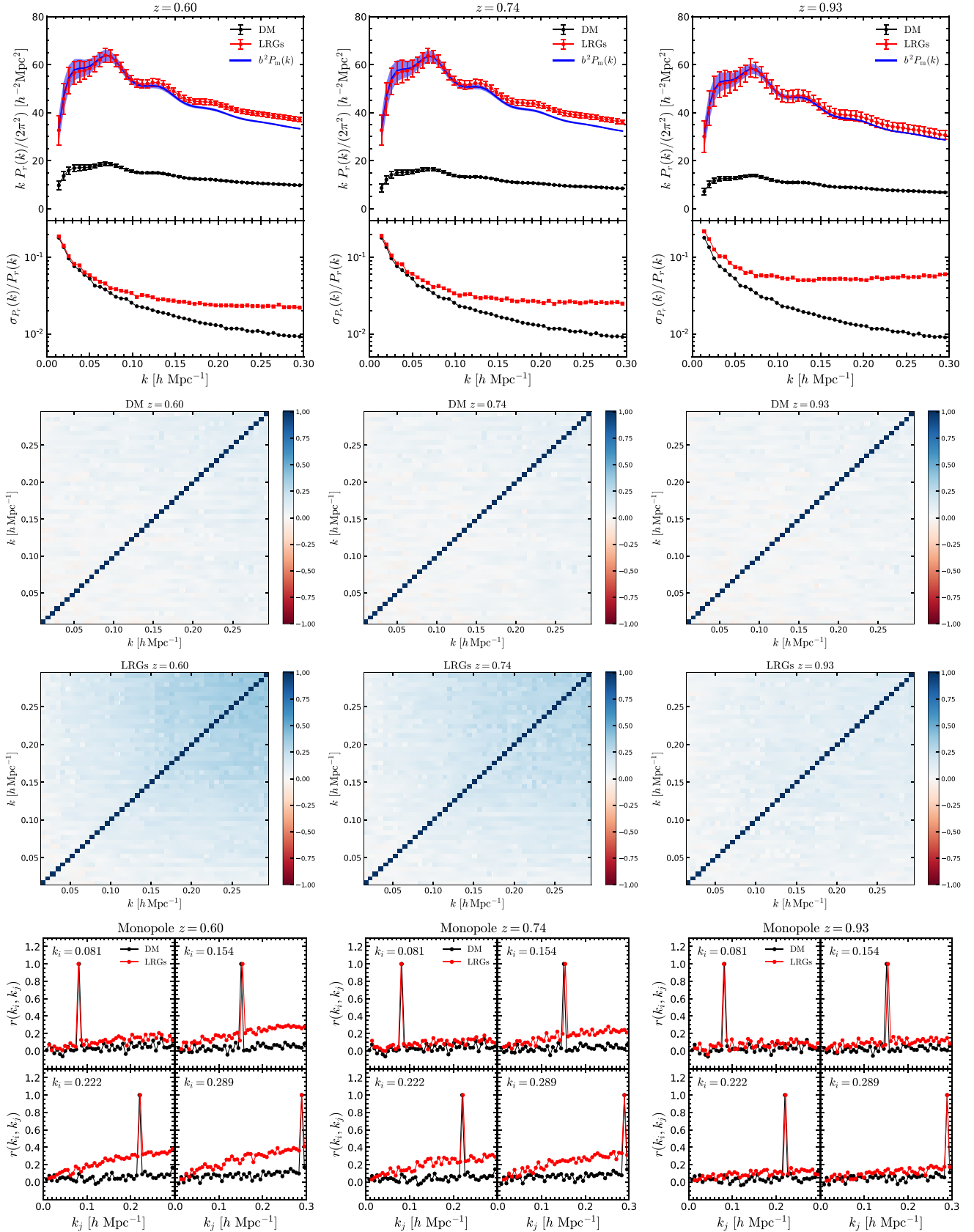
Figure 14. Covariance analysis of the real-space dark matter and LRG power spectra at $z = 0.6$, 0.74, and 0.93. Upper panels: Measured DM (black lines) and LRG (red lines) power spectrum from our GLAM simulations together with the DM power spectrum multiplied by the bias squared (blue lines; equation 6), the lower subpanels show the error contribution. Middle top and middle lower: Correlation matrices of the real-space power spectrum for the DM density field and LRGs, respectively. Bottom panels: Slices through the correlation matrices at different values of k_i in units of $[\,\mathrm{h\,Mpc^{-1}}]$.
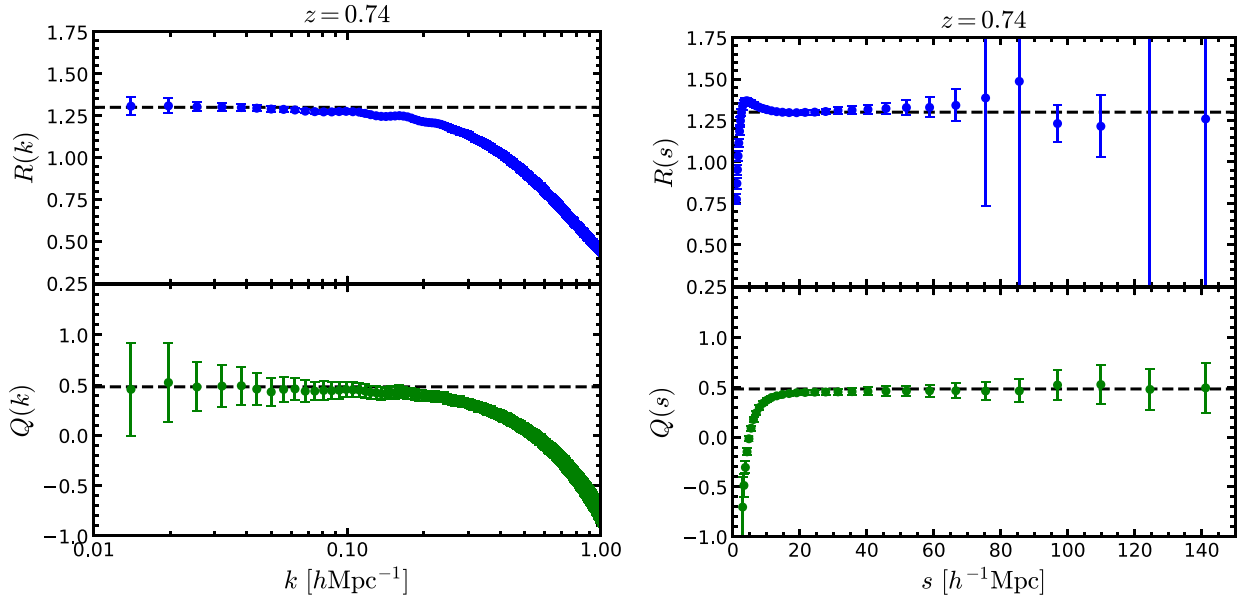
Figure 15. Estimators R (upper subpanels) (equation 23) and Q (lower subpanels), (equation 24) as function of separation in Fourier (left-hand panels) and configuration (right-hand panels) space at $z = 0.74$. Symbols with errorbars show the mean and standard deviation of the estimator measured from our 1000 GLAM catalogues. The black dashed line in each panel represents the fiducial linear theory value.
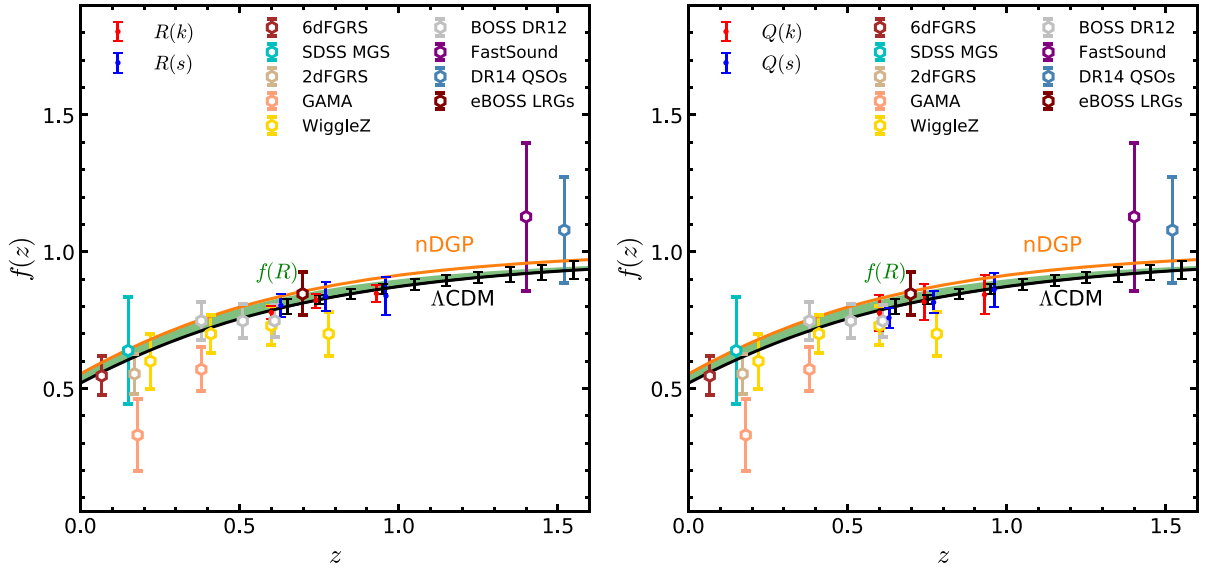

Figure 16. Evolution of the linear growth rate, f, as a function of redshift. Our estimations from R and Q are shown in the left- and right-hand panel in configuration (red dots with errorbars) and Fourier (blue dots with errorbars) space, respectively. The coloured symbols display measurements from different surveys at different redshifts as specified in the legend. Solid curves show the prediction for ΛCDM (black), nDGP (orange), and $f(R)$-gravity (green shaded region that represent wavenumbers $0.01 \leq k/[h\,\mathrm{Mpc}^{-1}] \leq 0.1$) models. The black errorbars over the ΛCDM prediction represent the DESI 14K forecast for $k_{\max} = 0.1\,h\,\mathrm{Mpc}^{-1}$ (see table 2.3 of DESI Collaboration et al. 2016).

Table 1. Results for the best-fitting values of the linear-growth rate, f, at redshifts 0.60, 0.74, and 0.93 for our estimators $R(k/s)$, equation (23), and $Q(k/s)$, equation (24).

Measurement	Redshift	Fiducial	$R(k)$	$Q(k)$	$R(s)$	$Q(s)$
f	$z = 0.60$	0.786	0.778 ± 0.024	0.776 ± 0.067	0.803 ± 0.041	0.758 ± 0.036
f	$z = 0.74$	0.823	0.822 ± 0.025	0.817 ± 0.066	0.839 ± 0.053	0.815 ± 0.039
f	$z = 0.93$	0.861	0.847 ± 0.033	0.842 ± 0.071	0.838 ± 0.070	0.861 ± 0.061

cosmology (see Section 2.1). In our fiducial cosmology, the values of $D_V(z)$ and r_d are

$$D_V^{\text{fid}}(z = 0.60) = 2141.07 \, \text{Mpc}, \tag{32}$$

$$D_V^{\text{fid}}(z = 0.74) = 2502.62 \, \text{Mpc}, \tag{33}$$

$$D_V^{\text{fid}}(z = 0.93) = 2926.11 \, \text{Mpc}, \tag{34}$$

$$r_d^{\text{fid}} = 148.13 \, \text{Mpc}. \tag{35}$$

The BAO scale can be extracted by fitting the monopole of the power spectrum (or correlation function) to a template that includes the dilation parameter. Therefore, the monopole of the power spectrum is modelled as the product of a smooth component and the BAO signal as (e.g. Anderson et al. 2014; Ross et al. 2015)

$$P_{0,\text{fit}}(k) = P_{\text{sm}}(k) O_{\text{damp}}(k/\alpha), \tag{36}$$

where $P_{\text{sm}}(k)$ is a smooth power spectrum, i.e. without any BAO feature, and $O_{\text{damp}}(k)$ represents the damped BAO signal (see below for the definitions of these quantities).

The smooth power spectrum component is modelled as (Anderson et al. 2014; Ross et al. 2015; Hernández-Aguayo et al. 2020)

$$P_{\text{sm}}(k) = B_p^2 P_{\text{nw}}(k) + A_1 k + A_2 + \frac{A_3}{k}, \tag{37}$$

where $P_{\text{nw}}(k)$ is a smooth 'de-wiggled' template obtained using the fitting formula of Eisenstein & Hu (1998), B_p is a large-scale bias parameter, and A_1, A_2, and A_3 are further free parameters.

The oscillatory component of the power spectrum is given by

$$O_{\text{damp}}(k) = 1 + \left(\frac{P_{\text{lin}}(k)}{P_{\text{nw}}(k)} - 1 \right) e^{-\frac{1}{2}k^2 \Sigma_{\text{nl}}^2}, \tag{38}$$

where Σ_{nl} is a damping parameter.

The monopole of the redshift-space correlation function is given by the model (Anderson et al. 2014; Ross et al. 2015)

$$\xi_{0,\text{fit}}(s) = B_s^2 \xi_{\text{lin, damp}}(\alpha s) + \frac{a_1}{s^2} + \frac{a_2}{s} + a_3, \tag{39}$$

where $\xi_{\text{lin, damp}}(s)$ is the Fourier transform of $P_{\text{nw}}(k) O_{\text{damp}}(k)$, B_s is the equivalent of B_p mentioned above, and a_1, a_2, and a_3 are polynomial free parameters.

To obtain the best-fitting α value, we use Bayesian statistics and maximize the likelihood, $\mathcal{L} \propto \exp(-\chi^2/2)$ (where χ^2 is defined by equation 13) by fitting the measurements of the monopole of the power spectrum on scales with $k < 0.3 \, h \, \text{Mpc}^{-1}$ and on scales with $s > 40 \, h^{-1} \text{Mpc}$ for the monopole of the correlation function. To find the best-fitting α value and its confidence levels, we again use the MCMC technique via the package EMCEE.

Fig. 17 displays the BAO feature in Fourier (left-hand panel) and configuration space (right-hand panel) at $z = 0.74$ (similar trends were found at $z = 0.60$ and 0.93). The BAO feature was isolated by dividing the best-fitting model and measurements of the monopole of the power spectrum by the smooth component of the best-fitting model. In the case of the monopole of the correlation function, we subtract the smooth component of the best-fitting model to the best-fitting model and measurements. We can see a clear BAO signal in both cases.

Our estimates of the dilation parameter are shown in Fig. 18 together with isotropic BAO measurements from the 6dFGRS at $z = 0.11$ (Beutler et al. 2011), the SDSS MGC at $z = 0.15$ (Ross et al. 2015), BOSS DR12 at $z = [0.38, 0.61]$ (Alam et al. 2017), WiggleZ at $z = [0.44, 0.6, 0.73]$ (Blake et al. 2011b), eBOSS DR16 LRGs at $z = 0.698$ (Bautista et al. 2021; Gil-Marín et al. 2020),

eBOSS DR14 LRGs at $z = 0.72$ (Bautista et al. 2018) and eBOSS DR14 QSO sample at $z = 1.52$ (Ata et al. 2018). The black errorbars are from the DESI forecast (see table 2.3 of DESI Collaboration et al. 2016).

Using the fiducial values of D_V^{fid}, equations (32)−(34), we convert our best-fitting α values into distance measurements via equation (30)

$$D_V(z = 0.60) = \begin{cases} 2140 \pm 28 \, (r_d/r_d^{\text{fid}}) \, \text{Mpc} & P_0(k), \\ 2145 \pm 27 \, (r_d/r_d^{\text{fid}}) \, \text{Mpc} & \xi_0(s), \end{cases} \tag{40}$$

$$D_V(z = 0.74) = \begin{cases} 2505 \pm 31 \, (r_d/r_d^{\text{fid}}) \, \text{Mpc} & P_0(k), \\ 2508 \pm 33 \, (r_d/r_d^{\text{fid}}) \, \text{Mpc} & \xi_0(s), \end{cases} \tag{41}$$

$$D_V(z = 0.93) = \begin{cases} 2927 \pm 38 \, (r_d/r_d^{\text{fid}}) \, \text{Mpc} & P_0(k), \\ 2926 \pm 40 \, (r_d/r_d^{\text{fid}}) \, \text{Mpc} & \xi_0(s). \end{cases} \tag{42}$$

We find good agreement between our estimates and the fiducial values of $D_V(z)$. The agreement is well within the 1σ level. In our case, the monopole of the correlation function gives slightly better constraints than the power spectrum. In general, we can estimate the isotropic BAO distance to better than 1.3 per cent in both spaces. Note that we did not scale our errorbars by the DESI volume.

6 SUMMARY AND CONCLUSIONS

We have presented predictions for the properties and clustering of LRGs selected using colour–magnitude cuts in the r, z, $W1$ bands that are similar to those that will be applied in the DESI LRG survey (DESI Collaboration et al. 2016). The predictions were made using the GALFORM semi-analytic model of galaxy formation run on the PMILL N-body simulation (Baugh et al. 2019) and a suite of low-resolution, larger volume simulations run with the Parallel-PM N-body code GLAM (Klypin & Prada 2018).

Our aim is to carry out a proof of concept study of our approach, using a non-trivial galaxy selection, to illustrate the power of using a physically motivated galaxy formation model. We made predictions for the abundance of 'DESI-like' LRGs and explore how the target selection cuts affect which galaxies are selected and how these populate haloes and subhaloes. We find that a small but important fraction of the most massive galaxies (those with stellar mass $\log_{10}(M_*/h^{-1} \text{M}_\odot) > 11.15$) are not selected as LRGs (see Fig. 4). A similar trend is seen in the galaxy luminosity function, and is most pronounced at shorter wavelengths: essentially all bright galaxies in the $W1$-band luminosity function are LRGs, but only roughly half of the galaxies in the bright end of the r-band luminosity function are LRGs. This shows that applying the full photometric selection is essential to reproduce LRGs in a galaxy formation model and that using a proxy, such as stellar mass, to select LRGs is at best an approximation. We explored the galaxy–(sub)halo connection of LRGs through the HOD and the subhalo mass function. We find that the shape of the HOD does not follow the canonical shape proposed by Zheng et al. (2005); in particular, the occupation of central galaxies does not reach unity for the most massive haloes (see Fig. 5), and drops with increasing mass.

We compared the HOD and the subhalo mass functions of galaxies selected by their stellar mass with those measured for the LRGs (see Fig. 6). By doing this exercise, we reaffirm that the DESI-LRG cuts affect the selection of subhaloes that are populated by LRGs. Mass alone is not enough to determine if a subhalo hosts an LRG. By comparing the clustering of these galaxy samples (Fig. 7), we found a difference that ranges from 10 per cent at $z = 0.6-0.74$ to up to 150 per cent at $z = 0.93$. Hence, we conclude that using galaxy stellar mass as a proxy for selecting LRGs could change the expected clustering signal.
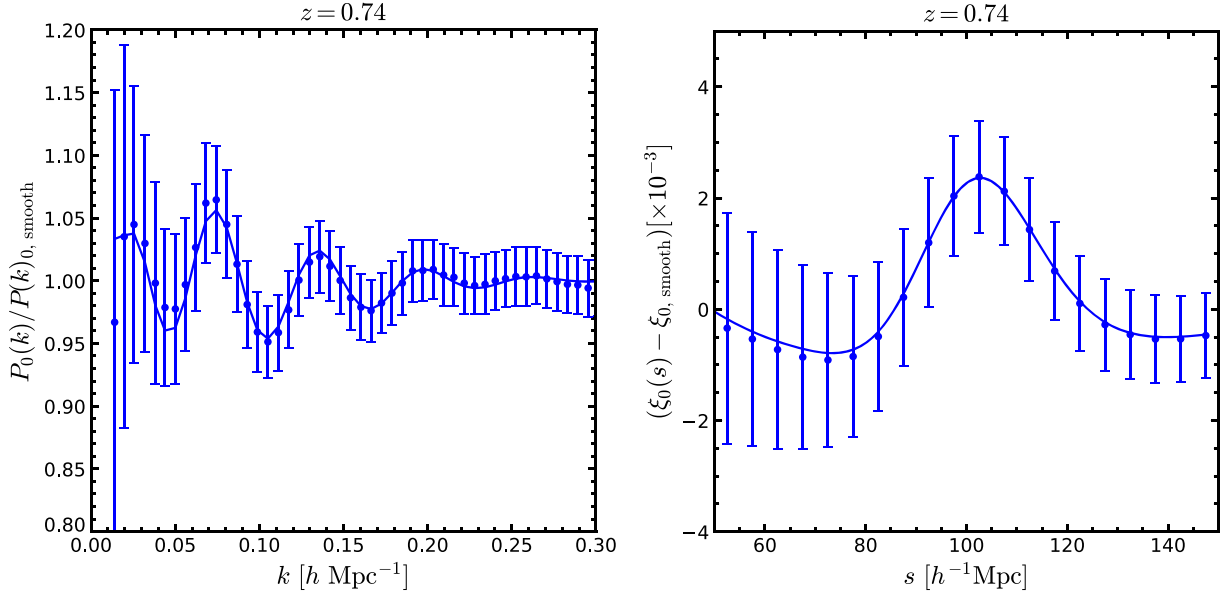
Figure 17. BAO signals in the monopole of the power spectrum (left-hand panel) and correlation function (right-hand panel) at $z = 0.74$. Blue dots with error bars come from the measurements from our GLAM-LRG catalogues. The solid lines show the predictions from the best-fitting BAO models. In order to highlight the BAO features, we have divided the $P(k)$ measurements and the best-fitting model by the no-wiggle power spectrum of the best-fitting model. In configuration space, we have also subtracted the smooth component of the best-fitting model.
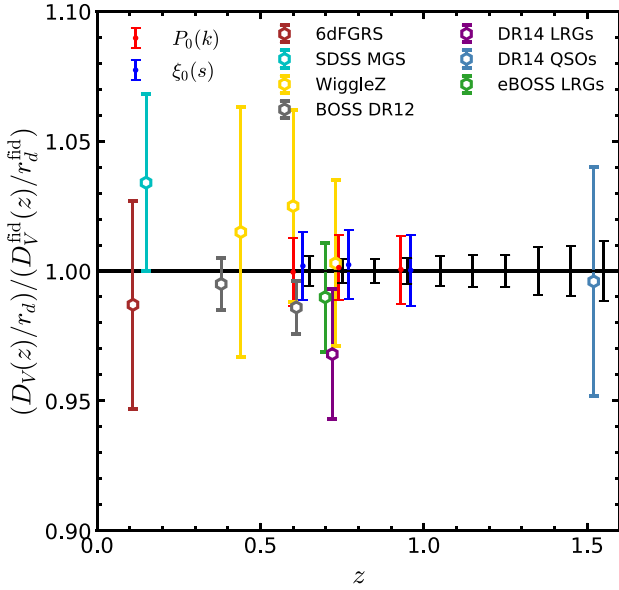
Figure 18. Isotropic BAO measurements as a function of redshift. Our estimates from the monopole of the power spectrum are shown by the red dots with errorbars while the blue dots with errobars represent the best-fitting values from the monopole of the correlation function. Note that we did not rescale our errorbars by the DESI survey volume.
We show the measurements from different galaxy surveys as labelled. The black errorbars show the DESI 14K forecast presented in DESI Collaboration et al. (2016).

To prepare for the clustering measurements of DESI, we ran 1000 GLAM simulations. When comparing the halo statistics between the GLAM simulation ensemble and the PMILL high-resolution run, we found good agreement between the halo mass functions, but

differences of ~ 10 per cent in the halo clustering (see Fig. 1). This difference can be attributed to the different halo finder used in the PMILL and GLAM simulations. Despite the difference in halo clustering, the galaxy clustering statistics measured from the GLAM-LRG catalogues are in good agreement with that in the PMILL-GALFORM LRG sample in both configuration and Fourier space. To populate the GLAM halo catalogues with DESI-like LRGs, we used the tabulated HODs obtained from GALFORM. We also found a good agreement between our clustering measurements in real-space with those reported by Kitanidis et al. (2019) and Zhou et al. (2020a) (see upper panels of Figs 8 and 9).

We extended the analysis of covariance and correlation matrices of GLAM simulations started by Klypin & Prada (2018) to galaxies and correlation functions (see Figs 10–14). We found that the galaxy correlation matrix shows a different and more complex pattern than its dark-matter counterpart.

We presented predictions for the large-scale clustering of DESI-like LRGs in configuration and Fourier space, by extracting the linear growth rate from the linear Kaiser RSD model and the BAO scale from the isotropic dilation parameter. In a follow-up project, we plan to extend this study to non-linear models of RSDs and an anisotropic analysis of the BAO scale, including the impact of the light-cone survey geometry and observational systematic.

Using our GLAM-LRG catalogues, we estimated the growth of structure from the ratio of the monopole in redshift space to the real-space power spectrum with a precision of $\sim 3 - 4$ per cent, and we can measure the BAO scale with a 1.3 per cent precision in both configuration and Fourier space. Nevertheless, if we want to compare the precision of our measurements with those expected from DESI (table 2.3 of DESI Collaboration et al. 2016), our error estimations should take into account the contribution from the expected volume covered by DESI (see Section 5.2 for details). However, the amplitude of the statistical errors estimated from our best-fitting search on the linear growth rate and BAO scale are consistent with the forecast presented by DESI Collaboration et al. (2016).

We conclude that the colour–magnitude cuts have a big impact on the properties and clustering of LRGs, showing that LRGs are different than stellar mass selected galaxies. But more importantly, the analysis presented in this paper provided accurate estimates on the galaxy clustering expected by DESI-LRGs thanks to our GLAM-HOD pipeline. The GLAM-LRG galaxy catalogues are made public at the Skies & Universes site.[3] Moreover, our pipeline can be easily adapted to the specifications of other next generation surveys such as Euclid, the Vera Rubin Observatory (formerly the LSST), PFS, and 4MOST.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The GLAM DESI-like and halo catalogues are available at http://www.skiesanduniverses.org/Products/MockCatalogues/GLAMDESILRG/

REFERENCES

Ade P. A. R. et al., 2014, A&A, 571, A16
Alam S. et al., 2015, ApJ Suppl., 219, 12
Alam S. et al., 2017, MNRAS, 470, 2617
Almeida C., Baugh C. M., Lacey C. G., 2007, MNRAS, 376, 1711
Almeida C., Baugh C., Wake D., Lacey C., Benson A., Bower R., Pimbblet K. A., 2008, MNRAS, 386, 2145
Anderson L. et al., 2014, MNRAS, 441, 24
Ata M. et al., 2018, MNRAS, 473, 4773
Aubourg É. et al., 2015, Phys. Rev. D, 92, 123516
Barreira A., Sánchez A. G., Schmidt F., 2016, Phys. Rev. D, 94, 084022
Baugh C. M., 2006, Rept. Prog. Phys., 69, 3101
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Baugh C. M. et al., 2019, MNRAS, 483, 4922
Baumgarten F., Chuang C.-H., 2018, MNRAS, 480, 2535
Bautista J. E. et al., 2018, ApJ, 863, 110
Bautista J. E. et al., 2021, MNRAS, 500, 736

Benson A. J., 2012, New Astron., 17, 175
Bernardi M. et al., 2003, Astron. J., 125, 1866
Bernardi M., Sheth R. K., Nichol R. C., Schneider D., Brinkmann J., 2005, Astron. J., 129, 61
Beutler F. et al., 2011, MNRAS, 416, 3017
Beutler F. et al., 2012, MNRAS, 423, 3430
Blake C., Collister A., Lahav O., 2008, MNRAS, 385, 1257
Blake C. et al., 2011a, MNRAS, 415, 2876
Blake C. et al., 2011b, MNRAS, 418, 1707
Blake C. et al., 2013, MNRAS, 436, 3089
Blot L. et al., 2019, MNRAS, 485, 2806
Bower R. G., Benson A., Malbon R., Helly J., Frenk C., Baugh C., Cole S., Lacey C., 2006, MNRAS, 370, 645
Brown M. J. I. et al., 2008, ApJ, 682, 937
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Cabre A., Gaztanaga E., 2009a, MNRAS, 393, 1183
Cabre A., Gaztanaga E., 2009b, MNRAS, 396, 1119
Campbell D. J. R. et al., 2015, MNRAS, 452, 852
Colavincenzo M. et al., 2019, MNRAS, 482, 4883
Cole S., Fisher K. B., Weinberg D. H., 1994, MNRAS, 267, 785
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Contreras S., Baugh C., Norberg P., Padilla N., 2013, MNRAS, 432, 2717
Contreras S., Baugh C. M., Norberg P., Padilla N., 2015, MNRAS, 452, 1861
Cowley W. I., Lacey C. G., Baugh C. M., Cole S., 2016, MNRAS, 461, 1621
Crocce M., Gaztañaga E., Cabré A., Carnero A., Sánchez E., 2011, MNRAS, 417, 2577
Dawson K. S. et al., 2013, AJ, 145, 10
Dawson K. S. et al., 2016, AJ, 151, 44
DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
Dey A. et al., 2019, AJ, 157, 168
Dvali G. R., Gabadadze G., Porrati M., 2000, Phys. Lett., B485, 208
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eisenstein D. J., Blanton M., Zehavi I., Bahcall N., Brinkmann J., Loveday J., Meiksin A., Schneider D., 2005a, ApJ, 619, 178
Eisenstein D. J. et al., 2005b, ApJ, 633, 560
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gil-Marín H. et al., 2020, MNRAS, 498, 2492
Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, MNRAS, 439, 264
Gonzalez-Perez V. et al., 2018, MNRAS, 474, 4024
Gonzalez-Perez V. et al., 2020, MNRAS, 498, 1852
Hamilton A. J. S., 1992, ApJ, 385, L5
Hamilton A. J. S., 1998, in Hamilton D., ed., Astrophysics and Space Science Library Vol. 231, The Evolving Universe. Kluwer, Dordrecht, p. 185
Hand N., Feng Y., Beutler F., Li Y., Modi C., Seljak U., Slepian Z., 2018, AJ, 156, 160
Hawkins E. et al., 2003, MNRAS, 346, 78
Hernández-Aguayo C., Hou J., Li B., Baugh C. M., Sánchez A. G., 2019, MNRAS, 485, 2194
Hernández-Aguayo C., Cautun M., Smith A., Baugh C. M., Li B., 2020, MNRAS, 494, 3120
Howlett C., Ross A., Samushia L., Percival W., Manera M., 2015, MNRAS, 449, 848
Hu W., Kravtsov A. V., 2003, ApJ, 584, 702
Hu W., Sawicki I., 2007, Phys. Rev., D76, 064004
Icaza-Lizaola M. et al., 2020, MNRAS, 492, 4189
Jenkins A., 2010, MNRAS, 403, 1859
Jenkins A., 2013, MNRAS, 434, 2094
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Jiménez E., Contreras S., Padilla N., Zehavi I., Baugh C. M., Gonzalez-Perez V., 2019, MNRAS, 490, 3532
Kaiser N., 1987, MNRAS, 227, 1
Kitanidis E. et al., 2020 , MNRAS, 496, 2262
Klypin A., Holtzman J., 1997, preprint (arXiv:astro-ph/9712217)
Klypin A., Prada F., 2018, MNRAS, 478, 4602

[3] http://www.skiesanduniverses.org

Klypin A., Yepes G., Gottlober S., Prada F., Hess S., 2016, MNRAS, 457, 4340

Lacey C. G. et al., 2016, MNRAS, 462, 3854

Lippich M. et al., 2019, MNRAS, 482, 1786

Manera M. et al., 2012, MNRAS, 428, 1036

Merson A., Wang Y., Benson A., Faisst A., Masters D., Kiessling A., Rhodes J., 2018, MNRAS, 474, 177

Merson A., Smith A., Benson A., Wang Y., Baugh C. M., 2019, MNRAS, 486, 5737

Mitchell P. D., Lacey C. G., Baugh C. M., Cole S., 2016, MNRAS, 456, 1459

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Okumura T. et al., 2016, Publ. Astron. Soc. Jap., 68, 38

Padmanabhan N., White M., Norberg P., Porciani C., 2009, MNRAS, 397, 1862

Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton University Press, USA

Percival W. J. et al., 2004, MNRAS, 353, 1201

Prakash A. et al., 2016, ApJS, 224, 34

Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, MNRAS, 449, 835

Samushia L., Percival W. J., Raccanelli A., 2012, MNRAS, 420, 2102

Sanchez A. G., Crocce M., Cabre A., Baugh C. M., Gaztanaga E., 2009, MNRAS, 400, 1643

Simha V., Cole S., 2017, MNRAS, 472, 1392

Smith A., Cole S., Baugh C., Zheng Z., Angulo R., Norberg P., Zehavi I., 2017, MNRAS, 470, 4646

Springel V., 2005, MNRAS, 364, 1105

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Stoppacher D. et al., 2019, MNRAS, 486, 1316

Tegmark M. et al., 2006, Phys. Rev., D74, 123507

Tröster T. et al., 2020, Astron. Astrophys., 633, L10

Wake D. et al., 2006, MNRAS, 372, 537

Wake D. A. et al., 2008, MNRAS, 387, 1045

Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435

Zarrouk P. et al., 2018, MNRAS, 477, 1639

Zehavi I. et al., 2005, ApJ, 621, 22

Zhai Z. et al., 2017, ApJ, 848, 76

Zheng Z. et al., 2005, ApJ, 633, 791

Zheng Z., Zehavi I., Eisenstein D. J., Weinberg D. H., Jing Y., 2009, ApJ, 707, 554

Zhou R. et al., 2020a, MNRAS, 501, 3309

Zhou R. et al., 2020b, Res. Notes Am. Astron. Soc., 4, 181

This paper has been typeset from a TEX/LATEX file prepared by the author.