



# Clustering in the simulated H $\alpha$ galaxy redshift survey from *Nancy Grace Roman Space Telescope*

Zhongxu Zhai,<sup>1</sup>★ Chia-Hsun Chuang<sup>1,2</sup>, Yun Wang,<sup>1</sup> Andrew Benson<sup>1,3</sup> and Gustavo Yepes<sup>4,5</sup>

<sup>1</sup>IPAC, California Institute of Technology, Mail Code 314-6, 1200 E. California Blvd., Pasadena, CA 91125, USA

<sup>2</sup>Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA

<sup>3</sup>Carnegie Observatories, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>4</sup>Departamento de Física Teórica, Módulo 8, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

<sup>5</sup>CIAFF, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

Accepted 2020 December 15. Received 2020 November 13; in original form 2020 July 15

## ABSTRACT

We present a realistic 2000 deg<sup>2</sup> H $\alpha$  galaxy mock catalogue with  $1 < z < 2$  for the *Nancy Grace Roman Space Telescope (Roman)* galaxy redshift survey, the High Latitude Spectroscopic Survey (HLSS), created using Galacticus, a semi-analytical galaxy formation model, and high-resolution cosmological N-body simulations. Galaxy clustering can probe dark energy and test gravity via baryon acoustic oscillation (BAO) and redshift space distortion (RSD) measurements. Using our realistic mock as the simulated *Roman* HLSS data, and a covariance matrix computed using a large set of approximate mocks created using EZMOCK, we investigate the expected precision and accuracy of the BAO and RSD measurements using the same analysis techniques used in analysing real data. We find that the *Roman* H $\alpha$  galaxy survey alone can measure the angular diameter distance with 2 per cent uncertainty, the Hubble parameter with 3–6 per cent uncertainty, and the linear growth parameter with 7 per cent uncertainty, in each of four redshift bins. Our realistic forecast illustrates the power of the *Roman* galaxy survey in probing the nature of dark energy and testing gravity.

**Key words:** methods: numerical – methods: statistical – galaxies: formation – cosmology: large-scale structure of Universe.

## 1 INTRODUCTION

Clustering of galaxies in the universe provides an important probe to constrain the relationship between galaxies and the underlying matter distribution. The accurate measurement can be used to describe the physics governing galaxy formation and evolution. In addition, the signal measured at different scales is also able to constrain the evolution of the large-scale structure, including the linear growth rate through the redshift space distortion (RSD) effect and cosmic distance scales through the baryon acoustic oscillation (BAO).

In the early universe, the photons and baryons are tightly correlated due to the high temperature. The pressure from the photons and gravitational interaction from the matter perturbations lead to oscillations in the coupled fluid. This feature is imprinted on the fluctuation of temperature and can be observed through cosmic microwave background (CMB), for instance the *WMAP* satellite (Bennett et al. 2003; Spergel et al. 2003) and *Planck* mission (Planck Collaboration I 2014a; Planck Collaboration XVI 2014b). This oscillating feature is also imprinted on the matter distribution, which can be measured through the power spectrum and correlation function of galaxies. Since the first detection of this BAO signal in the Sloan Digital Sky Survey (SDSS; York et al. 2000) data set (Eisenstein et al. 2005) and the two-degree Field Survey (Cole et al.

2005), the technique of measuring and analysing this phenomenon has evolved significantly in the past decade. Based on observations from spectroscopic surveys like 6-degree Field Galaxy Survey (6dFGS; Beutler et al. 2011), SDSS-II (Ross et al. 2015), SDSS-III (Dawson et al. 2013; Anderson et al. 2014; Font-Ribera et al. 2014; Delubac et al. 2015; Alam et al. 2017), SDSS-IV (Dawson et al. 2016; Ata et al. 2018; Bautista et al. 2018), and WiggleZ (Blake et al. 2011), the BAO signal has been reported in different types of tracers, including the luminous red galaxies, emission-line galaxies (ELGs), and quasars. The redshift range is also not limited to local galaxies. The measurement from Lyman- $\alpha$  forest is able to measure the BAO signal in auto- and cross-correlation at redshift up to 2.40 (Font-Ribera et al. 2014; Delubac et al. 2015). The compilation of these state-of-the-art measurements of the BAO signal and the cosmic distance scales can place constraint on the fundamental cosmological parameters. Combination with other cosmological data such as CMB, Type Ia supernovae (SNe), and so on can further constrain the cosmological model and its deviation from lambda cold dark matter ( $\Lambda$ CDM). This has been widely explored in the literature; see e.g. Wang & Mukherjee (2007), Planck Collaboration XIII (2016), Zhai et al. (2017), Zhai & Wang (2019), Planck Collaboration VI (2020), eBOSS Collaboration (2020), and references therein.

Due to peculiar velocities, the observed redshifts of the galaxies are different from the prediction for a homogeneous universe. This so-called RSD effect changes the observed galaxy density field and the

\* E-mail: [zhai@ipac.caltech.edu](mailto:zhai@ipac.caltech.edu)

inferred clustering signal (Kaiser 1987). However, this systematics also carries information of the universe and can be used to infer the linear growth rate of cosmic large-scale structure. With the galaxy catalogue from 6dFGS (Beutler et al. 2012), WiggleZ (Blake et al. 2011), SDSS-II (Chuang & Wang 2013), SDSS-III (Beutler et al. 2014), etc., we can have measurements over a wide redshift range that provide complementary constraints on the cosmological model. This dynamic probe is different than the geometrical probe such as BAO and Type Ia SNe, and can distinguish dark energy and modified gravity models from the standard  $\Lambda$ CDM model when they have the same expansion history.

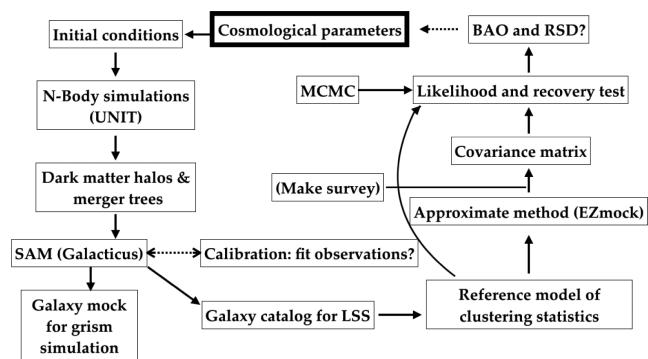
As NASA's flagship survey for next generation, *Nancy Grace Roman Space Telescope* (*Roman*; Dressler et al. 2012; Green et al. 2012; Spergel et al. 2015) will measure millions of ELGs in its planned 5 yr mission. The resulting data will provide uniform and powerful measurements of the BAO signal and the RSD effect, as well as information to study galaxy evolution and star formation histories. For future surveys like *Roman* and *Euclid* (Laureijs et al. 2011, 2012), we need to optimize the survey designs and evaluate their performance in constraining dark energy. In Merson et al. (2018) and Zhai et al. (2019a), we estimate the number densities of the H $\alpha$  and [O III] ELGs that *Roman* and *Euclid* will be able to observe based on semi-analytic model (SAM) of galaxy formation. This enables further investigation of the clustering properties of these galaxies (Merson et al. 2019). In this work, we move forward by making a realistic galaxy mock catalogue of the data expected from the *Roman* galaxy redshift survey (GRS), the High Latitude Spectroscopic Survey (HLSS), and investigate the expected precision and accuracy of the BAO and RSD measurements from it using the same analysis techniques used in analysing real data. The mock catalogues presented here can also be used to study other cosmic probes, such as clustering measurement at small scales, higher order statistics, cross-correlation with different tracers of dark matter distribution, and so on.

Note that the Fisher matrix-based method is often used to forecast the performance of future surveys, but it underestimates the uncertainties by construction (Tegmark 1997). In this paper, we apply the same methodology used in analysing actual data to the simulated *Roman* galaxy survey data. We utilize the calibrated SAM to generate a synthetic *Roman* H $\alpha$  galaxy catalogue. This catalogue has an area of  $2000 \text{ deg}^2$  to be consistent with the current baseline mission design. The observed cosmic volume can contain millions of galaxies for clustering measurements. In order to estimate the uncertainty of the clustering measurements, we compute the covariance matrix by adopting a quick and economic method to generate a large number of approximate mocks (EZmocks) that have clustering consistent with the SAM-derived mock (the simulated *Roman* data set).

Our paper is organized as follows: In Section 2, we introduce the method for producing the galaxy mocks from SAM and EZmocks. Section 3 describes the methodology for modelling the galaxy power spectrum, and Section 4 presents the results on the BAO/RSD measurements from the simulated *Roman* H $\alpha$  galaxy mock. We discuss and conclude in Section 5.

## 2 METHODOLOGY FOR PRODUCING GALAXY MOCKS

The analysis in this paper involves calibration and generation of multiple mocks for galaxy clustering, as well as the theoretical modelling to validate and interpret the result. Fig. 1 summarizes



**Figure 1.** Flowchart of the algorithm applied in this work, starting from a cosmological model. We assume a cosmology as applied in the UNIT simulation (Chuang et al. 2019) throughout the work, i.e. the Planck 2016 cosmology (Planck Collaboration XIII 2016). With the dark matter halo and merger trees constructed from the  $N$ -body simulation, we calibrate the SAM to produce a galaxy catalogue for the *Roman* galaxy survey. This catalogue is assumed as the reference model to calibrate EZMOCK that is used to generate a large number of approximate mocks for clustering analysis. We then estimate the covariance matrix for the clustering statistics of interest. This enables a recovery test to validate our simulation and modelling, as well as evaluate the precision and accuracy of BAO and RSD measurements from the *Roman* galaxy survey. Each step is explained in detail in the text.

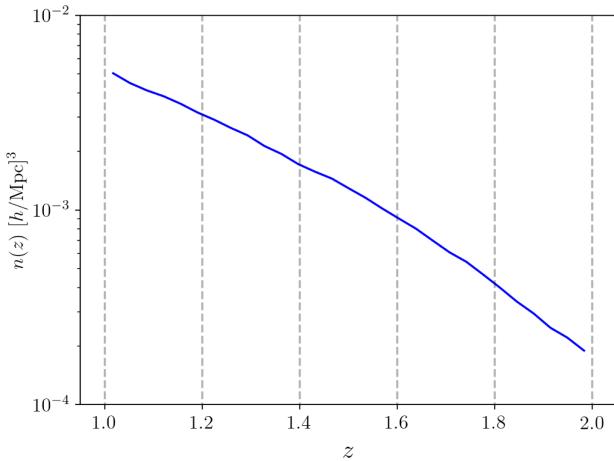
the whole algorithm in the analysis and each step is explained and described in following sections.

### 2.1 Reference catalogue from SAM

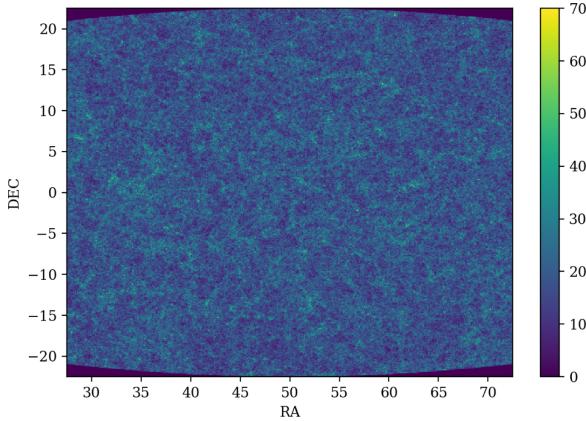
The synthetic galaxy catalogue can be constructed by populating high-resolution  $N$ -body simulations with galaxies. We adopt the SAM model based on Zhai et al. (2019a), which is used to predict the number counts of ELGs that *Roman* will observe. In particular, we apply Galacticus (Benson 2012) to model galaxy formation and evolution. In Zhai et al. (2019a), this model is calibrated based on observations of the H $\alpha$  luminosity function from the ground-based narrow-band High-z Emission Line Survey (HiZELS; Geach et al. 2008; Sobral et al. 2009, 2013). In this work, we adopt the same parameters of Galacticus and the dust model to produce a reference catalogue for clustering analysis. Note that Zhai et al. (2019a) compare the prediction from two dust models, which can match the data from HiZELS or WISP number counts (Mehta et al. 2015), respectively. For simplicity in this work, we just consider the dust model that can produce consistent result with HiZELS. The other dust model allows observation of more faint galaxies due to its weaker dust attenuation. We do not expect it to significantly change the clustering analysis in this work. More details of the SAM can be found from Zhai et al. (2019a), and the dark matter halo merger trees used as input for the SAM calculation are from the UNIT simulation<sup>1</sup> (Chuang et al. 2019). We refer the readers to these references for more information and details. In addition, we also use the same SAM to produce a small simulated galaxy catalogue along with spectral energy distributions as the input for the pixel level grism simulations for the *Roman* galaxy survey, to be presented in a separate paper.

The *Roman* HLSS H $\alpha$  mock generated from Galacticus covers an area of  $\sim 2000 \text{ deg}^2$ , consistent with the current baseline mission design. We select all the galaxies with dust-attenuated H $\alpha$  flux higher

<sup>1</sup><https://unitsims.ft.uam.es>



**Figure 2.** The number density as a function of redshift of the mock galaxies generated by Galacticus. The vertical dashed grey lines indicate the redshift slices in the clustering analysis.



**Figure 3.** The distribution of the mock galaxies generated by Galacticus for subsample at  $1.0 < z < 1.2$ , colour coded by density on coordinates RA and Dec. Other redshift bins are omitted to avoid redundancy.

than  $1 \times 10^{-16}$  erg s $^{-1}$  cm $^{-2}$ . This flux limit corresponds to the  $6.5\sigma$  nominal depth for the *Roman* GRS. The number density distribution of these galaxies is shown in Fig. 2. Note that the survey sensitivity has a direct impact on the number densities of large galaxies and thus influences the uncertainty of the measured clustering signal. In the redshift range  $1.0 < z < 2.0$ , the survey can observe 10 million H  $\alpha$ -emitting galaxies within a cosmic volume of  $7 (\text{Gpc}/h)^3$ . Fig. 3 displays the density distribution of mock galaxies at  $1.0 < z < 1.2$  on the sky, generated by Galacticus simulation. The other redshift bins have the same RA/Dec. distribution. For clustering analysis, we split the catalogue into five redshift bins with width  $\Delta z = 0.2$ . We first measure the two-point correlation function  $\xi_r$  in configuration space by the LS estimator (Landy & Szalay 1993)

$$\xi_r = \frac{\text{DD} - 2\text{DR} + \text{RR}}{\text{RR}}, \quad (1)$$

where DD, DR, and RR are suitably normalized numbers of data–data, data–random, and random–random pairs with separation  $r$ . The random catalogue is generated to match the sky coverage as the data catalogue, i.e. have the same distribution for right ascension and declination and uniformly distributed on the sky. The redshifts of the random catalogues are random draws from the data such that both

catalogues have consistent radial distributions. In order to obtain a stable measurement of clustering, the random catalogue is 10 times larger than the data.

Due to peculiar velocities, the spectroscopically observed redshift of a galaxy is the combination of two components:

$$z_{\text{spec}} = z_{\text{cosmo}} + \frac{v_p}{ac}, \quad (2)$$

where  $z_{\text{cosmo}}$  is the redshift if the universe is homogeneous,  $v_p$  is the peculiar velocity along the line of sight,  $a$  is the cosmic scale factor with  $a = 1/(1 + z_{\text{cosmo}})$ , and  $c$  is the speed of light. This second term leads to a distortion in the measured galaxy density field and thus the resultant clustering property, i.e. the RSD effect (Kaiser 1987). We add this effect to the galaxy catalogue and remeasure the clustering. On large scales where the linear perturbation theory applies, the RSD effect increases the clustering amplitude compared to the expectation in real space and affects the significance of the BAO peak.

The galaxy power spectrum is the two-point correlation function in Fourier space. For a given galaxy overdensity in Fourier space  $\delta_g(\mathbf{k})$ , the power spectrum is defined as

$$\langle \delta_g(\mathbf{k})\delta_g(\mathbf{k}') \rangle = (2\pi)^3 P(k)\delta^D(\mathbf{k} + \mathbf{k}'). \quad (3)$$

For both correlation function and power spectrum, we can expand them in harmonics and represent the measurements as Legendre multipoles. For the galaxy correlation function, we can measure it on a two-dimensional grid of galaxy pair separations perpendicular ( $r_p$ ) and parallel ( $\pi$ ) to the line of sight through

$$\pi = \frac{\mathbf{s} \cdot \mathbf{l}}{|\mathbf{l}|}, \quad r_p = \mathbf{s} \cdot \mathbf{s} - \pi^2, \quad (4)$$

where  $\mathbf{l} = (\mathbf{s}_1 + \mathbf{s}_2)/2$  and  $\mathbf{s} = \mathbf{s}_1 - \mathbf{s}_2$ , where  $\mathbf{s}_1$  and  $\mathbf{s}_2$  are the positions of a pair of galaxies (Davis & Peebles 1983; Fisher et al. 1994). The resultant Legendre multipoles are given as

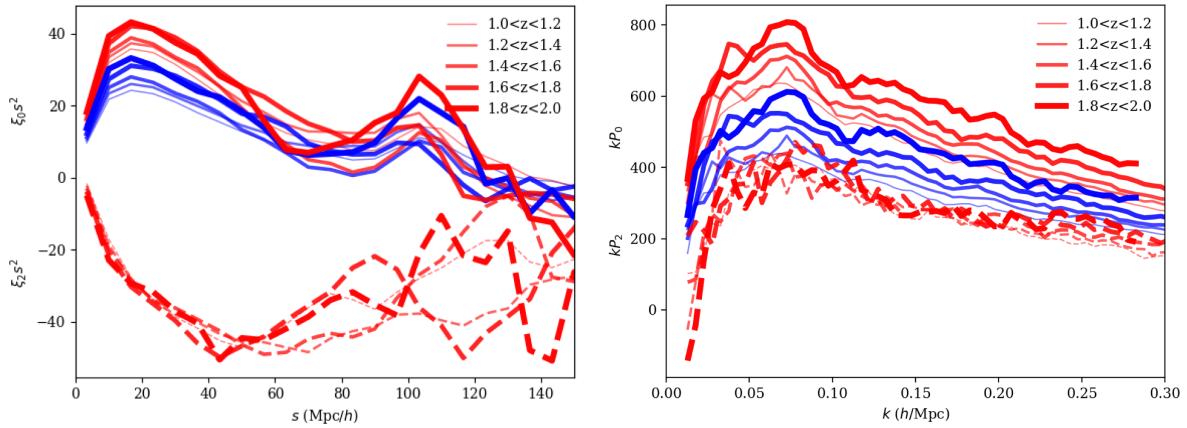
$$\xi_\ell(r) = \frac{2\ell + 1}{2} \int_{-1}^1 \xi(s, \mu) L_\ell(\mu) d\mu, \quad (5)$$

where  $L_\ell$  is the Legendre polynomial of order  $\ell$ , and  $\mu = \pi/s$ . The decomposition of the galaxy power spectrum can be done in a similar way.

The clustering measurements of the galaxies in our simulated *Roman* HLSS H  $\alpha$  mock are shown in Fig. 4, for both correlation function and power spectrum in real and redshift space. We have used the publicly available code CUTE (Alonso 2012) to calculate the two-point correlation function, and NBODYKIT (Hand et al. 2018) to calculate the power spectrum. This mock catalogue and its clustering measurement form the reference model in the construction of the large set of mock catalogues required for computing the covariance matrix.

## 2.2 Approximate mocks from EZMOCK

In order to assess the uncertainty of the measurements in a reliable manner, we need to estimate the covariance matrix of the measurements. The most robust approach is to employ a large number of mocks with clustering properties consistent with the data set. In principle, one can use the same SAM approach to generate the galaxy mocks. However, this requires a large number of  $N$ -body simulations to provide dark matter merger trees, which would be prohibitive in both cost and time using the current computational technology. Therefore, we need an economic but accurate way to generate a sufficient number of galaxy mocks to estimate the covariance of the observable. This is the approach currently used in the analysis of actual data.



**Figure 4.** Clustering measurements from the reference mock of *Roman* HLSS H $\alpha$  ELGs generated by Galacticus. Left: two-point correlation function; right: power spectrum. Solid lines show the monopole of the Legendre decomposition, while the dashed lines are the quadrupole. Blue is the result for real space, while red is in redshift space. Note that the Legendre multipoles higher than monopole in real space are zero for both correlation function and power spectrum. Different line thickness denotes the redshift bins as shown in the legend.

In this work, we use the EZMOCK code (Chuang et al. 2015) to generate the approximate mocks. EZmocks are constructed using the Zel'dovich approximation of the density field and effectively include stochastic scale-dependent, non-local and non-linear biasing contributions. This approach has been demonstrated to be able to produce accurate clustering properties comparable to  $N$ -body simulations, including one-point, two-point, and three-point statistics. This method has been widely used in the analysis of large-scale structure, for instance in the BAO measurement from eBOSS quasar observations (Ata et al. 2018) and the clustering analysis for SDSS DR16 (Bautista et al. 2021; de Mattia et al. 2020; Raichoor et al. 2021; Tamone et al. 2020; Zhao et al. 2020).

EZMOCK has several parameters that need to be calibrated to fit the provided reference model. Since our reference catalogue from SAM covers a wide redshift range of  $1.0 < z < 2.0$ , a single parameter set is not able to describe the evolution of the clustering. One solution is to measure the clustering in wider and overlapping redshift bins, measure the EZMOCK parameter values at different redshifts, and then interpolate to determine values at arbitrary redshifts where the clustering is measured; see for instance Ata et al. (2018). In this work, we calibrate the EZMOCK parameters at each redshift bin where the clustering signal is directly measured as shown in Section 2.1.

In order to calibrate the EZMOCK parameters, i.e. find a parameter set for EZMOCK such that it can reproduce clustering signal as the reference model, we adopt an emulator-like method. Traditional interpolation method requires thousands or more sampling in the parameter space to achieve reasonable and accurate prediction at arbitrary point. The emulator method based on Gaussian process (GP) or other machine-learning algorithm can significantly reduce the required number of training data. We find that for modelling of the EZMOCK within a four-dimensional parameter space, a training set with less than 100 data points is able to achieve an acceptable accuracy. At a given redshift, we first generate a parameter set through a Latin hyper-cube algorithm uniformly distributed in the parameter space as the training set and use the EZMOCK code to produce cubic mock for each model. In this calibration process, we use smaller box with boxsize  $750 \text{ Mpc } h^{-1}$  and grid size of  $240^3$  to accelerate the computation. Our test shows that it takes 10 CPU min to produce one mock. Due to finite size and galaxy number density, the clustering measured from these mocks is noisy. We model the corresponding uncertainty by randomly choosing a training model and regenerate

multiple mocks with different random seeds. We assume this can represent the level of error for the training set. The construction of the emulator follows the GP-based method as presented in Zhai et al. (2019b). Since the dimension of the parameter space of the model is significantly lower than the modelling of galaxy clustering at small scale, a simpler kernel function in the GP construction is sufficient in this work. We choose the squared exponential covariance function

$$k_{\exp}(r) = \exp\left(-\frac{r^2}{2l^2}\right), \quad (6)$$

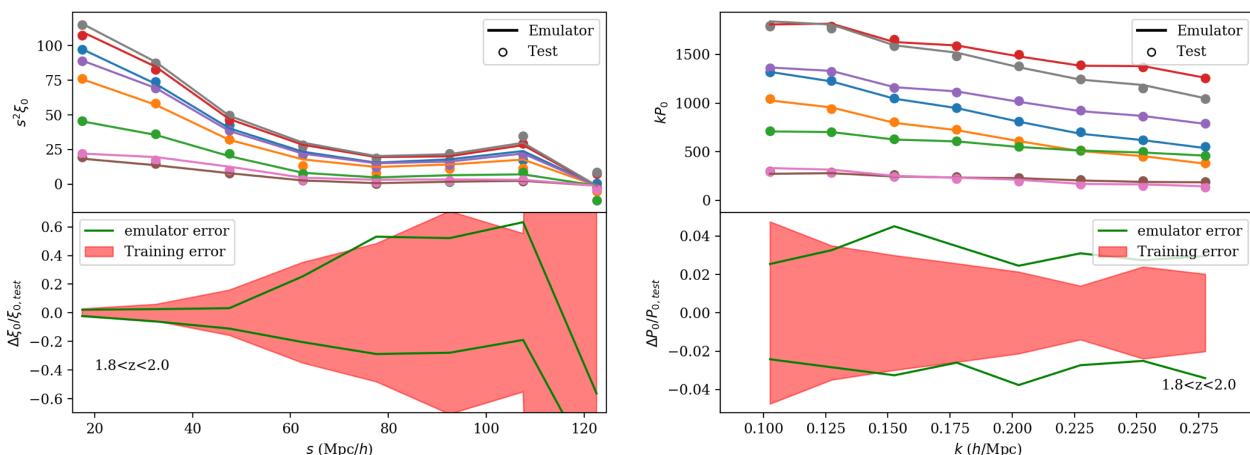
where the hyperparameter  $l$  defines the characteristic length-scale. We refer the readers to Zhai et al. (2019b) for more details in the construction of the emulator.

In order to test the accuracy of the emulator, we also generate 50–100 test points randomly distributed in the parameter space. Each model is also calculated by the EZMOCK code with the same boxsize and number of grid. In this work, we focus on the two-point statistics of the mock galaxies. So we construct the emulators to predict the clustering statistics including the monopoles of correlation function and power spectrum in both real and redshift space. Note that in the analysis of actual data, only the redshift space is used in both calibration and interpretation. In particular, we summarize the following statistics for the emulator construction:

(i)  $\xi_0(s)$ : monopole of 2-point correlation function (2PCF) in both real and redshift space, with a galaxy pair separation range of  $20 < s < 120 \text{ Mpc } h^{-1}$ .

(ii)  $P_0(k)$ : monopole of power spectrum in both real and redshift space, within range  $0.1 < k < 0.3 \text{ hMpc}^{-1}$ .

Adding higher Legendre multipoles, like quadrupole and hexadecapole, or higher order statistics like bispectrum can improve the calibration of the EZMOCK parameter. We find that the clustering statistics described above are sufficient for the current purpose. In Fig. 5, we show an example of the emulator performance for one redshift bin. The other redshift bins have similar results. In the top panel, we compare the results of the test sample that are directly calculated from EZMOCK as the truth, and the prediction from emulator. The overall error of the emulator is estimated as the 68 per cent distribution of the fractional error based on 50 test models. For 2PCF, we find that the emulator performance is better than the input training error at almost all the scales. The accuracy



**Figure 5.** Performance of the emulator built to calibrate the EZMOCK parameters. Only clustering statistics in redshift space for subsample within redshift  $1.8 < z < 2.0$  is shown for illustration. The results for other redshift bins are similar due to the same algorithm. Left: monopole of 2PCF. Right: monopole of power spectrum. The top panel shows a few test models randomly chosen in the parameter space. Dots stand for the measurement from the EZMOCK directly, while the solid lines are the prediction of the emulator. The bottom panels show the overall performance of the emulator as the solid green line. The shaded area is  $1\sigma$  uncertainty estimated by multiple generations with different random seeds, also as the training error for the emulator. The emulator error is 68 per cent error estimated from 50 test models.

for the power spectrum is worse; however, the uncertainty is better than 5 per cent at the scales of interest. We expect this error does not bias the calibration of the EZMOCK. In addition, we find that the error at large scale is significantly larger than intermediate and small scales; this is due to the sample variance of the finite volume, since the EZMOCK catalogue constructed for calibration has a boxsize of  $0.75 \text{ Gpc } h^{-1}$ .

With the emulators to generate predictions of 2PCF and power spectrum for an arbitrary EZMOCK parameter set, we can search for the optimized model that can produce consistent result as the reference galaxy mock from SAM. This can be easily done through a  $\chi^2$  computation

$$\chi^2 = (O_{\text{EZ}} - O_{\text{ref}})C^{-1}(O_{\text{EZ}} - O_{\text{ref}}), \quad (7)$$

where  $O_{\text{EZ,ref}}$  denotes the clustering statistics (2PCF or power spectrum) for EZMOCK prediction from emulator or reference model, and  $C$  is the corresponding covariance matrix. For simplicity, we just use the training error as for the emulator to get the covariance matrix here. We note that better estimate of this covariance matrix is possible by increasing the number of populations and can improve the search for the best-fitting parameters. However, this is not of critical importance as the calibration is only used to match the reference model as close as possible. In addition, we note that equation (7) can be directly used in a simple  $\chi^2$ -minimization algorithm, where the EZMOCK prediction is from the population itself instead of GP interpolation. However, this method may need more evaluations than the training and testing of the emulator for the model in this paper. The sample variance measured from this small volume can be significant and destroys the smoothness of the fitting function at some part in the parameter space, which makes the fitting function not well defined. On the other hand, the GP-based method has better behaviour and the toolkit is already well tested and implemented.

Since the EZMOCK only gives cubic mock, we need to truncate and convert the mock to have the same radial and angular distribution as the reference mock. We use the MAKE\_SURVEY toolkit (White, Tinker & McBride 2014) to rotate and trim the cubic mock. In order to embed the whole survey volume of Roman HLSS for H $\alpha$ -emitting galaxies in the simulation, we produce EZMOCK catalogue with a boxsize of  $3 \text{ Gpc } h^{-1}$  and a grid size of  $960^3$ . Using the

code MAKE\_SURVEY, we construct the mock containing the whole survey range within redshift  $1 < z < 2$  without reusing any region of the box. For each subsample of the clustering mock, we generate 1000 EZmocks, which is sufficient to estimate the covariance matrix for the clustering statistics.

In Fig. 6, we present the final measurement of the power spectrum from the calibrated EZmocks, which is truncated to have the same angular and radial distribution as the reference catalogue from Galacticus. The result shows the measurement of monopole, quadrupole, and hexadecapole in both real and redshift space. Note that the calibration of EZMOCK only uses the monopole; the consistency of higher Legendre multipoles between EZMOCK and Galacticus validates our calibration and modelling of these approximate mocks.

### 3 GALAXY CLUSTERING MODELLING

#### 3.1 Power spectrum model

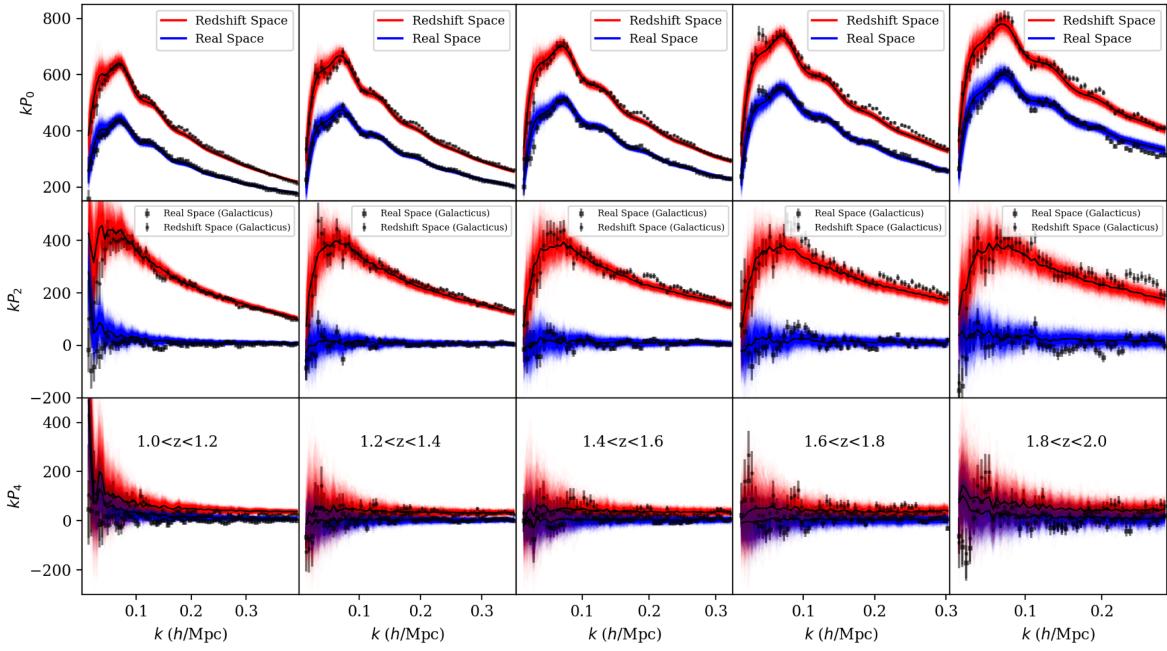
The Roman HLSS H $\alpha$  mock galaxy catalogue and its associated approximate galaxy mocks we have built enable the validation of the algorithm for extracting cosmological parameters. For investigating BAO/RSD measurements, we use a theoretical power spectrum to model the various known effects and compare with data, in deriving the measurement uncertainties on the cosmic distance scales and linear growth rate of large-scale structure. We model the galaxy power spectrum as Chuang et al. (2013), Wang (2014), and Hemantha, Wang & Chuang (2014)

$$P_{\text{dw}}^S(k, \mu, z) = P_{\text{dw}}(k, \mu, z) \frac{(1 + \beta \mu^2)^2}{1 + \frac{1}{2}(k \mu \sigma_v)^2}, \quad (8)$$

where  $\beta$  is the linear redshift distortion parameter and  $\sigma_v$  is the pairwise peculiar velocity (Hamilton 1998).  $P_{\text{dw}}(k, \mu, z)$  is the dewiggled linear galaxy power spectrum given by

$$P_{\text{dw}} = G^2 P_0 k^{n_s} T_{\text{dw}}(k, \mu, z), \quad (9)$$

where  $G(z)$  is the linear growth factor at redshift  $z$ , and  $n_s$  is the power-law index of the primordial power spectrum. In order to construct the de-wiggled transfer function, we apply the method as in Wang, Chuang & Hirata (2013)



**Figure 6.** The power spectrum measured from the reference catalogue and calibrated EZmocks. The EZmocks are generated with  $\text{boxsize} = 3 \text{ Gpc } h^{-1}$  to embed the whole *Roman* HLSS with  $1 < z < 2$ , and further truncated to have the same angular and radial distribution as the reference model. The results for different redshift bins are shown in each column. The top row shows monopole of galaxy power spectrum, the middle row shows quadrupole, and bottom row shows hexadecapole. The red lines are measurement from EZmocks in redshift space and blue lines are in real space. Note that the quadrupole and hexadecapole in real space are zero, which are also shown for completeness. The solid lines are the mean of the 1000 EZmocks and the error bars in the measurement of reference model are uncertainties estimated from EZmocks. The calibration of the EZMOCK only utilizes information from monopole; the consistency of all the three multipoles with the reference model shown here validates the modelling for the galaxy clustering.

$$T_{\text{dw}}(k, \mu, z) = T_{\text{lin}}^2(k, z) \exp\{-g_\mu k^2/k_*^2\} \quad (10)$$

$$+ T_{\text{nw}}^2(k, z)(1 - \exp\{-g_\mu k^2/k_*^2\}), \quad (11)$$

where the linear transfer function  $T_{\text{lin}}$  is calculated using CAMB (Lewis, Challinor & Lasenby 2000), the ‘no wiggle’ transfer function  $T_{\text{nw}}(k, z)$  is from equation (29) of Eisenstein & Hu (1998),  $k_*$  represents the scale of the non-linear effect on the BAO scales, and  $g_\mu$  describes the enhanced damping along the line of sight due to the enhanced power given by

$$g_\mu = G^2(z) \left\{ 1 - \mu^2 + \mu^2 \left[ 1 + f_g^*(z) \right]^2 \right\}, \quad (12)$$

where  $f_g^*$  is related to the linear growth rate.

Due to the non-linear evolution of the galaxy power spectrum, we use the model introduced by Cole et al. (2005) to correct the linear matter power spectrum. The resultant galaxy power spectrum can be written as

$$P_{\text{nl}}^S(k, \mu, z) = \frac{1 + Qk^2}{1 + Ak + Bk^2} P_{\text{dw}}^S(k, \mu, z), \quad (13)$$

where  $A$ ,  $B$ , and  $Q$  are constants. Sánchez, Baugh & Angulo (2008) fix  $B = Q/10$  in the modelling of the non-linear power spectrum at small scales. In the following analysis in this work, we also allow  $B$  to be a free parameter to make the non-linear correction more flexible.

We note that the RSD model for the galaxy power spectrum, equation (8), is not unique in the analysis of galaxy clustering, and different models have been utilized in literature. In this work, we also adopt a different model to compare the results. In particular, we

compare a model as demonstrated in Wang (2017):

$$P_{\text{dw}}^S(k, \mu, z) = P_{\text{dw}}(k, \mu, z)(1 + \beta \bar{W}(k, z)\mu^2)^2 \exp\left[-\frac{(k\mu\sigma_v)^2}{2}\right], \quad (14)$$

where the window function takes the form (Zhang, Pan & Zheng 2013; Zheng et al. 2013)

$$\bar{W}(k, z) = \frac{1}{1 + \Delta\alpha(z)\Delta^2(k, z)}, \quad (15)$$

where  $\Delta\alpha(z)$  is a parameter to be determined by observational data,  $\Delta^2(k, z) = k^3 P_{\text{lin}}/(2\pi^2)$ , and  $P_{\text{lin}}$  is the linear matter power spectrum. Compared with the model in equation (8), this model changes the RSD modelling in the galaxy clustering, and the galaxy peculiar velocity distribution is also changed. We denote equation (8) as Model A and equation (14) as Model B in our analysis. In Model A, the galaxy peculiar velocity distribution takes the usual form  $f(v) = (\sigma_v \sqrt{2})^{-1} \exp(-\sqrt{2}|v|/\sigma_v)$  which in Fourier space corresponds to one over the denominator in equation (8). In Model B, the galaxy peculiar velocity distribution takes the Gaussian form  $f(v) = 1/(\sigma_v \sqrt{2\pi}) \exp(-v^2/(2\sigma_v^2))$ .

In the actual measurement of the galaxy power spectrum, we need to assume a cosmological model to convert galaxy’s position from angular and redshift coordinates into 3D comoving coordinates. Therefore, a reference model or fiducial model is applied. However, it is possible that the fiducial model is different than the true cosmology; this can lead to a distortion in the observed galaxy power spectrum than the true power spectrum, known as AP effect (Alcock & Paczynski 1979). In order to take into account this effect in the modelling of the galaxy power spectrum, one can introduce two

scaling parameters

$$\alpha_{||} = \frac{H^{\text{fid}}(z)r_s^{\text{fid}}(z_d)}{H(z)r_s(z_d)}, \quad \alpha_{\perp} = \frac{D_A(z)r_s^{\text{fid}}(z_d)}{D_A^{\text{fid}}(z)r_s(z_d)}, \quad (16)$$

where  $H^{\text{fid}}(z)$  and  $D_A^{\text{fid}}$  are the Hubble parameter and angular diameter distance for the fiducial model, respectively, and  $r_s^{\text{fid}}(z_d)$  is prediction from the fiducial model for the sound horizon at the drag epoch. These two scaling parameters can relate the true wave vector component  $k'_{||}$  and  $k'_{\perp}$  with the observed ones:  $k'_{||} = k_{||}/\alpha_{||}$  and  $k'_{\perp} = k_{\perp}/\alpha_{\perp}$ . This can further transfer to the coordinates of  $k$  and  $\mu$  via

$$k' = \frac{k}{\alpha_{\perp}} \left[ 1 + \mu^2 \left( \frac{1}{F^2} - 1 \right) \right]^{1/2} \quad (17)$$

$$\mu' = \frac{\mu}{F} \left[ 1 + \mu^2 \left( \frac{1}{F^2} - 1 \right) \right]^{-1/2}, \quad (18)$$

where  $F = \alpha_{||}/\alpha_{\perp}$ . Note that the AP effect can also change the overall amplitude of the power spectrum multipoles due to the volume difference compared with the fiducial model, which is degenerate with the amplitude of the power spectrum template  $P_0$ .

Following Wang et al. (2013), we model the constraints on the growth of large-scale structure by measuring the *linear growth parameter*

$$f_g(z)\sigma_m(z) \equiv f_g(z)G(z)\sqrt{P_0}, \quad (19)$$

where  $P_0$  is the linear matter power spectrum normalization. The advantage of using  $f_g(z)\sigma_m(z)$  over  $f_g(z)\sigma_8(z)$  is that it is independent of the Hubble constant  $h$ , but contains the same amount of information on the growth of large-scale structure. In parallel with  $\alpha_{||}$  and  $\alpha_{\perp}$ , we define

$$\begin{aligned} \alpha_g &\equiv \frac{f_g(z)\sigma_m(z)}{f_{g,\text{fid}}(z)\sigma_{m,\text{fid}}(z)} = \frac{f_g(z)G(z)\sqrt{P_0}}{f_{g,\text{fid}}(z)G_{\text{fid}}(z)\sqrt{P_{0,\text{fid}}}} \\ &= \frac{\beta(z)\sqrt{P_n(z)}}{f_{g,\text{fid}}(z)G_{\text{fid}}(z)\sqrt{P_{0,\text{fid}}}}, \end{aligned} \quad (20)$$

where  $P_n(z) = P_0G(z)b(z)$  is the normalization factor of the galaxy power spectrum at each redshift slice, with  $b(z)$  denoting the galaxy bias factor, and  $\beta(z) = f_g(z)/b(z)$ . Both  $P_n(z)$  and  $\beta(z)$  can be measured from observational data.

### 3.2 Survey window function

Since the survey area in the simulation and actual observation is not regular, its effect on the measured galaxy power spectrum is not negligible. Applying this survey window function correction in configuration space is easy and straightforward, but becomes complicated in Fourier space. In our analysis, we adopt the correcting method for the Baryon Oscillation Spectroscopic Survey (BOSS) data (Beutler et al. 2017a,b) as suggested by Wilson et al. (2017).

The correction can be summarized as follows. We first calculate the power spectrum multipoles of a given model and Fourier transform them to get the correlation function multipoles  $\xi_{\ell}$ . Then, we multiply  $\xi_{\ell}$  with the window function to obtain the ‘convolved’ correlation function multipoles  $\xi_{\ell}^c$ . Finally, we Fourier transform back  $\xi_{\ell}^c$  to get the convolved power spectrum multipoles  $P_{\ell}^c$  that can be compared with the measurements from galaxy mocks or observational data. The multipoles of correlation function and power spectrum are connected through Hankel transformation, i.e.

$$P_{\ell}^c = 4\pi(-i)^{\ell} \int ds s^2 \xi_{\ell}^c(s) j_{\ell}(sk), \quad (21)$$

where  $j_{\ell}$  is the spherical Bessel function of order  $\ell$ .

For actual observational data or galaxy mock as in this paper, we can calculate the multipoles of the window function from the random pair counts  $RR(s, \mu)$  as (Beutler et al. 2017b)

$$W_{\ell}^2(s) \propto RR(s, \mu)L_{\ell}(\mu), \quad (22)$$

and the normalization is  $W_0^2(s \rightarrow 0) = 1$ . The application of the window function multipoles to the correlation function is given by

$$\xi_0^c = \xi_0 W_0^2 + \frac{1}{5} W_2^2 + \text{higher order term} \quad (23)$$

$$\xi_2^c = \xi_0 W_0^2 + \xi_2 \left[ W_0^2 + \frac{2}{7} W_2^2 \right] + \text{higher order term}. \quad (24)$$

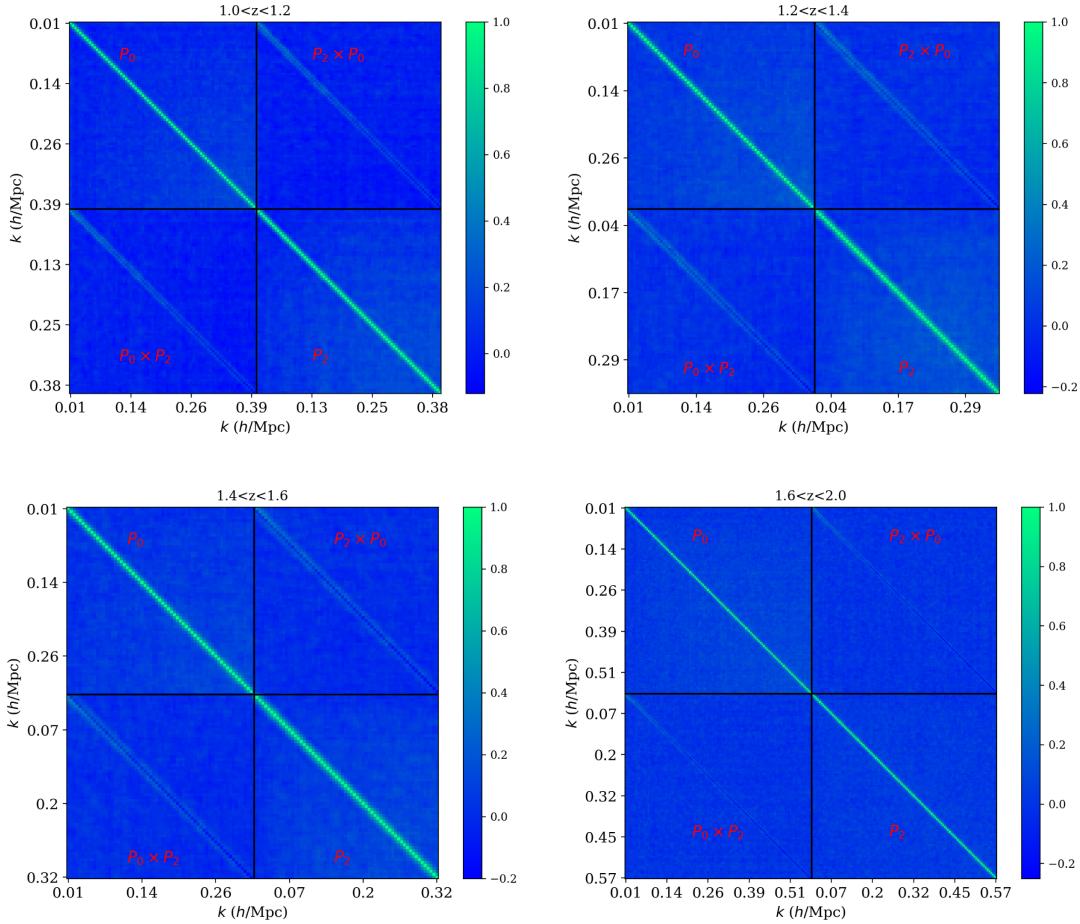
We are only interested in the monopole and quadrupole in this paper, so we ignored the contribution from hexadecapole and higher order multipoles to the correction. In this work, we perform the Hankel transformation of the correlation function and power spectrum by using the PYTHON package `Hankel` (Murray & Poulin 2019). In addition to the window function correction, we also consider the integral constraint bias as in Beutler et al. (2017b). For the galaxy mock in this paper, we find that this correction is only important at large scales. Since our following recovery analysis is limited to  $k > 0.02 h \text{ Mpc}^{-1}$ , the results are not affected significantly by this cut.

## 4 DATA ANALYSIS RESULTS

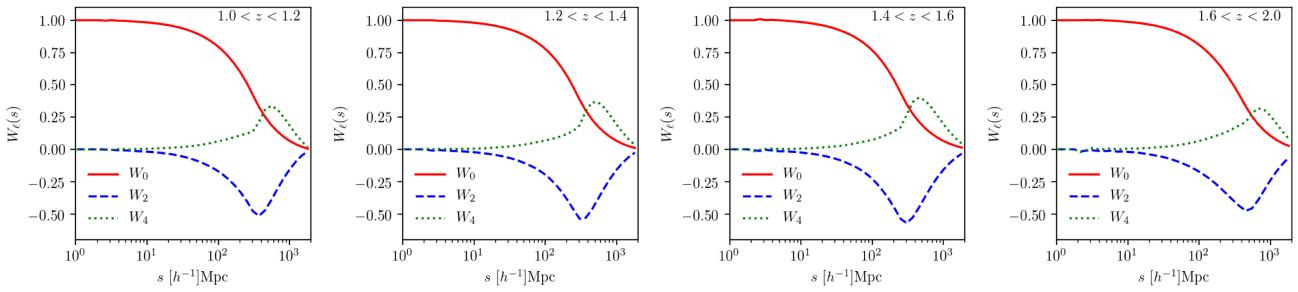
Given the mock catalogue for the estimate of the covariance matrix and the galaxy power spectrum model in Section 3.1, we perform a Markov chain Monte Carlo (MCMC) analysis to derive the parameter constraints. We choose the monopole and quadrupole of the galaxy power spectrum in redshift space as the clustering statistics. We note that the observables in configuration space can provide complementary information, and higher order statistics can tighten the cosmological constraints; we leave those for future work. Since the number densities of galaxy at redshifts  $1.6 < z < 1.8$  and  $1.8 < z < 2.0$  are significantly lower than those of the lower redshift counterparts, the measurement of the power spectrum is much noisier as in Fig. 6. Thus, we combine these two redshift bins in the analysis below.

In Fig. 7, we present the correlation matrix, i.e. normalized covariance matrix of monopole and quadrupole of the galaxy power spectrum in redshift space for each subsample. For each panel, the upper left corner shows the correlation between bins in monopole, the lower right corner presents correlation between bins in quadrupole, and the lower left and upper right show cross-correlation between monopole and quadrupole. In redshift space, we find that the  $k$  modes are not independent, and this correlation can increase as we go to smaller scales.

In Fig. 8, we present the first few multipoles of the window function as described in the previous section. The result shows similar behaviour as in the BOSS analysis (Beutler et al. 2017a, b). Obviously, the effect of the window function on the galaxy power spectrum can be non-trivial and scale dependent. As a sanity test, we present in Fig. 9 a comparison of the galaxy power spectrum with and without applying the window function correction. The result shows that the window function has a significant impact at large scale and becomes minor at  $k > 0.1 h \text{ Mpc}^{-1}$ , since the volume of the survey can limit the large-scale modes. We note that without window function correction, the modelled galaxy power spectrum is higher than the truth (corrected), which means a model needs lower amplitude of power spectrum at linear scale to fit the data. This impact can lead to an underestimate of the relevant quantity, for instance the linear growth rate. We will present a relevant comparison later in this paper.



**Figure 7.** Correlation matrix for monopole and quadrupole of the galaxy power spectrum in redshift space estimated from the EZmocks. The colour indicates the level of correlation. In each figure, the correlation of monopole is shown in the upper left corner, while the correlation for the quadrupole is in the lower right panels. The off-diagonal corner shows the cross-correlations.



**Figure 8.** Window function multipoles for the mock galaxy catalogues calculated through random pair counting. It shows the first few multipoles that are relevant to the analysis in this work. Each panel corresponds to a redshift slice as indicated in the upper right corner.

With the window function correction, we can compare the theoretical model with the observational data (here, the simulated *Roman* HLSS H $\alpha$  mock galaxy catalogue). We use MCMC likelihood analysis in the recovery test. The likelihood function can be written as

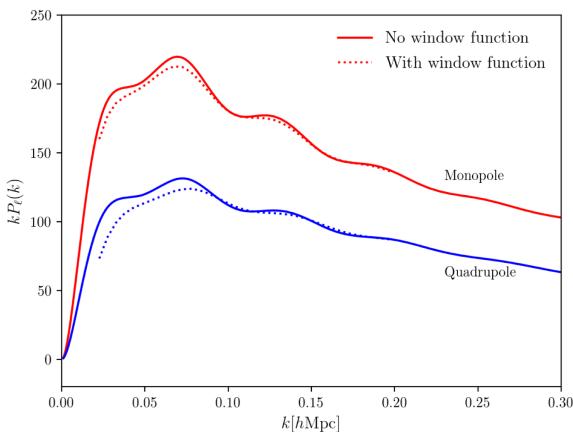
$$\chi^2 = \sum_{i,j} (P_{\text{obs},i} - P_{\text{th},i}) C_{ij}^{-1} (P_{\text{obs},j} - P_{\text{th},j}), \quad (25)$$

where  $P_{\text{obs}}$  is the galaxy power spectrum measured from the reference Galactus catalogue,  $P_{\text{th}}$  is the (window function corrected) prediction from the theoretical template as in Section 3.1, and  $C$  is the covariance matrix estimated from EZmocks. Since the estimated

covariance matrix is from mock catalogues, the inverse  $C^{-1}$  is biased compared with the true inverse covariance matrix (Hartlap, Simon & Schneider 2007). Thus, we need to correct for this bias by rescaling the inverse covariance matrix as

$$C_{\text{Hartlap}}^{-1} = \frac{N_s - n_b - 2}{N_s - 1} C^{-1}, \quad (26)$$

where  $N_s = 1000$  is the number of mocks used in the estimate of the covariance matrix, and  $n_b$  is the number of bins of galaxy power spectrum used in the likelihood analysis. For the galaxy samples analysed in this work, this correction can increase the final uncertainty by a few per cent.



**Figure 9.** Effect of the window function on the modelling of galaxy power spectrum, including monopole and quadrupole. The solid line is calculated from the power spectrum template as in Section 3.1, and the dotted line is after applying the window function correction as described in Section 3.2.

We use the PYTHON package EMCEE (Foreman-Mackey et al. 2013) to run the MCMC test and get the posterior of the parameters. In our method, the parameter set for the galaxy power spectrum is  $\{\alpha_{||}, \alpha_{\perp}, k_*, P_n, \beta, \sigma_v, A, Q, B, f_g^*\}$ . Note that  $f_g^*$  is only used in equation (12), and not used in measuring the growth rate  $f_g$ . We use uninformative and flat priors on these parameters throughout this work:  $\alpha_{||} = [0.2, 1.8]$ ,  $\alpha_{\perp} = [0.2, 1.8]$ ,  $k_* = [0.01, 0.8] h \text{Mpc}^{-1}$ ,  $\log P_n = [2, 10]$ ,  $\beta = [0.01, 2.0]$ ,  $\sigma_v = [0.1, 1000] \text{km s}^{-1}$ ,  $A = [0.5, 15]$ ,  $Q = [2.0, 30.0]$ ,  $B = [0.2, 10.0]$ , and  $f_g^* = [0.01, 2.0]$ . For Model B, we have an additional parameter  $\Delta\alpha(z)$  with prior  $[0, 2]$ .

We restrict the use of the galaxy power spectrum at  $0.02 < k < 0.3 h \text{Mpc}^{-1}$  in the analysis, indicating that the information at small scale is also used. In Fig. 10, we present the constraints on the scaling parameters  $\alpha_{||}$ ,  $\alpha_{\perp}$ , and  $\alpha_g$  after marginalizing over other parameters. The fiducial values of these parameters are recovered at  $1\sigma$ – $2\sigma$  level, which validates our construction of the galaxy mock and theoretical modelling. The results show that the two different models of the galaxy power spectrum can give parameter constraints that are consistent within  $1\sigma$ . For Model B, the uncertainty of the measurement is increased slightly due to the addition of an extra parameter.

Fig. 11 shows the 1D uncertainties of  $\alpha_{||}$ ,  $\alpha_{\perp}$ , and  $\alpha_g$  expected from the *Roman* HLSS H $\alpha$  galaxies. Within redshift range  $1 < z < 2$ , we find that the clustering analysis can give 2 per cent measurement of  $\alpha_{\perp}$ , i.e. the angular diameter distance. For  $\alpha_{||}$ , i.e. the Hubble parameter, the measurement is accurate at 3–6 per cent level depending slightly on the analysis template. For  $\alpha_g$  (see equation 20), i.e. the linear growth parameter  $f_g \sigma_m$  (see equation 19), the measurement is at roughly 7 per cent. Compared with the current measurements, *Roman* can give accurate measurement of these dark energy observables using a single tracer over a wide redshift coverage.

In Fig. 12, we investigate the effect of window function on the cosmological parameter constraint. As an example, we present the result from galaxies with  $1.0 < z < 1.2$  with Model A for the galaxy power spectrum. It shows that the window function has some impact on the estimate of radial distance scale and linear growth rate. For the former, the reason is that the maximum distance scale in the line-of-sight direction is smaller due to the thickness of the redshift shell, which means the window function starts to work at smaller scales than the angular direction. This leads to an overestimate of the Hubble

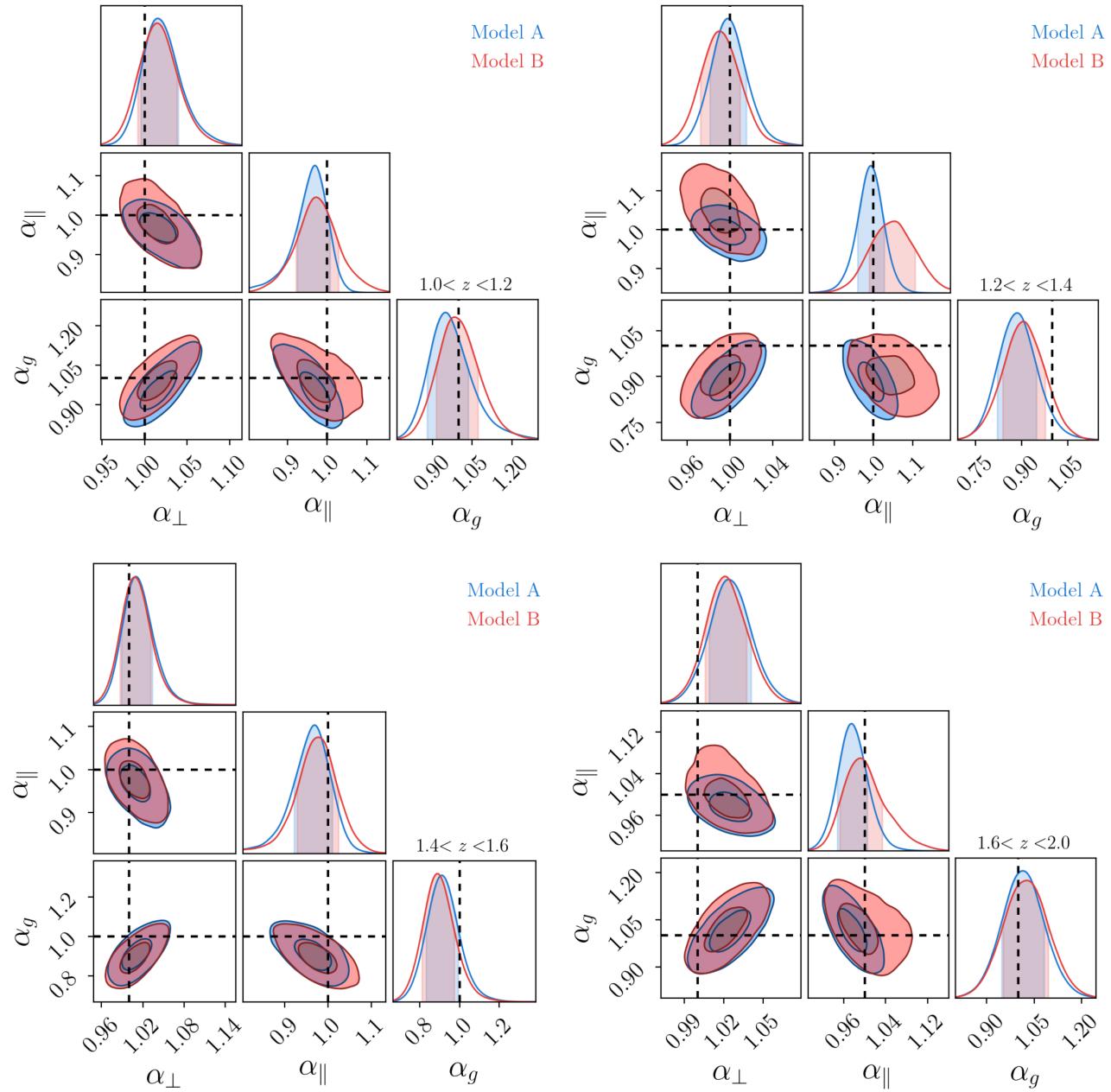
parameter at per cent level. The effect on the linear growth rate can be seen from Fig. 9 as well. Without window function correction, the model galaxy power spectrum can have a lower amplitude. According to equation (20), this leads to an underestimate of the linear growth rate. For galaxy mock in this paper, this change can be as large as 5 per cent. Although the effect of the window function on all the parameters is smaller than  $1\sigma$ , this is non-negligible for precision cosmology. In addition, the analysis in this paper uses information up to  $k = 0.3 h \text{Mpc}^{-1}$ ; the window function effect is reduced to some extent by incorporating this non-linear information. Therefore, it is likely that a smaller  $k_{\max}$  can lead to more biased result.

## 5 DISCUSSION AND CONCLUSION

Clustering of galaxies provides important information of the evolution of the universe, and is a powerful probe of dark energy. In particular, galaxy clustering measured from GRSs enables the measurement of cosmic expansion history and linear growth rate as functions of redshift, which can be used to determine the true nature of the observed cosmic acceleration. For next-generation surveys like those planned for *Roman* and *Euclid*, the observation of millions of galaxies can enable accurate measurements of the galaxy clustering signal, and the resulting cosmological measurements can help to constrain the dark energy and modified gravity models, or any deviation from standard  $\Lambda$ CDM model. In this paper, we use Galacticus, a calibrated SAM to generate a realistic synthetic catalogue of H $\alpha$  ELGs that *Roman* will observe in its HLSS. This catalogue has the same area on the sky, redshift range, and flux limit as specified by the current mission baseline design, and can serve as a reasonable reference to evaluate the expected performance. In particular, we have applied the calibrated SAM Galacticus of galaxy formation evolution from Zhai et al. (2019a) to the dark matter merger trees from the UNIT simulation described in Chuang et al. (2019) and built a light-cone catalogue of ELGs within redshift  $1 < z < 2$ . After applying a dust attenuation model, we select all the galaxies with H $\alpha$  flux higher than  $1 \times 10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$  to form our final galaxy sample.

We have measured the clustering statistics of this mock catalogue, including 2PCF and power spectrum. In order to validate our algorithm and forecast the dark energy measurements from *Roman* HLSS, we use the EZMOCK code to produce thousands of approximate but fast galaxy mocks to estimate the covariance matrix for the statistics of interest. These EZmocks are calibrated to match the clustering of the reference SAM catalogue. In particular, we employ an emulator approach to search the parameter space of EZMOCK and the resulting model can fit the monopole of 2PCF and power spectrum in both real and redshift space. The calibrated EZmocks are further truncated to have the same angular and radial distribution as the reference model, which makes the estimate of the covariance matrix straightforward. We note that the calibration of EZMOCK is not perfect given the uncertainty in the clustering measurement; this can be further improved if higher order statistics are taken into account.

With these mock galaxy catalogues, we perform a recovery test using the clustering signal in Fourier space. In particular, we compare two RSD models of the observed galaxy power spectrum and perform a likelihood analysis to get the posterior of the parameters. The result shows that the *Roman* GRS can measure the angular diameter distance at close or better than 2 per cent, the Hubble parameter can be measured at 3–6 per cent accuracy, and the error of the linear growth parameter is at 7 per cent. The comparison of the results from these two different RSD models indicates that RSD modelling has a



**Figure 10.** 2D and 1D distributions of the scaling parameters  $\alpha_{\parallel}$ ,  $\alpha_{\perp}$ , and  $\alpha_g$  from the MCMC analysis, marginalized over other parameters. We show contours at 68 and 95 per cent confidence levels with  $k_{\max} = 0.3$  for two different templates of galaxy power spectrum. The dashed lines represent no deviation from the fiducial model  $\alpha_{\parallel} = \alpha_{\perp} = \alpha_g = 1$ . The overall result validates the simulation and modelling of galaxy clustering from the Galacticus catalogue. The fiducial cosmology can be recovered at  $1\sigma$ – $2\sigma$  level for all the subsamples.

significant impact on the accuracy and precision of the dark energy measurements.

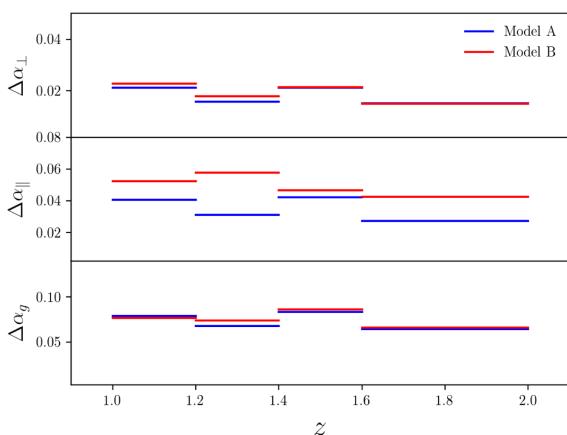
Our analysis has been performed using power spectrum due to the simplicity in calculating its theoretical prediction. It is also possible to analyse in configuration space, which can provide a consistency check on the cosmological constraints.

We have included the window function correction in the interpretation of the galaxy power spectrum. We find that not accounting for the window function correction leads to the underestimate of parameter errors, and a biased estimate of the linear growth parameter. Note that the actual *Roman* GRS will have a different footprint from what we have assumed here, and thus a different window function. In

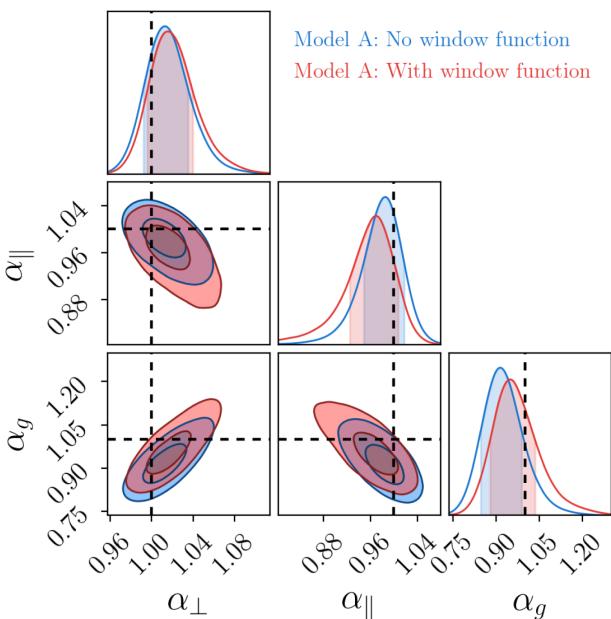
addition, systematic effects from slitless grism observations, such as line misidentification and spectral overlap, will need to be added to make our mock more realistic.

In future work, we plan to add realistic observational systematic effects to the *Roman* GRS mocks, and develop the theoretical model required to accurately model the *Roman* GRS data to extract precise and unbiased BAO/RSD measurements. The models used in this paper to fit the mock data illustrate the dependence of the results on the models.

The galaxy mock constructed in this work can be used beyond the two-point statistics analysis presented here. The constraining power on cosmological models from *Roman* H $\alpha$  ELGs can be further



**Figure 11.** Uncertainties of the scaling parameters measured from H $\alpha$ -emitting galaxies by RST, after projection from the constraint in the previous section. We display estimates of both templates of galaxy power spectrum.



**Figure 12.** Comparison of the constraint on the parameters  $\alpha_{||}$ ,  $\alpha_{\perp}$ , and  $\alpha_g$  with and without window function correction. This illustration uses galaxies with  $1.0 < z < 1.2$  and Model A for the template of galaxy power spectrum.

explored by higher order statistics, cross-correlation with other dark matter tracers, galaxy–galaxy lensing, statistical properties of the void distribution in the universe, and so on. Our mock catalogue enables the investigation of these topics for the *Roman* HLSS data. Such analysis will greatly strengthen the dark energy constraining power of *Roman* HLSS.

Note that our recovery test on the galaxy mock has been performed to measure the BAO and RSD signals. The actual measurements from *Roman* can be impacted by multiple factors. The dust attenuation can change the number and density of galaxies we can observe, which in turn can change the uncertainty of the clustering measurement, as well as reducing or increasing the shot noise, the survey volume may have an effect on the scales that we can probe and use, contamination from other emission lines such as [N II] may change the identification of H $\alpha$  ELGs and affect the bias of the galaxy sample, the actual survey strategy may be different from that assumed in this paper, a

shallower or deeper flux sensitivity can change the observation of faint galaxies, and the resulting galaxy sample due to these low-mass haloes is also affected. While a comprehensive exploration of these systematic effects on the cosmological constraint from the *Roman* GRS is beyond the scope of this paper, the *Roman* galaxy mock we have presented here enables realistic studies of cosmological constraints that can be expected from *Roman* HLSS.

We note that for the purpose of forecasting the cosmological constraints from a galaxy survey, the simple Fisher matrix formalism can be used to investigate the impact of changing the galaxy density, survey volume, scales of clustering in the analysis, and different types of dark energy properties and modified gravity theories. However, this method usually gives a lower limit of the uncertainties on the parameters of interest, which means that the forecast for the cosmological parameters is more idealistic than the simulation-based method as we applied in this paper. Our *Roman* HLSS H $\alpha$  galaxy mock can be used in carrying out sanity checks for Fisher matrix-based studies. Finally, the pipeline for the mock construction and data analysis presented in this work is directly applicable to the analysis of real observational data; thus, it lays a solid foundation for probing dark energy using future spectroscopic surveys.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for the valuable comments and suggestions that have helped us improve this paper. This work is supported in part by NASA grant 15-WFIRST15-0008, Cosmology with the High Latitude Survey Roman Science Investigation Team (SIT). GY would like to thank MICIU/FEDER (Spain) for financial support under project grant PGC2018-094975-B-C21. The UNIT simulations have been done in the MareNostrum Supercomputer at the Barcelona Supercomputing Center (Spain) thanks to the cpu time awarded by PRACE under project grant number 2016163937. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562 (Towns et al. 2014).

Software: PYTHON, MATPLOTLIB (Hunter 2007), NUMPY (van der Walt, Colbert & Varoquaux 2011), SCIPY (Jones et al. 2001), GEORGE (Ambikasaran et al. 2015), EMCEE (Foreman-Mackey et al. 2013), and HANKEL (Murray & Poulin 2019).

## DATA AVAILABILITY

The original dark matter halo catalogues are available from the UNIT simulation website. The galaxy mocks are available by request. A public webpage presenting the mocks will be available at a later time.

## REFERENCES

- Alam S. et al., 2017, *MNRAS*, 470, 2617
- Alcock C., Paczynski B., 1979, *Nature*, 281, 358
- Alonso D., 2012, preprint ([arXiv:1210.1833](https://arxiv.org/abs/1210.1833))
- Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O’Neil M., 2015, *IEEE Trans. Pattern Anal. Mach. Intell.*, 38, 252
- Anderson L. et al., 2014, *MNRAS*, 441, 24
- Ata M. et al., 2018, *MNRAS*, 473, 4773
- Bautista J. E. et al., 2018, *ApJ*, 863, 110
- Bautista J. E. et al., 2021, *MNRAS*, 500, 736
- Bennett C. L. et al., 2003, *ApJS*, 148, 97
- Benson A. J., 2012, *New Astron.*, 17, 175
- Beutler F. et al., 2011, *MNRAS*, 416, 3017
- Beutler F. et al., 2012, *MNRAS*, 423, 3430
- Beutler F. et al., 2014, *MNRAS*, 443, 1065

- Beutler F. et al., 2017a, *MNRAS*, 464, 3409  
 Beutler F. et al., 2017b, *MNRAS*, 466, 2242  
 Blake C. et al., 2011, *MNRAS*, 418, 1707  
 Chuang C.-H. et al., 2013, *MNRAS*, 433, 3559  
 Chuang C.-H. et al., 2019, *MNRAS*, 487, 48  
 Chuang C.-H., Wang Y., 2013, *MNRAS*, 435, 255  
 Chuang C.-H., Kitaura F.-S., Prada F., Zhao C., Yepes G., 2015, *MNRAS*, 446, 2621  
 Cole S. et al., 2005, *MNRAS*, 362, 505  
 Davis M., Peebles P. J. E., 1983, *ApJ*, 267, 465  
 Dawson K. S. et al., 2013, *AJ*, 145, 10  
 Dawson K. S. et al., 2016, *AJ*, 151, 44  
 de Mattia A. et al., 2020, preprint ([arXiv:2007.09008](https://arxiv.org/abs/2007.09008))  
 Delubac T. et al., 2015, *A&A*, 574, A59  
 Dressler A. et al., 2012, preprint ([arXiv:1210.7809](https://arxiv.org/abs/1210.7809))  
 eBOSS Collaboration, 2020, preprint ([arXiv:2007.08991](https://arxiv.org/abs/2007.08991))  
 Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605  
 Eisenstein D. J. et al., 2005, *ApJ*, 633, 560  
 Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J., 1994, *MNRAS*, 266, 50  
 Font-Ribera A. et al., 2014, *J. Cosmol. Astropart. Phys.*, 5, 027  
 Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306  
 Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, *MNRAS*, 388, 1473  
 Green J. et al., 2012, preprint ([arXiv:1208.4012](https://arxiv.org/abs/1208.4012))  
 Hamilton A. J. S., 1998, in Hamilton D., eds, *Linear Redshift Distortions: A Review*, The Evolving Universe. Astrophysics and Space Science Library, vol. 231, Springer, Dordrecht. p. 185  
 Hand N., Feng Y., Beutler F., Li Y., Modi C., Seljak U., Slepian Z., 2018, *AJ*, 156, 160  
 Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399  
 Hemantha M. D. P., Wang Y., Chuang C.-H., 2014, *MNRAS*, 445, 3737  
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90  
 Jones E. et al., 2001, SciPy: Open Source Scientific Tools for Python. available at: <http://www.scipy.org/>  
 Kaiser N., 1987, *MNRAS*, 227, 1  
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64  
 Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))  
 Laureijs R. et al., 2012, in Mark C. C., Giovanni G. F., Howard A. M., Jacobus M. O., eds, Proc. SPIE Conf. Ser. Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 84420T  
 Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473  
 Mehta V. et al., 2015, *ApJ*, 811, 141  
 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., 2019, *MNRAS*, 486, 5737  
 Murray S. G., Poulin F. J., 2019, *J. Open Source Softw.*, 4, 1397  
 Planck Collaboration I, 2014a, *A&A*, 571, A1  
 Planck Collaboration XVI, 2014b, *A&A*, 571, A16  
 Planck Collaboration XIII, 2016, *A&A*, 594, A13  
 Planck Collaboration VI, 2020, *A&A*, 641, A6  
 Raichoor A. et al., 2021, *MNRAS*, 500, 3254  
 Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, *MNRAS*, 449, 835  
 Sánchez A. G., Baugh C. M., Angulo R. E., 2008, *MNRAS*, 390, 1470  
 Sobral D. et al., 2009, *MNRAS*, 398, 75  
 Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, 428, 1128  
 Spergel D. N. et al., 2003, *ApJS*, 148, 175  
 Spergel D. et al., 2015, preprint ([arXiv:1503.03757](https://arxiv.org/abs/1503.03757))  
 Tamone A. et al., 2020, *MNRAS*, 499, 5527  
 Tegmark M., 1997, *Phys. Rev. Lett.*, 79, 3806  
 Towns J. et al., 2014, *Comput. Sci. Eng.*, 16, 62  
 van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22  
 Wang Y., 2014, *MNRAS*, 443, 2950  
 Wang Y., 2017, *MNRAS*, 464, 3005  
 Wang Y., Mukherjee P., 2007, *Phys. Rev. D*, 76, 103533  
 Wang Y., Chuang C.-H., Hirata C. M., 2013, *MNRAS*, 430, 2446  
 White M., Tinker J. L., McBride C. K., 2014, *MNRAS*, 437, 2594  
 Wilson M. J., Peacock J. A., Taylor A. N., de la Torre S., 2017, *MNRAS*, 464, 3121  
 York D. G. et al., 2000, *AJ*, 120, 1579  
 Zhai Z., Wang Y., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 005  
 Zhai Z., Blanton M., Slosar A., Tinker J., 2017, *ApJ*, 850, 183  
 Zhai Z., Benson A., Wang Y., Yepes G., Chuang C.-H., 2019a, *MNRAS*, 490, 3667  
 Zhai Z. et al., 2019b, *ApJ*, 874, 95  
 Zhang P., Pan J., Zheng Y., 2013, *Phys. Rev. D*, 87, 063526  
 Zhao C. et al., 2020, preprint ([arXiv:2007.08997](https://arxiv.org/abs/2007.08997))  
 Zheng Y., Zhang P., Jing Y., Lin W., Pan J., 2013, *Phys. Rev. D*, 88, 103510

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