Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **479,** 94–110 (2018)
Advance Access publication 2018 May 28

doi:10.1093/mnras/sty1389

Dark Energy Survey Year-1 results: galaxy mock catalogues for BAO

S. Avila,[1,2,3★] M. Crocce,[4] A. J. Ross,[5] J. García-Bellido,[2,3] W. J. Percival,[1] N. Banik,[6,7,8,9] H. Camacho,[10,11] N. Kokron,[10,11,12,13] K. C. Chan,[4,14] F. Andrade-Oliveira,[11,15] R. Gomes,[10,11] D. Gomes,[10,11] M. Lima,[10,11] R. Rosenfeld,[11,12] A. I. Salvador,[2,3] O. Friedrich,[16,17] F. B. Abdalla,[18,19] J. Annis,[7] A. Benoit-Lévy,[18,20,21] E. Bertin,[20,21] D. Brooks,[18] M. Carrasco Kind,[22,23] J. Carretero,[24] F. J. Castander,[4] C. E. Cunha,[25] L. N. da Costa,[11,26] C. Davis,[25] J. De Vicente,[27] P. Doel,[18] P. Fosalba,[4] J. Frieman,[7,28] D. W. Gerdes,[29,30] D. Gruen,[25,31] R. A. Gruendl,[22,23] G. Gutierrez,[7] W. G. Hartley,[18,32] D. Hollowood,[33] K. Honscheid,[5,34] D. J. James,[35] K. Kuehn,[36] N. Kuropatkin,[7] R. Miquel,[24,37] A. A. Plazas,[38] E. Sanchez,[27] V. Scarpine,[7] R. Schindler,[31] M. Schubnell,[30] I. Sevilla-Noarbe,[27] M. Smith,[39] F. Sobreira,[11,40] E. Suchyta,[41] M. E. C. Swanson,[23] G. Tarle,[30] D. Thomas,[1] A. R. Walker[42] and (The Dark Energy Survey Collaboration)[42]

Affiliations are listed at the end of the paper

Accepted 2018 May 21. Received 2018 April 20; in original form 2017 December 16

ABSTRACT

Mock catalogues are a crucial tool in the analysis of galaxy surveys data, both for the accurate computation of covariance matrices, and for the optimization of analysis methodology and validation of data sets. In this paper, we present a set of 1800 galaxy mock catalogues designed to match the Dark Energy Survey Year-1 BAO sample (Crocce et al. 2017) in abundance, observational volume, redshift distribution and uncertainty, and redshift-dependent clustering. The simulated samples were built upon HALOGEN (Avila et al. 2015) halo catalogues, based on a 2*LPT* density field with an empirical halo bias. For each of them, a light-cone is constructed by the superposition of snapshots in the redshift range $0.45 < z < 1.4$. Uncertainties introduced by so-called photometric redshifts estimators were modelled with a *double-skewed-Gaussian* curve fitted to the data. We populate haloes with galaxies by introducing a hybrid halo occupation distribution–halo abundance matching model with two free parameters. These are adjusted to achieve a galaxy bias evolution $b(z_{\rm ph})$ that matches the data at the 1σ level in the range $0.6 < z_{\rm ph} < 1.0$. We further analyse the galaxy mock catalogues and compare their clustering to the data using the angular correlation function $w(\theta)$, the comoving transverse separation clustering $\xi_{\mu<0.8}(s_\perp)$ and the angular power spectrum C_ℓ, finding them in agreement. This is the first large set of three-dimensional {RA,Dec.,z} galaxy mock catalogues able to simultaneously accurately reproduce the photometric redshift uncertainties and the galaxy clustering.

Key words: large-scale structure of Universe – cosmology: theory – cosmology: observations – methods: numerical.

1 INTRODUCTION

The large-scale structure (LSS) of the Universe has proven to be a very powerful tool to study cosmology. In particular, distance measurements of the Baryonic Acoustic Oscillation (BAO) scale

★ E-mail: santiagoavilaperez@gmail.com

© 2018 The Author(s)
Published by Oxford University Press on behalf of the Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

(Peebles & Yu 1970; Sunyaev & Zeldovich 1970) can be used to infer the expansion history of the Universe and, hence, to constrain dark energy properties. Whereas most BAO detections have been performed by spectroscopic galaxy surveys, able to estimate radial positions with great accuracy (Cole et al. 2005; Eisenstein et al. 2005; Beutler et al. 2011; Blake et al. 2011; Bautista et al. 2017; Ross et al. 2017a; Ata et al. 2018), the size and depth of the Dark Energy Survey[1] (DES) gives us the opportunity to measure the BAO angular distance $D_A(z)$ with competing constraining power using only photometry. Photometric galaxy surveys provide moderately accurate estimates of the redshift of galaxies from the magnitudes observed through a number of filters (5, in the case of DES; Hoyle et al. 2018), making it more difficult to obtain BAO measurements. However, the fidelity with which the BAO is observed is boosted by the photometric survey's capability to explore larger areas of the sky (1318 deg^2 for the BAO DES sample from data taken in the first year, ∼5000$ deg^2 for the complete survey) and larger number density of galaxies, reducing the shot noise. The current Year-1 (Y1) DES data already allow us to probe a range $0.6 < z < 1$ poorly explored with BAO physics.

This paper is released within a series of studies devoted to the measurement of the BAO scale with the DES Y1 data. The main results are presented in DES Collaboration (2017, hereafter DES-BAO-MAIN), including a ∼4 per cent precision D_A BAO measurement. Crocce et al. (2017) defines the sample selection optimized for BAO analysis (hereafter DES-BAO-SAMPLE). A photometric redshift validation over the sample is performed in Gaztañaga et al. (in preparation, hereafter DES-BAO-PHOTOZ). A method to extract the BAO from angular clustering in tomographic redshift bins is presented in Chan et al. (2018, from now DES-BAO-θ-METHOD). Ross et al. (2017b, DES-BAO-$s_\perp$-METHOD in the remainder) explains a method to extract the BAO information from the comoving transverse distance clustering. Camacho et al. (in preparation; DES-BAO-ℓ-METHOD from now), presents a method to extract the BAO scale from the angular power spectrum. This paper will be devoted to the simulations used in the analysis.

In order to analyse the data, we need an adequate theoretical framework. Even though there are analytic models that can help us understand the structure formation of the Universe (Zel'dovich 1970; Press & Schechter 1974; Kaiser 1984, 1987; Bond et al. 1991; Moutarde et al. 1991; Cooray & Sheth 2002), most realistic models are based on numerical simulations. Simulations have the additional advantages that they allow us to easily include observational effects such as masks and redshift uncertainties and can realistically mimic how these couple with other sources of uncertainty such as cosmic variance or shot noise. For the estimation of the covariance matrices of our measurements, we need a number of the order of hundreds to thousands of simulations, depending on the size of the data vector analysed (Dodelson & Schneider 2013), in order that the uncertainty in the covariance matrices is subdominant for the final results. As full *N*-Body simulations require considerable computing resources, running that number of *N*-Body simulations is unfeasible. Approximate mock catalogues are an alternative to simulate our data set in a much more computationally efficient way (Coles & Jones 1991; Bond & Myers 1996; Scoccimarro & Sheth 2002; Manera et al. 2013; Monaco et al. 2013; Tassev, Zaldarriaga & Eisenstein 2013; White, Tinker & McBride 2014; Avila et al. 2015; Chuang et al. 2015a,b; Kitaura et al. 2016; Monaco 2016). These methods are limited in accuracy at small scales, however, these methods have been shown to reproduced accurately the large scales (Chuang et al. 2015b). Alternatively, we can use a lower number of mock catalogues combining them with theory using hybrid methods (Scoccimarro 2000; Pope & Szapudi 2008; Taylor & Joachimi 2014; Friedrich & Eifler 2017), or methods that can re-sample *N*-Body simulations (e.g. Schneider et al. 2011). These alternatives would still rely on ∼100–200 simulations (see DES-BAO-θ-METHOD, or Schneider et al. 2011), in order to have subdominant noise in the covariance.

Galaxy mock catalogues are important in LSS studies not only for the computation of covariance matrices, but also crucial when optimizing the methodology and understanding the significance of any particularity found in the data itself, and learn how to interpret/deal with it (see e.g. Appendix A in DES-BAO-MAIN).

In this paper, we present a set of 1800 mock catalogues designed to statistically match those properties of the DES Y1-BAO sample. The main properties from the simulations that we need to match to the data in order to correctly reproduce the covariance are: the galaxy abundance, the galaxy bias evolution, the redshift uncertainties, and the shape of the sampled volume (angular mask and redshift range). The definition of the reference sample is summarized in Section 2. As a first step, we use the halo generator method called HALOGEN (Avila et al. 2015, summarized in Section 3.1), to create dark matter halo catalogues in Cartesian coordinates and fixed redshift. We then generate a light-cone (Section3.2) by transforming our catalogues to observational coordinates {RA, Dec., $z_{\rm sp}$}, accounting for redshift evolution, and implement the survey mask (Section 3.3). In Section 4, we model and implement the redshift uncertainties introduced in the sample by the photometric redshift techniques. The galaxy clustering model is described in Section 5, where we introduce a redshift evolving hybrid halo occupation distribution (HOD)–halo abundance matching (HAM) model. Finally, in Section 6 we analyse the set of mock catalogues, comparing their covariance matrices with the theoretical model in DES-BAO-θ-METHOD, and we compare the clustering measurements in angular configuration space ($w_i(\theta)$), three-dimensional configuration space ($\xi_{<\mu 0.8}(s_\perp)$) and angular harmonic space (C_l^i) of our mock catalogues with the data and theoretical models. We conclude in Section 7.

2 THE REFERENCE DATA

The aim of this paper is to reproduce in a cosmological simulation all the properties relevant for BAO analysis of the DES Y1-BAO sample. We describe how this sample is selected below in Section 2.1, and how the redshifts of that sample are obtained in Section 2.2. We also describe how we compute correlation functions from data or simulations in Section 2.3.

2.1 The Y1-BAO sample

The Y1-BAO sample is a subsample of the Gold Catalogue (Drlica-Wagner et al. 2018) obtained from the first Year of DES observations (Diehl et al. 2014). The Gold Catalogue provides 'clean' galaxy catalogues and photometry as described in Drlica-Wagner et al. (2018). A footprint quantified using a HEALPIX (Górski et al. 2005) map with $n_{\rm side} = 4096$ is provided with all the areas with at least 90 s of exposure time in all the filters g, r, i, and z, summing up to ∼1800 deg^2. After vetoing bright stars and the Large Magellanic Cloud, the area is reduced to ∼1500 deg^2.

[1] www.darkenergysurvey.org

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

The Y1-BAO sample selection procedure was optimized to obtain precise BAO measurements at high redshift and is fully described in DES-BAO-SAMPLE. The Y1-BAO sample is obtained applying three main selection criteria:

$$\begin{aligned} &17.5 < i_{\rm auto} < 19.0 + 3.0\, z_{\rm BPZ-MA} \\ &(i_{\rm auto} - z_{\rm auto}) + 2.0(r_{\rm auto} - i_{\rm auto}) > 1.7 \\ &0.6 < z_{\rm ph} < 1.0 \end{aligned} \tag{1}$$

with $X_{\rm auto}$ being the MAG_AUTO magnitude in the band X, $z_{\rm BPZ\text{-}MA}$ being the photometric redshift obtained by BPZ (Benítez 2000) using MAG_AUTO photometry, and $z_{\rm ph}$ being the photometric redshift (either $z_{\rm BPZ\text{-}MA}$ or $z_{\rm DNF\text{-}MOF}$; see below). Apart from the three main cuts in equation (1), we remove outliers in colour space and perform a star–galaxy separation. Further veto masks are applied to the Y1-BAO sample guaranteeing at least a 80 per cent coverage of each pixel in the four bands, requiring sufficient depth-limit in different bands and removing 'bad regions'. The final Y1-BAO sample after all the veto masks have been applied covers an effective area of 1318 deg^2 (see more details in DES-BAO-SAMPLE).

2.2 Photometric redshifts

The redshift estimation for each galaxy is based on the magnitude observed in each filter. For this paper, we will use two combinations of photometry and photo-z code, respectively: MAG_AUTO with BPZ (BPZ-MA), and MOF with DNF (DNF-MOF).

MAG_AUTO photometry is derived from the flux of the co-added image, as measured by the SEXTRACTOR software (Bertin & Arnouts 1996) from each of the bands. On the other hand, the MOF approach (Multi-Object Fitting, Drlica-Wagner et al. 2018) makes a multi-epoch, multiband fit to the shape of the object instead of on the co-added image as well as subtracting the light of neighbouring objects. The flux is fit with this common shape for each band separately.

A thorough description and comparison of both photometric redshift methods utilized here can be found in Hoyle et al. (2018). First, we have BPZ (Bayesian Photometric Redshift; Benítez 2000; Benítez et al. 2004), which is a method based on synthetic templates of spectra convolved with the DES filters, and makes use of Bayesian inference. On the other hand, we have DNF (Directional Neighbourhood Fitting; De Vicente, Sánchez & Sevilla-Noarbe 2016), which is a training-based method.

Both methods take the results from the chosen photometry in the four bands and give a probability distribution function (PDF) for the redshift of each galaxy: $P(z)$. As a full PDF for each galaxy would build up a very large data set to transfer and work with, here we take two quantities from each PDF: the mean $z_{\rm ph} \equiv \langle P(z) \rangle$, and a random draw from the distribution $z_{\rm mc}$. We will explain in Section 4 how we will model the effect of photometric redshift in our simulations.

By default, the reference data will use the DNF-MOF redshift $z_{\rm DNF\text{-}MOF}$, since this is the one used in DES-BAO-MAIN for the fiducial results. We will only include $z_{\rm BPZ\text{-}MA}$ in Section 4, since it was the reference redshift when part of the methodology presented in that section was designed.

2.3 Two-point correlation functions

Throughout this paper, we analyse two-point correlation functions in repeated occasions. In all cases, we use the Landy–Szalay estimator (Landy & Szalay 1993):

$$\Psi(\mathbf{x}) = \frac{DD(\mathbf{x}) - 2DR(\mathbf{x}) + RR(\mathbf{x})}{RR(\mathbf{x})} \tag{2}$$

with DD, DR, and RR being, respectively, the number of Data–Data, Data–Random, and Random–Random pairs separated by a distance $\boldsymbol{x}$. The correlation Ψ refers to either an angular correlation denoted by w, or a three-dimensional correlation ξ. The variable $\boldsymbol{x}$ may correspond to the angular separation θ projected on to the sky, or the three-dimensional comoving separation r. In the three-dimensional case, we will sometimes study the anisotropic correlation, distinguishing between the distance parallel to the line of sight and perpendicular to it, having $\boldsymbol{x} = \{s_{\parallel}, s_{\perp}\}$.

The data D may refer to observed data or simulated data. Random catalogues R are produced by populating the same sampled volume as the data with randomly distributed points. All the correlation function presented here were computed with the public code CUTE (Alonso 2012).[2]

3 HALO LIGHT-CONE CATALOGUES

Prior to the generation of the galaxy catalogues, we need to construct the field of dark matter haloes. For this, we will use the technique called HALOGEN (Avila et al. 2015), a technique that produces halo catalogues with Cartesian coordinates embedded in a cube and at a given time slice (*snapshot*). By superposing a series of HALOGEN snapshots, we construct an observational catalogue with angular coordinates and redshift {RA, Dec., $z_{\rm sp}$}: a light-cone halo catalogue. Finally, we describe how we implement the survey mask in the mock catalogues in order to statistically reproduce the angular distribution of the data.

3.1 HALOGEN

HALOGEN[3] is a fast approximate method to generate halo mock catalogues. It was designed and described in Avila et al. (2015), and compared with other methods in Chuang et al. (2015b). We summarize it here as four major steps:

(i) Generate a distribution of dark matter particles with second-order Perturbation Theory (Moutarde et al. 1991; Bouchet et al. 1995, 2LPT) at fixed redshift in a box of size $L_{\rm box}$. Distribute those particles on to a grid with cells of size $l_{\rm cell}$.

(ii) Produce a list of halo masses M_h from a theoretical/empirical halo mass function (HMF).

(iii) Place the haloes at the position of particles with a probability dependent on the cell density and halo mass as $P_{\rm cell} \propto \rho_{\rm cell}^{\alpha(M_h)}$. Within cells we choose random particles, while imposing an exclusion criterion to avoid halo overlap (using the $R_{200,\rm crit}$ derived from the halo mass). Mass conservation is ensured within cells by not allowing more haloes once the mass of the haloes surpasses the original dark matter mass.

(iv) Assign the velocities of the selected particles to the haloes rescaled through a factor: $\boldsymbol{v}_{\rm halo} = f_{\rm vel}(M_h) \cdot \boldsymbol{v}_{\rm part}$

There are one parameter, and two functions of halo mass that have been introduced in the method and need to be set for each run. We set the cell size $l_{\rm cell} = 5\,h^{-1}$ Mpc as in Avila et al. (2015). The parameter $\alpha(M_h)$ controls the halo bias and is fitted to a reference N-Body simulation to match the mass-dependent clustering. The factor $f_{\rm vel}(M_h)$ is also calibrated against an N-Body simulation in order to reproduce the variance of the halo velocities, crucial for the

[2] https://github.com/damonge/CUTE
[3] https://github.com/savila/HALOGEN

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

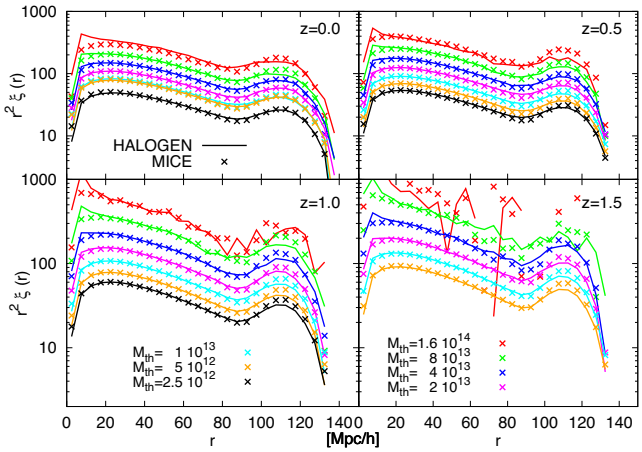

Figure 1. Two-point correlation function of MICE versus HALOGEN haloes in the simulation box at the snapshots $z = 0.0$, 0.5, 1.0, and 1.5 as labelled. We display the different mass thresholds $M_{\rm th}$ used during the fit (finding higher correlations for higher $M_{\rm th}$). Note that correlations have been multiplied by r^2 to highlight the large scales.

redshift-space distortions. For this study, we use the MICE simulation as a reference for this calibration.

The MICE Grand Challenge simulation, described in Fosalba et al. (2015a), Crocce et al. (2015), and Fosalba et al. (2015b), is based on a cosmology with parameters: $\Omega_{\rm M} = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_{\rm b} = 0.044$, $h = 0.7$, $\sigma_8 = 0.8$, $n_s = 0.95$ matching early *WMAP* data (Hinshaw et al. 2007). We will use this fiducial cosmology throughout the paper. The box size of the simulation is $L_{\rm box} = 3072\,h^{-1}$ Mpc and made use of 4096^3 particles. The HALOGEN catalogues use a lower mass resolution with 1280^3 particles in order to reduce the required computing resources. We use the same box size and cosmology for HALOGEN.

For the calibration, we used the same phases of the initial conditions as the *N*-Body simulation and fitted the HALOGEN parameters with the snapshots at $z_{\rm snap} = 0$, 0.5, 1.0, 1.5. We input to HALOGEN a hybrid HMF, using the HMF directly measured from MICE catalogues at low masses, while using an analytic expression (Watson et al. 2013) generated with HMFCALC[4] (Murray, Power & Robotham 2013) for the large masses, where the HMF from the MICE catalogues are noise dominated.

We fitted the clustering of the haloes in logarithmic mass bins (with factor 2 in mass threshold), this being a slight variation with respect to the method used in Avila et al. (2015). The minimum halo mass that we probe is $M_h = 2.5 \times 10^{12}\,{\rm M}_\odot\,h^{-1}$ for snapshots at $z \leq 1.0$, whereas we use a minimum mass of $M_h = 5.0 \times 10^{12}\,{\rm M}_\odot\,h^{-1}$ for higher redshift snapshots. Once the parameter calibration is finished, we find a good agreement between MICE and HALOGEN correlation functions as a function of redshift and mass, as shown in Fig. 1.

3.2 Light-cone

We place the observer at the origin (i.e. one corner of the box), so that we can simulate one octant of the sky, and transform to spherical coordinates. We use the notation $z_{\rm sp}$ for the redshift of a halo or galaxy as it would be observed with a spectroscopic survey (i.e. with negligible uncertainty). We have that

$$z_{\rm sp} = z(r) + \frac{1 + z(r)}{c}\boldsymbol{u}\cdot\hat{r}\,, \tag{3}$$

with $\boldsymbol{r} = \{X, Y, Z\}$ the comoving position, $\boldsymbol{u}$ the comoving velocity, $r = |\boldsymbol{r}|$, $\hat{r} = \boldsymbol{r}/r$ and $z(r)$ the inverse of

$$r(z) = c\int_0^z \frac{{\rm d}z'}{H(z')}\,. \tag{4}$$

The first term in equation (3) corresponds to the redshift due to the Hubble expansion, whereas the second term is the contribution from the peculiar velocity of the galaxy.

Given the redshift range covered by DES, we need to allow for redshift-dependent clustering, and hence we will let the HALOGEN parameters (α, $f_{\rm vel}$, and the HMF) vary as a function of redshift.

For that, we interpolate α, $f_{\rm vel}$, and $\log n$ (the logarithm of the HMF) using cubic splines from the reference redshifts $z_{\rm snap} = 0$, 0.5, 1, 1.5 at which parameters were fitted to our output redshifts $z_{{\rm snap},\,i} = 0.3$, 0.55, 0.625, 0.675, 0.725, 0.775, 0.825, 0.875, 0.925, 0.975, 1.05, and 1.3, then HALOGEN was run at those redshifts. We build the light-cone from the superposition of spherical $z_{\rm sp}$ shells drawn from the snapshots by setting the edges at the intermediate redshifts. So each snapshot i contributes galaxies whose redshift is in the interval $z_{\rm sp} \in [\frac{z_{{\rm snap},i-1}+z_{{\rm snap},i}}{2}, \frac{z_{{\rm snap},i}+z_{{\rm snap},i+1}}{2}]$, also imposing the edges of the light-cone at $z_{\rm sp} = 0.1$ and $z_{\rm sp} = 1.42$ (which is the maximum redshift reachable given the chosen geometry and cosmology). *A priori*, there will be a relatively sharp transition of the clustering properties at the edges of the $z_{\rm sp}$ shells, but, once we have introduced the redshift uncertainties in Section 4, those transitions will be smoothed. Throughout Sections 3–5, we will focus the analysis in eight photometric redshift bins with width $\Delta z_{\rm ph} = 0.05$ between $z_{\rm ph} = 0.6$ and $z_{\rm ph} = 1.0$. When dealing with true redshift space, we will need to extend the boundaries to the range $0.45 < z_{\rm sp} < 1.4$ (see Section 4).

Finally, we compare the resulting HALOGEN light-cone with the halo light-cone generated by MICE in Fig. 2. Note that the MICE simulated light-cone is constructed from fine timeslices ($\Delta z = 0.005$–0.025) built on-the-fly from a full *N*-Body simulation and using the velocity of the particles to extrapolate their positions at the precise moment they cross the light-cone (Fosalba et al. 2008). Remarkably, despite the large differences in the methodology, the angular correlation functions from both light-cones show very good agreement at all redshifts, independently of whether the HALOGEN parameters were fitted or interpolated. At large scales sampling variance becomes dominant, modifying stochastically the shape of the correlation function, but since we imposed the same phases of the initial conditions, it enables us to make a one-to-one comparison.

3.3 Angular selection

In Section 3.2, we placed the observer at one corner, providing a sample covering an octant of the sky. In Section 2, we described how we obtained a footprint covering the effectively observed area of the sky. This mask spreads over a large fraction of the sky and cannot be fitted into a single octant. However, using the periodic boundary conditions of the box, we can put eight replicas of the box together to build a larger cube, and extract a full sky light-cone catalogue, albeit with a repeating pattern of galaxies.

In Fig. 3, we show how we can draw eight mock catalogues with the Y1 footprint from the full sky catalogue by performing rotations on the sphere. We see that the footprint has a complicated shape with two disjoint areas: one passing by {RA = 0°, Dec. = 0°} known

[4]http://hmf.icrar.org

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

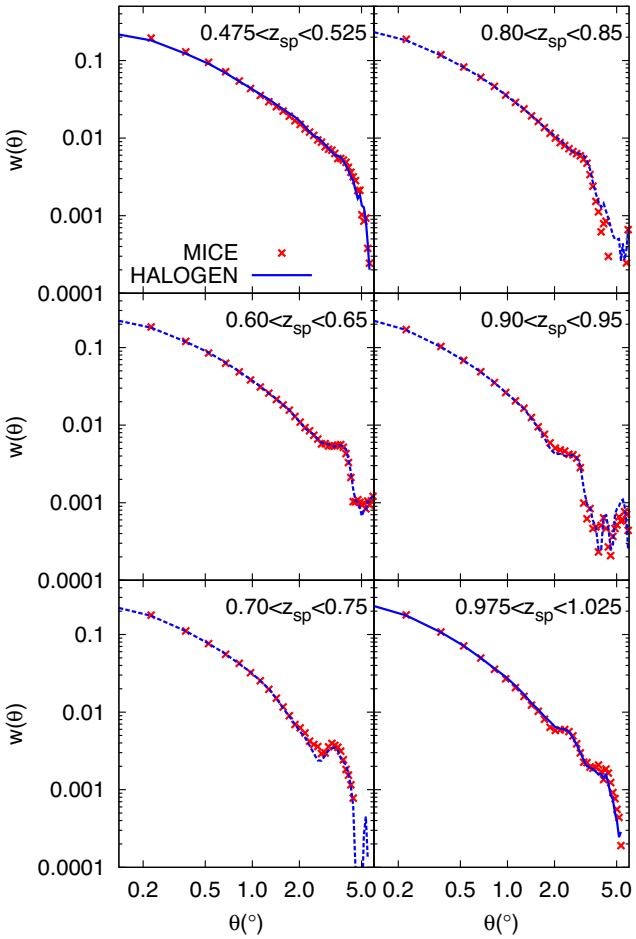

Figure 2. Angular correlation function of haloes from the MICE (crosses) and HALOGEN (lines) light-cones. The different panels correspond to different redshift bins, as labelled, with width $\Delta z_{\rm sp} = 0.05$. We mark in solid lines the redshifts at which HALOGEN parameters were fitted, and in dashed lines results from interpolated parameters. Correlation functions shown correspond to mass cuts at $M_h = 2.5 \times 10^{12}\,{\rm M}_\odot\,h^{-1}$.

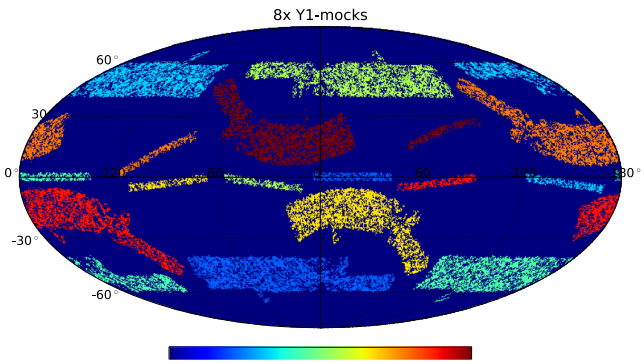

Figure 3. Mask Rotations. Representation of the full sky with the eight rotations of the mask used to generate the mock catalogues. Each mask, with two disjoint parts, is represented in a different colour, being the original data mask the one passing through $\{0^\circ, 0^\circ\}$ and $\{0^\circ, -60^\circ\}$ in {RA,Dec.}. The mock galaxies selected by the rotated masks are, then, rotated to the position of the original mask. Having the correct angular selection is one of the key ingredients to ensure that the mocks will give us the correct covariance matrices.

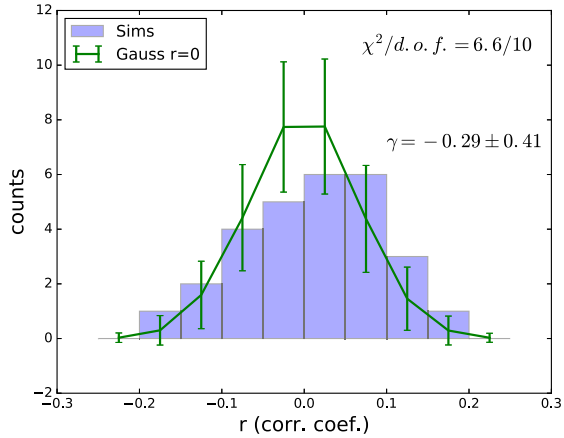

Figure 4. Distribution of correlation coefficients of α (see the text) between different mask rotations for the HALOGEN mock catalogues ('Sims'). We compare it to the expected distribution of correlations if α followed a Gaussian distribution and results from different masks were uncorrelated ('Gauss r=0'). We find that this hypothesis (with no free parameter) to be compatible with our simulations. We also find that the observed skewness in the distribution γ is within the expected statistical noise level under the same hypothesis.

as Stripe-82 (that overlaps with many other surveys), and another passing by $\{{\rm RA} = 0^\circ, {\rm Dec.} = -60^\circ\}$, known as SPT-region (due to the overlap with the South Pole Telescope observations). While designing the rotations depicted in Fig. 3, we made sure that every pair of footprints would not overlap and that the two disjoint areas are separated by more than the maximum scale of interest ($\sim 6^\circ$, see Section 6.2).

Including this angular selection to the mock catalogues will be essential since, as shown in Section 6, it has an important effect in the covariance matrices.

Given the repetition of boxes, one could be concerned about the effect it could have in our measurements. Qualitatively, we do not expect this to be important for a series of reasons. First, the repetition occurs at very large scales ($L = 3072\,h^{-1}$ Mpc) and we only use eight replicas. This makes difficult for structures to be observed more than once, and if they are, it will always be done from a different orientation and at a different redshift. On top of that, there are three stochastic processes that will make the hypothetically repeated structure appear differently: the halo biasing (since the structure would be at different redshift, it would be drawn from a different snapshot, see Section 3.1), the redshift uncertainties (Section 4) and the galaxy assignment to haloes (Section 5).

More quantitatively, we study the correlation coefficients between the measured BAO scale α (defined in equation 20 of DES-BAO-MAIN) from mocks coming from the same box but different mask rotation. The distribution of the 28 (=7 × 8/2) correlation coefficients indicate very small correlation ranging from $r = -0.2$ to $r = 0.2$ as shown (as a histogram) in Fig. 4. In order to study if these correlations r represent any significance given the number of mocks we used ($N_{\rm mocks} = 1800$), we generate $N_{\rm mocks}$ Gaussian realizations of α, distribute them in eight groups and compute the correlation coefficients between them. We repeat this process $N_{\rm rep} = 1000$ times, computing for each realization the distribution of r. The mean (and 1σ error bar) is also shown in that panel. We compute the goodness of the model using the covariance between the $N_{\rm rep}$ realizations and find $\chi^2/{\rm dof} = 6.6/10$, showing that the distribution of r in our

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

simulations is completely consistent with the null hypothesis of α being uncorrelated. We note the r distribution of the simulations is skewed towards positive values: $\gamma = -0.29$. Nevertheless, this γ value is compatible with simply being a statistical fluctuation, since its absolute value is lower than the standard deviation $\Delta\gamma = 0.41$ of our $N_{\rm rep}$ realizations.

4 PHOTOMETRIC REDSHIFT MODELLING

A photometric survey like DES can determine the redshift of a galaxy with a limited precision. Typically $\sigma_z/(1+z) \approx 0.017$–$0.05$ (Rozo et al. 2016; Drlica-Wagner et al. 2018), depending on the sample and the algorithm. This will have a major impact on the observed clustering and has to be included in the model. The aim of this section is to apply the effect of the redshift uncertainties to the mock catalogues. In particular, we model the number density of galaxies as a function of *true* redshift $z_{\rm sp}$ and *observed* redshift $z_{\rm ph}$: $n(z_{\rm ph}, z_{\rm sp})$.

In Section 2.2, we explained that we will work with two sets of photometric redshifts with different properties: BPZ-MA and DNF-MOF. As we do not know the true redshift $z_{\rm sp}$ of the galaxies in the data, we need to be careful when estimating $n(z_{\rm ph}, z_{\rm sp})$ from the data. For both sets of photometric redshifts, we will follow the same methodology starting from the PDF (or $P(z_{\rm sp})$) of each galaxy.

As we mentioned in Section 2.2, we take the *observed* redshift $z_{\rm ph}$ as the mean of the PDF $z_{\rm ph} = \langle P(z_{\rm sp})\rangle$, and we also select a Monte Carlo random value from the distribution $P(z_{\rm sp}) \rightarrow z_{\rm mc}$, which will be used to estimate $n(z_{\rm ph}, z_{\rm sp})$. This way, stacking the PDFs of a large number of galaxies will be statistically equivalent to taking the normalized histogram of their $z_{\rm mc}$. Hence, we can use

$$n(z_{\rm ph}, z_{\rm sp}) = n(z_{\rm ph}, z_{\rm mc})\ ,\ , \tag{5}$$

being able to estimate the right-hand side from the data and apply the left-hand side to the simulations.

One could alternatively use the training sample to estimate directly $n(z_{\rm ph}, z_{\rm sp})$. However, the technique explained above has the advantage of dividing the problem into two distinct steps: one in which the photo-z codes are calibrated and validated, and one in which the science analysis from the photo-z products is performed. We discuss at the end of this section the validation of the chosen method performed in DES-BAO-PHOTOZ. Note also that, for the mocks design, the actual definition of $z_{\rm ph}$ is not relevant as long as the $P(z_{\rm ph}|z_{\rm sp})$ is known.

We select from the data thin bins of width $\Delta z_{\rm mc} = 0.01$ and measure $\partial N/\partial z_{\rm ph}$. We denote N as the total number of observed galaxies in our sample (or equivalently in our mock catalogues) and $n = {\rm d}N/{\rm d}V$ as the number density or abundance. Equivalently to equation (5), we can use

$$\left.\frac{\partial N}{\partial z_{\rm ph}}\right|_{z_{\rm sp}} = \left.\frac{\partial N}{\partial z_{\rm ph}}\right|_{z_{\rm mc}} \quad . \tag{6}$$

From now we drop the $z_{\rm mc}$ notation, since we will always be looking at distributions of $z_{\rm mc}$ (never individual values), which are equivalent to the distributions of $z_{\rm sp}$. Additionally, the focus of this paper is the simulations for which we will only have $z_{\rm sp}$.

In the top panel of Fig. 5, we present different fits to the BPZ-MA data for the case $\left.\frac{\partial N}{\partial z_{\rm ph}}\right|_{z_{\rm sp}=0.85}$. The fitting functions can be described

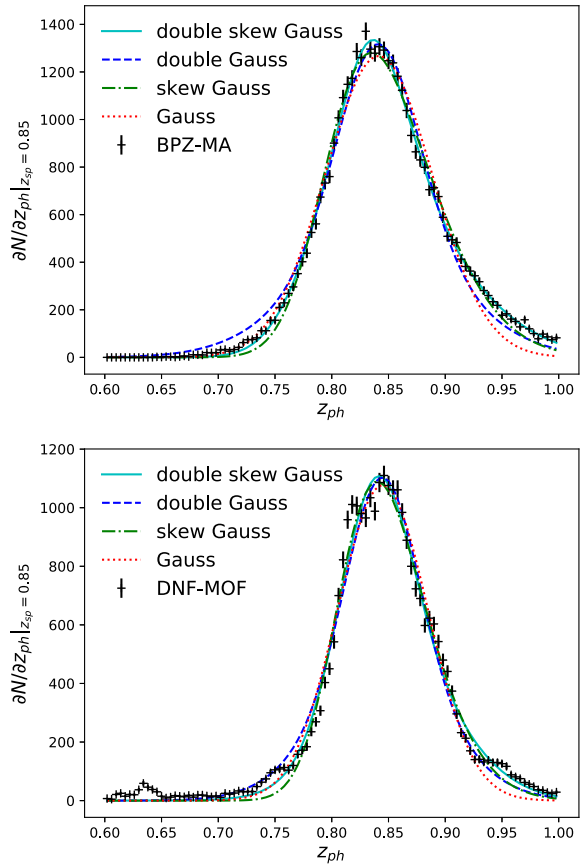

Figure 5. Abundance of galaxies at fixed true redshift $z_{\rm sp}$ as a function of photometric redshift $z_{\rm ph}$ (estimated from the joint $\{z_{\rm mc}, z_{\rm ph}\}$ distribution, see text and equation 6). Comparison of different fitting functions from equation (7) (lines) to the data (error bars representing the Poison noise). In the top panel we fit the BPZ-MA data, whereas in the bottom panel we fit DNF-MOF (default) data.

by

$$\begin{aligned}
&\left.\frac{\partial N}{\partial z_{\rm ph}}\right|_{z_{\rm sp}} = A\cdot P(z_{\rm ph}|z_{\rm sp}) \quad \text{with}\\
&P(z_{\rm ph}|z_{\rm sp}) = \frac{1-r}{\sqrt{2\pi\sigma_1^2}}e^{-(z_{\rm ph}-\mu)^2/(2\sigma_1^2)}\\
&+\ \frac{r}{\sqrt{2\pi\sigma_2^2}}e^{-(z_{\rm ph}-\mu)^2/(2\sigma_2^2)}\cdot {\rm erf}\left(\frac{\gamma-\mu}{\sqrt{2}\cdot\sigma_2}\right),\\
&r = 0, 0.5, 1
\end{aligned} \tag{7}$$

with different choices of parameters. All A, μ, σ_1, σ_2, γ and r depend implicitly on redshift $z_{\rm sp}$.

The simplest case is a Gaussian ($\gamma = 0$, $\sigma_2 = 0$, $r = 0$). In order to include a term of kurtosis to the fit, we can extend the curve to a double Gaussian ($\gamma = 0$, $\sigma_2 \neq 0$, $r = 0.5$). We can also introduce skewness with the *skewed* Gaussian ($\gamma \neq 0$, $\sigma_2 = 0$, $r = 1$). Finally, the most general case considered here is the *skewed double* Gaussian ($\gamma \neq 0$, $\sigma_2 \neq 0$, $r = 0.5$). Note that for skewed curves, the parameter μ does not represent the mean.

Fig. 5 and Table 1 show that the skewed double Gaussian does make a significant improvement to the goodness of fit for the BPZ-MA data (improving a factor of ∼3 to ∼7). However, the improvement is less significant for the DNF-MOF data, which shows more

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

Table 1. Relative goodness of the fits shown in Fig. 5. The reduced χ^2 were computed only taking into account Poisson noise of the histograms. Hence, their absolute values are not significant, but their relative values (normalized to the minimum in the table) give us an idea of the improvement in the fits when adding more parameters. For BPZ-MA, the improvement is significant when including more degrees-of-freedom, whereas for DNF-MOF the improvement is less pronounced.

Curve	$\chi^2_{\rm BPZ-MA}/\chi^2_{\rm ref}$	$\chi^2_{\rm DNF-MOF}/\chi^2_{\rm ref}$
Gaussian	7.7	5.9
Double Gaussian	4.9	5.0
Skewed Gaussian	3.0	4.1
Skewed Double Gaussian	1	3.7

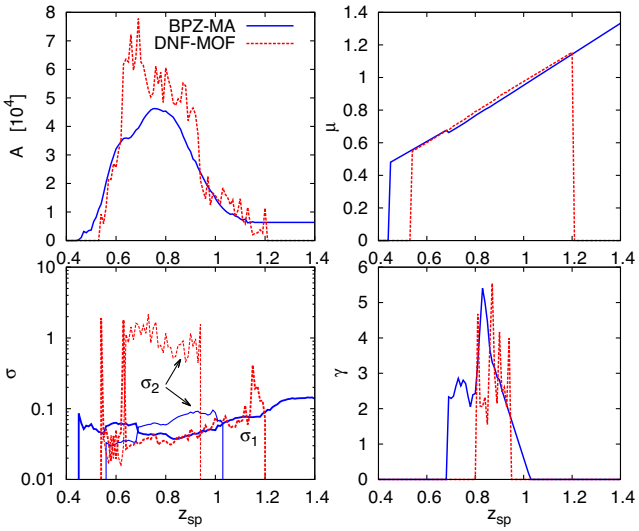

Figure 6. Best-fitting parameters of the fitting function described in equation (7) as a function of true redshift $z_{\rm sp}$ for both data sets. See the text for details.

small-scale structure in the curves and is consequently more difficult to model. We discuss at the end of the section the origin of this structure.

We repeat the fit performed in Fig. 5 for $z_{\rm sp} = 0.85$ at all redshifts $z_{\rm sp}$. In some cases, the degeneracy between parameters makes two adjacent $z_{\rm sp}$ bins have quite different set of best fit parameters, even if the shape of the curves are similar. In order to mitigate that, we reduce the degrees of freedom in the fits when possible. We restrict the values of r to 0, 0.5, and 1.0 according to the values of σ_1 and σ_2. We also parametrize the evolution with redshift of the parameters and fix the values of these additional parameters. For example, we fix $\mu(z_{\rm sp})$ to a straight line for most of the $z_{\rm sp}$ range, and we also set $\gamma = 0$ or $\sigma_2 = 0$ where they stop improving the fits. Fig. 6 shows the evolution of the fitted parameters that we obtain. For BPZ-MA, the fits converge more easily and we can find relatively smooth evolution. For DNF-MOF, the curves have more small-scale structure, and the evolution of the best-fitting parameters inherits that structure.

The amplitude $A(z_{\rm sp})$ in equation (7) gives us the abundance of objects by simply dividing by the survey area,

$$\frac{{\rm d}n_A(z_{\rm sp})}{{\rm d}z} = \frac{1}{\Omega_{\rm sur}} A(z_{\rm sp}) \qquad (8)$$

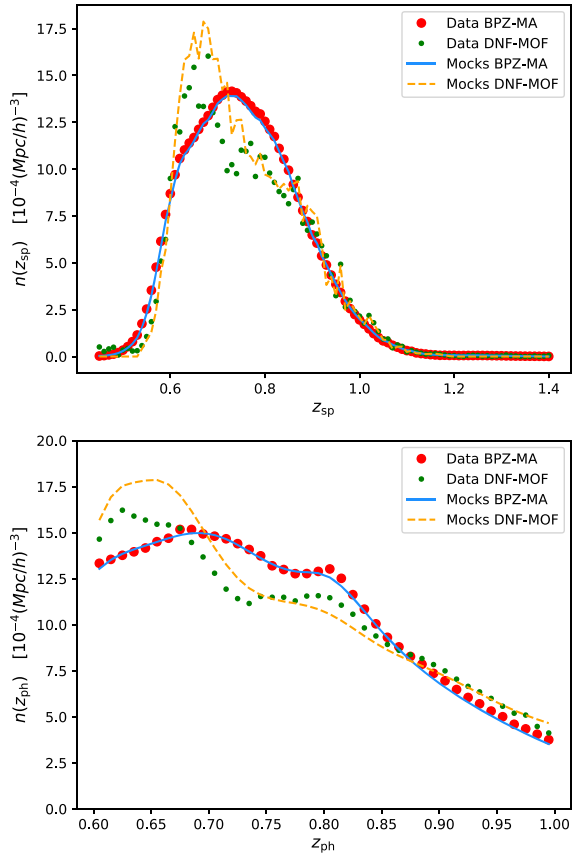

Figure 7. Number density of galaxies as a function of true redshift $z_{\rm sp}$ (Top) and photometric redshift $z_{\rm ph}$ (Bottom). We compare the results from the mocks (averaged over 1800 realizations) against their reference data, for the two different data sets. This is one of the key ingredients to ensure that the mocks will give us the correct covariance matrices.

with n_A being the number of objects per unit area. Note that ${\rm d}n_A(z_{\rm sp})/{\rm d}z$ contains the same information as the volume number density n, given that the cosmology is known.

At this stage, we can set mass thresholds $M_{\rm th}(z_{\rm sp})$ to our halo catalogues to obtain the abundance specified by the amplitude $A(z_{\rm sp})$ of the fits and equation (8). In Section 5, we will equivalently select galaxies by setting luminosity thresholds $l_{\rm th}(z_{\rm sp})$, but when looking at abundance and redshift distributions, this methodology yields the same results for haloes or galaxies. Once we fix those thresholds and apply them to our mock catalogues, we can also apply the redshift uncertainty by Monte Carlo sampling the $P(z_{\rm ph}|z_{\rm sp})$ distribution (from equation 7). This give us the $z_{\rm ph}$ for each halo/galaxy, and then we select the haloes/galaxies in the range $0.6 < z_{\rm ph} < 1.0$. Although the binning in $z_{\rm sp}$ was already small compared to the typical redshift uncertainties, we interpolate the value of $\mu(z_{\rm sp})$ to carry the information for $z_{\rm sp}$ beyond the precision of 0.01.

The resulting catalogues have an abundance of haloes/galaxies as shown in Fig. 7. The abundance as a function of true redshift $n(z_{\rm sp})$ (top panel) matches the data by construction, as we have set thresholds to force it to satisfy $n(z_{\rm sp})$ derived from equation (8). The small differences found in $n(z_{\rm sp})$, come simply from the uncertainty in the fit to equation (7). Note that the shape of $A(z_{\rm sp})$ and $n(z_{\rm sp})$ is similar, but still differ due to the cuts imposed in $z_{\rm ph}$. When analysing the abundance in $z_{\rm ph}$ space (bottom panel), we also find an overall good agreement, with some differences due to equation (7) not capturing completely the $\partial N/\partial z_{\rm ph}$ distributions

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

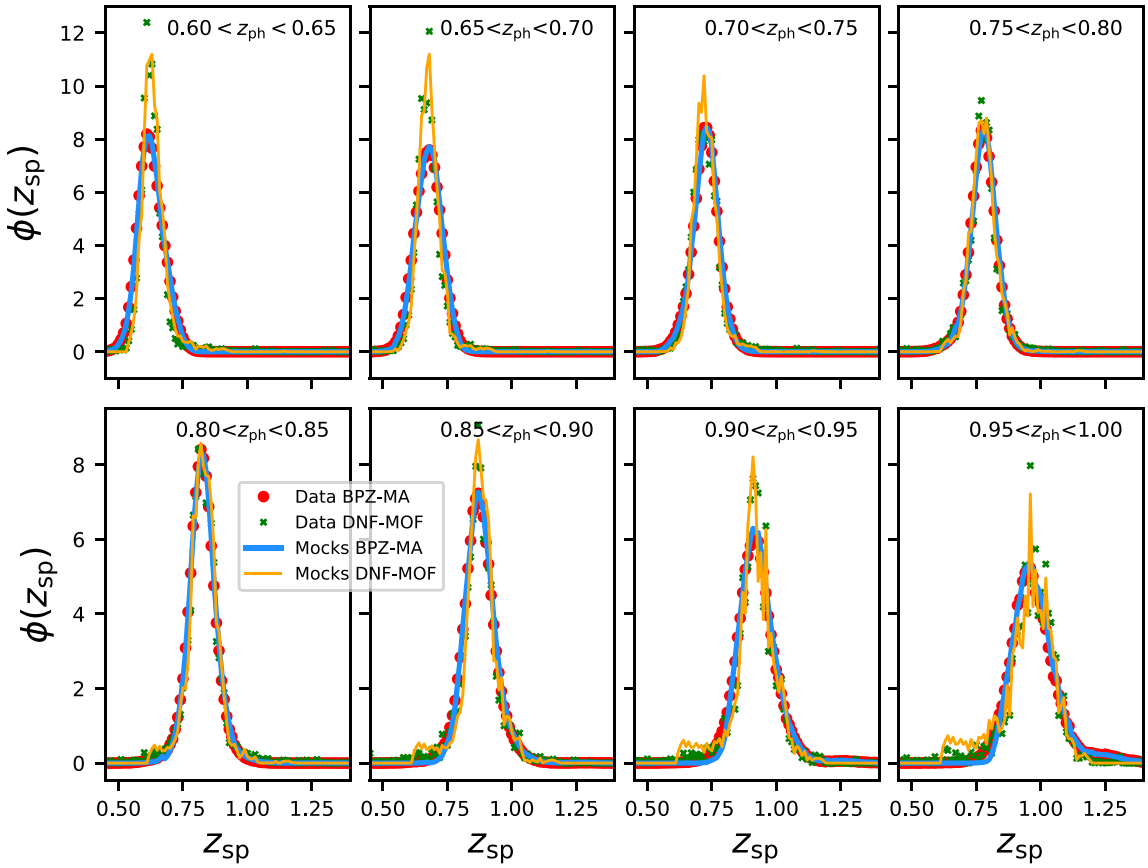

Figure 8. True redshift z_{sp} distribution in each of the eight z_{ph} bins in the interval $0.6 < z_{ph} < 1.0$ of data (points) versus mocks (line) for the two different data sets. Curves are normalized to have an integral of unity (equation 9). This is one of the key ingredients to ensure that the mocks will give us the correct covariance matrices.

from the data, this effect being more pronounced for the DNF-MOF data.

From the point of view of angular clustering analysis, the most relevant way of quantifying the photometric redshift uncertainty is to study the redshift distribution $\phi(z_{sp})$ in z_{ph} bins i, which we define as

$$\phi_i(z_{sp}) = \frac{1}{N_i^{tot}} \frac{dN_i}{dz_{sp}} \tag{9}$$

with N_i^{tot} being the total number of galaxies in the redshift bin so that the integral of ϕ equals unity.

We show in Fig. 8 the z_{sp} distribution for eight equally spaced z_{ph} bins for both the data and mock catalogues. We find that data and mocks approximately agree in their $\phi_i(z_{sp})$ for both data sets.

In DES-BAO-PHOTOZ, we study the $\phi_i(z_{sp})$ distributions of the DES Y1-BAO sample, by comparing the results estimated from z_{mc} against the distribution directly observed with spectroscopic data from the COSMOS field (and corrected for sample variance). We found that results from BPZ-MA were biased and underestimated the redshift uncertainties, whereas DNF-MOF was found to give accurate $\phi(z_{sp})$ distributions from the z_{mc} estimations. Hence, the BAO analysis in DES-BAO-MAIN uses the DNF-MOF data set for the main results. For that reason, in the following sections, where we model and analyse the clustering of the mock catalogues, we will only show results for DNF-MOF (although similar clustering fits were obtained for BPZ-MA).

In order to quantify the level of agreement in the redshift uncertainties modelling, we repeated the BAO analysis performed in DES-BAO-MAIN (with DNF-MOF), but assuming $\phi(z_{sp})$ from the mocks instead of the $\phi(z_{mc})$ from the data (the results for z_{mc} are denoted as 'z uncal' in table 5 of that paper). We obtain the same best-fitting value and uncertainty for the BAO scale and only $\Delta\chi^2/\mathrm{dof} = +1/43$, indicating that the accuracy achieved in $\phi(z_{sp})$ is excellent for the purposes of our analysis. We will also show at the end of next section (Fig. 11) the effect of the small differences in $\phi(z_{sp})$ in the amplitude of the clustering, finding them small compared to the error bars.

The functional form of the fits for $n(z_{ph}, z_{sp})$ from the data presented here had originally been optimized for the shape of the $n(z_{ph}, z_{sp})$ of BPZ-MA data. That should be taken into account when comparing the fits of both. BPZ-MA appears to give smoother distributions easier to fit with smooth curves (equation 7). However, given the poorer performance of the photo-z validation by BPZ-MA, we believe that the smoothness in the top panel of Fig. 5 is not realistic, but rather an oversimplification. Galaxy spectra contain features that translates into structure in the redshift distribution when redshifts are estimated with photometric redshift codes using broad-band filters. These features are well captured by DNF z_{mc}, as demonstrated by the lines in Fig. 5. These curves can also be approximately fitted with equation (7), but with some small structures on top of the smooth curve. We leave the modelling of these structures for a future study: the level of agreement in $\phi(z_{sp})$ (shown in Fig. 8) suggests that fitting any of the two data sets with the equation (7) is

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

a good approximation, and avoids the subtle problem of overfitting the data.

5 GALAXY CLUSTERING MODELLING

In Section 3, we described how to generate halo catalogues that sample the same volume as the DES Y1-BAO sample in the observational coordinates: {RA, Dec., $z_{\rm sp}$}. In Section 4, we described how we introduce the uncertainty in the estimation of redshift, leading to a catalogue that reproduces the abundance as a function of true and estimated redshift $\partial^2 n/\partial z_{\rm ph}/\partial z_{\rm sp}$ of the data. The next and last step, which is the focus of this section, is to produce a galaxy catalogue, able to reproduce the data clustering. For this, we will introduce hybrid HOD-HAM modelling.

Halo abundance matching (*HAM*). So far, all the clustering measurements shown throughout this paper were obtained from halo catalogues at a given mass threshold. But observed clustering is typically measured from galaxy catalogues with a magnitude-limited sample with the associated selection effects and, more generally, with redshift-dependent colour and magnitude cuts.

The basis of the HAM model is to assume that the most massive halo in a simulation would correspond to the most luminous galaxy in the observations and that we could do a one-to-one mapping in rank order. This is certainly very optimistic and realistic models need to add a scatter in the Luminosity–Mass relation (*L–M*) that will decrease the clustering for a magnitude-limited sample (Conroy, Wechsler & Kravtsov 2006; Behroozi, Conroy & Wechsler 2010; Trujillo-Gomez et al. 2011; Nuza et al. 2013; Guo et al. 2016).

HALOGEN was designed to only deal with main haloes, neglecting subhaloes. This limits the potential of HAM, as we cannot use its natural extension to subhaloes SHAM, where there is more freedom in the modelling by treating separately satellite and central galaxies (see e.g. Favole et al. 2015).

Nevertheless, substructure can be easily added to a main halo catalogue using an HOD technique.

Halo occupation distribution (*HOD*). We know that haloes can host more than one galaxy, especially massive haloes which match galaxy clusters. If we attribute a number of galaxies $N_{\rm gal}$ that is an increasing function of the halo mass ($M_{\rm h}$) to a halo mock catalogue, the clustering will be enhanced, since massive haloes will be over-represented (as occurs in reality for a magnitude-limited sample). This is the basis of the HOD methods (Jing, Mo & Börner 1998; Peacock & Smith 2000; Berlind & Weinberg 2002; Zheng et al. 2005; Skibba & Sheth 2009; Zehavi et al. 2011; Guo, Zehavi & Zheng 2012; Carretero et al. 2015; Rodríguez-Torres et al. 2016).

The details of the HOD and HAM need to be matched to observations via parameter fitting. This process can be particularly difficult if one aims at having a general model that serves for any sample with any magnitude and colour cut at any redshift (e.g. Carretero et al. 2015), with the added difficulty in our case that the redshift uncertainties would also vary with colour and magnitude. Additionally, the HOD implementation will determine the small-scale clustering corresponding to the correlation between galaxies of the same halo (Cooray & Sheth 2002). However, this is beyond the scope of this paper and we will only aim to match the large-scale clustering of the Y1-BAO sample.

In this paper, we combine the two processes: first we add substructure with an HOD model, and in a later step we select the galaxies that enter into our sample following an HAM prescription.

With regards to the HOD model we assign to each halo one central galaxy

$$N_{\rm cent} = 1\ ,\ , \tag{10}$$

and $N_{\rm sat}$ satellite galaxies given by a Poisson distribution with mean

$$\langle N_{\rm sat}\rangle = \frac{M_h}{M_1}\ ,\ , \tag{11}$$

where M_h is the mass of the halo, and M_1 is a free HOD parameter. Note that $N_{\rm cen}$ and $N_{\rm sat}$ above do not correspond to the final occupation distribution of the haloes, after applying the HAM (selecting only a subsample of these galaxies) in the later step.

Central galaxies are placed at the centre of the halo, whereas satellite galaxies are placed following an NFW (Navarro, Frenk & White 1996) profile. The concentration (a parameter of the NFW profile) is determined from the mass by the mass–concentration relation given in Klypin et al. (2016). The velocities of the central galaxies are taken from the host halo, whereas the velocities of the satellite have an additional dispersion:

$$v_{{\rm sat\ i}} = v_{{\rm halo\ i}} + \frac{1}{\sqrt{3}}\sigma_v(M_h)\cdot R^{\rm gauss}_{\mu=0\ \sigma=1}\ ,\ ,\qquad i = x, y, z\ ,\ , \tag{12}$$

where $R^{\rm gauss}_{\mu=0\ \sigma=1}$ is a random number drawn from a Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$, and σ_v is the dispersion expected from the virial theorem:

$$\sigma_v(M_h) = \sqrt{\frac{1}{5}\frac{GM_h}{R_{\rm vir}}}\ . \tag{13}$$

In this case we use $R_{\rm vir} = R_{200,\,\rm crit}$, since the mass definition used M_h is that included inside this radius.

Following the concepts of HAM, we assign a pseudo-luminosity l_p to the galaxies, which represents the entire set of selection criteria (in our case equation 1) compressed into a one-dimensional criterion. This l_p is not intended to represent a realistic luminosity. It is used to set the thresholds to match the abundance of the data in a simple way while accounting not only for the intrinsic Luminosity–Mass scatter, but also for the incompleteness of the sample. We model l_p (in arbitrary scales) with a Gaussian scatter around the halo mass M_h in logarithmic scales:

$$\log_{10}(l_p) = \log_{10}(M_h) + \Delta_{LM}\cdot R^{\rm gauss}_{\mu=0\ \sigma=1}, \tag{14}$$

where Δ_{LM} is a free parameter of the HAM model that controls the amount of scatter.

The abundance is then fixed by setting luminosity thresholds $l_p^{\rm th}(z_{\rm sp})$ that give us $n(z_{\rm sp})$ matching equation (8). Note that $l_p^{\rm th}(z_{\rm sp})$ is implicitly another HOD parameter, since it will depend on the other two parameters (M_1 and Δ_{LM}), but is not let free as it is defined by construction to match the abundance. Once these HOD–HAM steps are complete, we generate a photometric redshift for each galaxy as explained in Section 4.

In Fig. 9, we demonstrate the influence on clustering, under the assumption of fixed abundance and redshift distribution, of the two HOD parameters that we have introduced: M_1 and Δ_{LM}. We do so by studying the angular correlation function.

In order to create a galaxy catalogue identical to the halo catalogue, we would implement $M_1 = \infty$, $\Delta_{LM} = 0$. Deviations from those parameters control the clustering, as follows (Fig. 9):

(i) M_1: by lowering this parameter, we oversample the most massive haloes increasing the linear bias. It also introduces a one-halo term that fades away as we increase M_1.

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

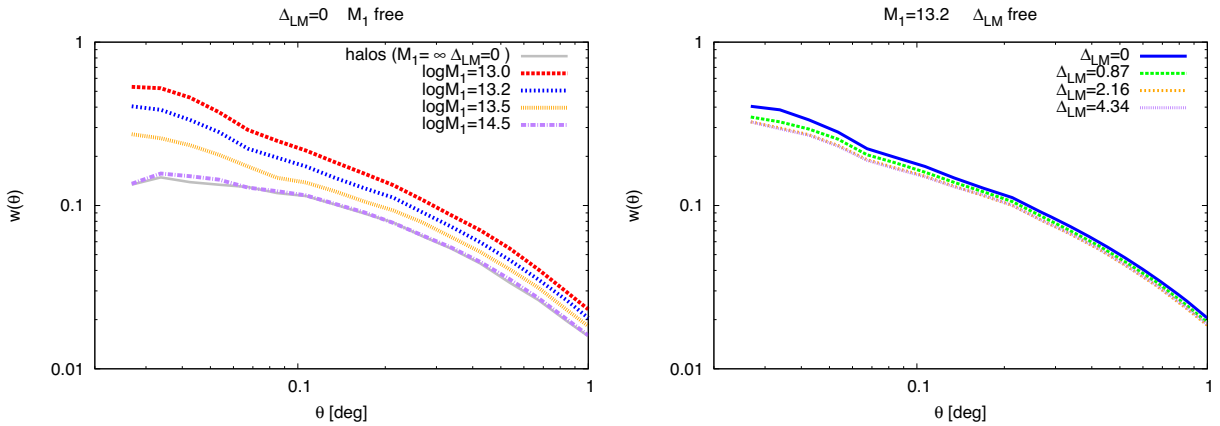

Figure 9. Effect on angular clustering in the redshift bin $0.6 < z_{\rm ph} < 0.65$ of the HOD parameters at fixed abundance and redshift distribution $n(z_{\rm sp}, z_{\rm ph})$. Left: effect produced by M_1. Right: effect produced by $\Delta_{\rm LM}$.

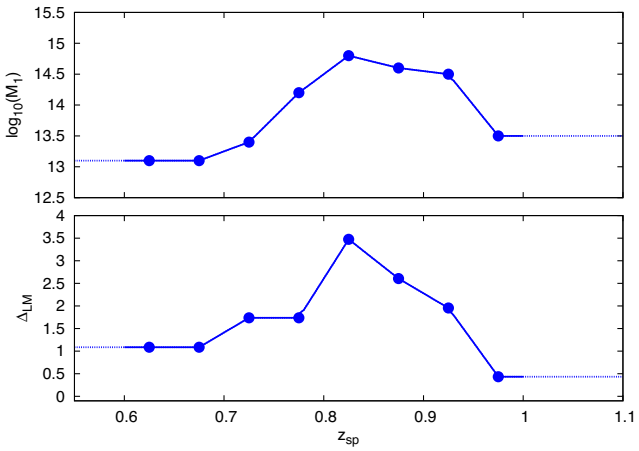

Figure 10. HOD parameter evolution with redshift $z_{\rm sp}$. Top: evolution of $\log M_1$ (in $\mathrm{M}_\odot\, h^{-1}$). Bottom: evolution of Δ_{LM}. The evolution of the parameters is assumed smoothed, and they are linearly interpolating between the eight pivots: the solid circles. Although the focus of the modelling and analysis in this paper is in the range $0.6 < z < 1$, there are contributions from galaxies with $z_{\rm sp}$ ranging from 0.45 to 1.4. A flat evolution of the HOD parameters is assumed before and after the first and last pivot, respectively.

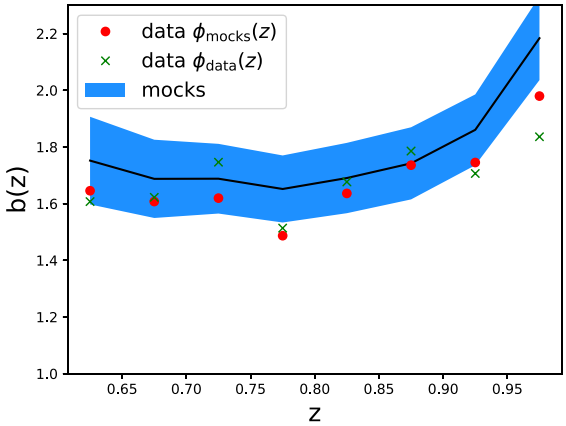

Figure 11. Bias evolution with redshift as measured from the mocks (shaded region representing the mean and standard deviation) and data. For all cases, we fit the bias for the angular correlation in $z_{\rm ph}$ bins of width $\Delta z = 0.05$. For the mocks, we assume their own $\phi(z_{\rm sp})$ distribution when fitting b, whereas for the data we test the difference between assuming $\phi_{\rm mocks}(z_{\rm sp})$ or $\phi_{\rm data}(z_{\rm sp})$. At this stage, we only consider the diagonal part of the covariance.

(ii) Δ_{LM}: as we increase this value, lower mass haloes enter into our selection and higher mass haloes escape it, lowering the bias.

We can find in the literature more complex $N_{\rm sat}(M_h)$ functions but to avoid huge degeneracies in the parameters, we choose to only have two free parameters M_1 and Δ_{LM}. We do let them evolve with redshift in order to adapt for the clustering of the data, including selection effects.

We impose a smooth evolution of these parameters with redshift by making linear interpolations between eight pivots at the centre of each redshift bin, and making the evolution flat beyond the centre of the first and last redshift bin, since our measurements (in eight $z_{\rm ph}$ bins in $0.6 < z_{\rm ph} < 1.0$) will not be able to constrain the parameter evolution beyond these points. With these constraints, we find that an HOD parameter evolution as shown in Fig. 10, gives us a good match to the evolution of the amplitude of the clustering of the data. We have only explored large scales, since we focus on BAO physics. This leaves some degeneracy between the two HOD parameters, which could be broken by including the small scales. We leave these studies for the next internal data release, which will include data from the first three years of the survey.

Once the HOD parameters are fixed, we measure the bias of the mock catalogues by fitting the angular correlation to theoretical predictions (see how we compute the theoretical $w(\theta)$ in Section 6.2). We fit correlations at angles that correspond to $20 < r < 60h^{-1}$ Mpc scales, different for each redshift bin. These measurements are meant to be a simple verification of the fit on the amplitude of the clustering, and hence we assume a simple diagonal covariance, measured from the mocks (unlike we will do in Section 6.2). In Fig. 11, we show the bias evolution recovered from the mock catalogues, with the blue band representing the 1σ region computed as the standard deviation of best fit of each mock.

We also present in Fig. 11 the bias measured from the data following the same procedure. In red, we show the measured bias if we assume the same redshift distribution $\phi_i(z_{\rm sp})$ as in the mocks, whereas in green we show the measured bias assuming the $\phi_i(z_{\rm sp})$ estimated from the data itself. The main conclusion of this plot is that the amplitude of the clustering of the data (red circles) agrees within 1σ with the mocks. We notice a trend of the mocks hav-

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

ing a slightly higher amplitude than the data, this effect is small compared to the uncertainties, but will be studied in more detail for future releases. This plot also shows us that the small differences between $\phi_{\rm mocks}$ and $\phi_{\rm data}$ in Fig. 8 do not have a significant effect in the clustering, since all shifts are smaller than the σ (and most of them being much smaller).

Looking at the mean halo mass of the catalogues we find that our Y1-BAO sample probes halo masses between $1.3 \times 10^{13}\,{\rm M}_\odot\,h^{-1}$ and $2.1 \times 10^{13}\,{\rm M}_\odot\,h^{-1}$ depending on the redshift range.

6 CLUSTERING RESULTS

In previous sections, we designed a method to reproduce all the relevant properties of the Y1-BAO sample for clustering analysis. In this section, we fix all the modelling and parameters and run $N_{\rm mocks} = 1800$ realizations with different initial conditions. We analyse the clustering of all those mocks using different estimators in different spaces ($w_i(\theta)$, $\xi(r)$, C_l)), and their covariance matrices. We also compare those statistics with theoretical models, and measurements from the data.

6.1 Theoretical modelling

In this section, we compare the clustering of the mock catalogues to several theoretical predictions. The baseline model we use is the linear theory with non-linear BAO damping, a Kaiser factor (Kaiser 1987) and linear bias:

$$P(k,\mu) = (b+\mu^2 f)^2\left[\left(P_{\rm lin}(k) - P_{\rm nw}(k)\right)e^{-k^2\Sigma_{\rm NL}^2} + P_{\rm nw}(k)\right]. \quad (15)$$

$P_{\rm lin}(k)$ is the linear power spectrum and $P_{\rm nw}(k)$ is a smoothed non-BAO wiggle version of it obtained with the fitting formulas from Eisenstein & Hu (1998). We use the same linear power spectrum from MICE cosmology that we used for the 2LPT. The parameter $\Sigma_{\rm NL}$ represents the damping scale of the BAO due to non-linear evolution. We will discuss in Sections 6.2 a nd6.3 its best-fitting value.

The details of the theoretical framework can be found in (DES-BAO-θ-METHOD). But, going from $P(k,\,\mu)$ to $\xi(r,\,\mu)$ can be summarized in three steps. First, decomposing the power spectrum in multipoles by convolving it with the Legendre Polynomial L_l:

$$P_l(k) = \frac{2l+1}{2}\int_{-1}^{1} {\rm d}\mu\, P(k,\mu)L_l(\mu). \quad (16)$$

Secondly, Fourier transforming the multipoles using the Spherical Bessel functions j_l:

$$\xi_l(s) = \frac{i^l}{2\pi^2}\int {\rm d}k\, k^2 P_l(k) j_l(ks). \quad (17)$$

And, finally, from the ξ_l multipoles we can recover the full anisotropic correlation function:

$$\xi(s,\mu) = \sum_l \xi_l(s)L_l(\mu). \quad (18)$$

In the following sections, we will see how to project $\xi(s,\,\mu)$ to obtain different estimators.

6.2 Angular clustering: $w\,(\theta)$

Here, we study the angular correlation function of the final mock catalogues. We divide our catalogues in four redshift bins as we do in DES-BAO-MAIN, and compute the correlation functions using the method given in Section 2.3. We compare in Fig. 12 the mean correlation function $\bar{w}(\theta)$ of the mock catalogues to the correlation from the data. The error bars attached to the data were computed as the square root of the diagonal of the covariance matrix, obtained from the mock catalogues with:

$$C_{ij} = \frac{1}{(N_{\rm mocks}-1)}\sum_{k=0}^{N_{\rm mocks}}[\bar{w}(\theta_i) - w_k(\theta_i)]\cdot[\bar{w}(\theta_j) - w_k(\theta_j)]. \quad (19)$$

We compute the χ^2 of the data with respect to the mean of the mocks using the covariance from the mocks using

$$\chi^2 = \sum_i\sum_j[\bar{w}(\theta_i) - w_{\rm data}(\theta_i)]C^{-1}{}_{ij}[\bar{w}(\theta_j) - w_{\rm data}(\theta_j)], \quad (20)$$

and we display in Fig. 12 the goodness of the fit $\chi^2_{\rm red} = \chi^2/d.o.f$ between the mocks to the data for each $z_{\rm ph}$ bin. We find values near 1, indicating a good fit to the data (with p-values of 0.430, 0.819, 0.375, and 0.147). Note the strong correlations between θ bins that move coherently up or down from realization to realization. This makes less intuitive the visual comparison (if one naively assumes a diagonal covariance matrix) between the curves to estimate the goodness of the fit: for example, the second bin has the best χ^2, even if the data points appear to be systematically below the mocks line. We also bear in mind the cosmology from the mocks is not compatible with current cosmological constraints (e.g. Planck Collaboration XIII 2016), and this could introduce an extra χ^2 contribution. However, the main difference due to cosmology would be the BAO position, and for the level of uncertainty that we have in a single $z_{\rm ph}$ bin (as opposed to combining the four, as we do in DES-BAO-MAIN), this contribution is expected to be negligible.

Typically, we would need to correct the χ^2 values by the factor in Hartlap, Simon & Schneider (2007), due to noise in the inverse of the covariance matrix caused by having a finite number of mocks. We do include those factors, but find a negligible effect. This is because, for the results presented here, we do not consider the covariance between different $z_{\rm ph}$ bins, and, hence, our data vector is small compared to the large number of simulations used.

In Fig. 12, we also show theoretical angular correlation functions fitted to the mocks. The theoretical $w(\theta)$ can be computed by projecting $\xi(s,\,\mu)$ from equation (18), and taking into account the $z_{\rm sp}$-distributions within the bins (equation 9):

$$w(\theta) = \int {\rm d}z_{\rm sp}\phi(z_{\rm sp})\int {\rm d}z'_{\rm sp}\phi(z'_{\rm sp})\xi(r(z_{\rm sp}, z'_{\rm sp}, \theta), \mu(z_{\rm sp}, z'_{\rm sp}, \theta)). \quad (21)$$

We denote $r(z_{\rm sp}, z'_{\rm sp}, \theta)$ as the comoving distance between two galaxies, respectively, at redshift z and z', and separated by a projected angle on the sky of θ, and $\mu(z_{\rm sp}, z'_{\rm sp}, \theta)$ the cosine of the orientation of the pair of galaxies with respect to the line of sight.

In Section 6.1, we left two parameters free: b and $\Sigma_{\rm NL}$. The choice of the damping factor is discussed in DES-BAO-θ-METHOD: we use a constant $\Sigma_{\rm NL} = 5.2\,h^{-1}$ Mpc, which fits well the correlation functions of the mocks and it is within the theoretical expectations (Seo & Eisenstein 2007). The linear bias b is then fit for each redshift bin to the mean of the mocks, obtaining the values written in Fig. 12. We overplot the theoretical $w(\theta)$ that best fits the mock catalogues, this best fit should not be confused with the best fit to the data presented in DES-BAO-MAIN.

In Fig. 13, we study the form of the covariance matrix obtained from the mock catalogues (equation 19) and compare it to the theoretical model discussed in Crocce, Cabré & Gaztañaga (2011) and DES-BAO-θ-METHOD. For the mock covariance, we used two sets of mocks in which the only difference is the mask we apply

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

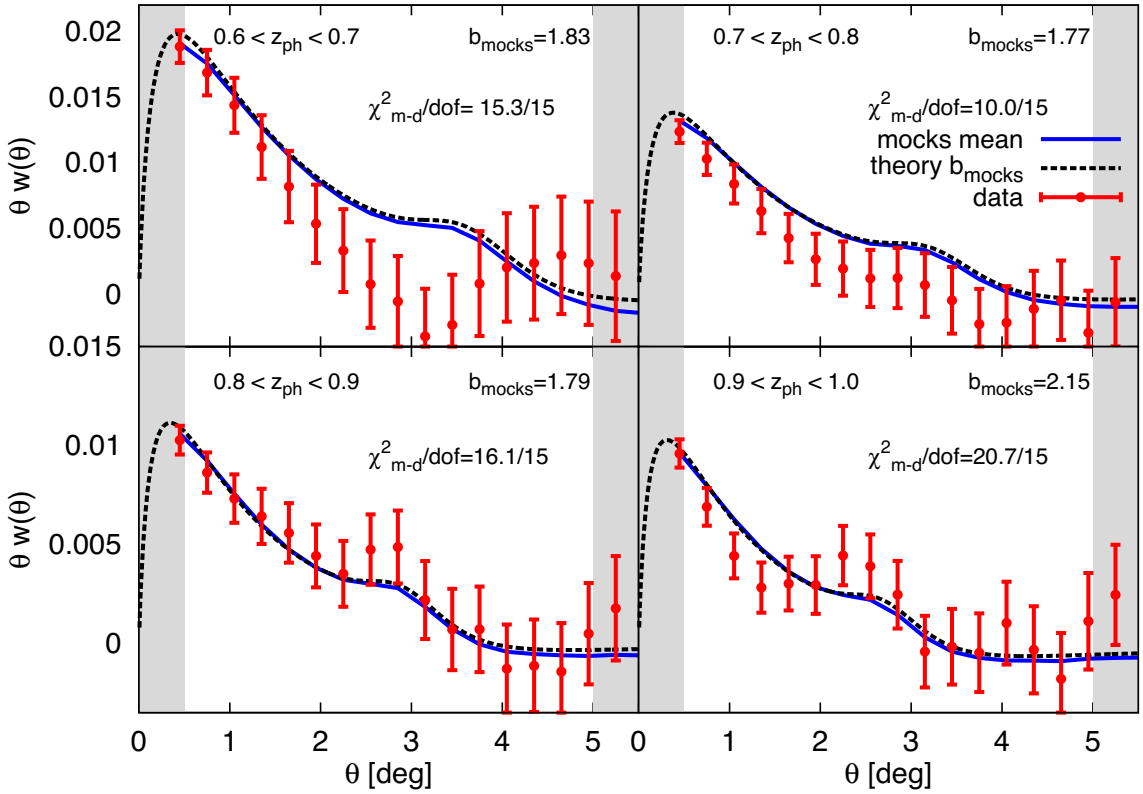

Figure 12. Angular clustering of mocks against data in four photometric redshift bins with width $\Delta z = 0.1$. The red points with error bars represent the Y1-DES DNF-MOF data with the error bar computed as the standard deviation of the 1800 mocks. The mean over the mock catalogues is represented by the blue solid line. We also show the theoretical curve that best fits the mocks. We show the χ^2 of the fit between the mocks and the data given by equation (20) and the degrees of freedom (dof, corresponding to the number of θ bins). Note that neighbouring θ bins are highly correlated. These χ^2 show a good level of agreement in the clustering, this is one of the key ingredients to ensure that the mocks will give us the correct covariance matrices.

to obtain eight mocks from the full sky catalogue. The one we denote as 'Y1', depicted in Fig. 3, has the same footprint as the data and it is the standard one used for all the analysis throughout this paper. The 'Square' mocks[5] have a continuous mask with $\Delta\mathrm{RA} = \Delta\sin(\mathrm{Dec.}) = \sqrt{\Omega_{\mathrm{sur}}}$.

We find that the theoretical predictions for the diagonal terms agree very well with the results from the 'Square' mask mocks except for the lowest scales ($\theta < 1^\circ$). When applying the 'Y1' mask we find that a tilt is introduced in the diagonal of the covariance matrix $\sigma(\theta)$. Studying the non-diagonal covariance, we find again a good agreement between the 'Square' mocks and the theory. In this case, the effect of introducing the 'Y1' mask is to modify significantly the covariance at small scales. A more detailed analysis on the covariance matrices and their effects for BAO is discussed in DES-BAO-θ-METHOD.

We conclude that the theoretical method agrees very well with the covariances from the mock catalogues if the geometry of the survey is ignored. But also that the geometry of the survey has a significant contribution to the covariance.

6.3 3D clustering: $\xi_{\mu<0.8}(s_\perp)$

In photometric surveys most of the radial BAO information is lost due to redshift uncertainties. However, we showed in DES-BAO-$s_\perp$-METHOD that also due to redshift uncertainties the angular BAO information is spread to apparent $\mu > 0$ modes. More specifically, most of the BAO information is spread nearly homogeneously in the range $0 < \mu < 0.8$. Hence, a nearly optimal way to study the BAO is analysing the following estimator (DES-BAO-$s_\perp$-METHOD):

$$\xi_{\mu<0.8}(s_\perp) \equiv \frac{1}{0.8}\int_0^{0.8} \mathrm{d}\mu\, \xi\left(s_\perp,\ s_\parallel(s_\perp,\mu)\right), \tag{22}$$

being $s_\parallel$ ($s_\perp$) the projection of s along (perpendicular to) the line of sight and

$$s_\parallel(s_\perp,\mu) = s_\perp\sqrt{\frac{\mu^2}{1-\mu^2}}. \tag{23}$$

Equivalently to Feldman, Kaiser & Peacock (1994), we apply inverse-covariance weighting to the galaxies to optimize the signal to noise of our measurements:

$$w_{\mathrm{FKP}}(z) = \frac{b(z)D(z)}{1 + n_{\mathrm{eff}}(z)P_{\mathrm{lin}}(k_{\mathrm{eff}}, z=0)b^2(z)D^2(z)}, \tag{24}$$

where $n_{\mathrm{eff}}(z)$ is the effective number density accounting for the redshift uncertainty (see equation15 in Ross et al. 2017b), and $k_{\mathrm{eff}} \approx 0.12h\mathrm{Mpc}^{-1}$ the scale where the BAO information is effectively coming from.

We compute the 3D correlation function $\xi(s_\perp, s_\parallel)$ of the data and mocks from pair-counts as explained in Section 2.3, and integrating it using equation (22). The results are shown in Fig. 14, finding again a good agreement between mocks and data (with a p-value of 0.20). Again, the error bars attached to the data come from the diagonal

[5]Note that these square mocks are not the same as the ones in DES-BAO-$s_\perp$-METHOD, which had a simpler modelling.

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

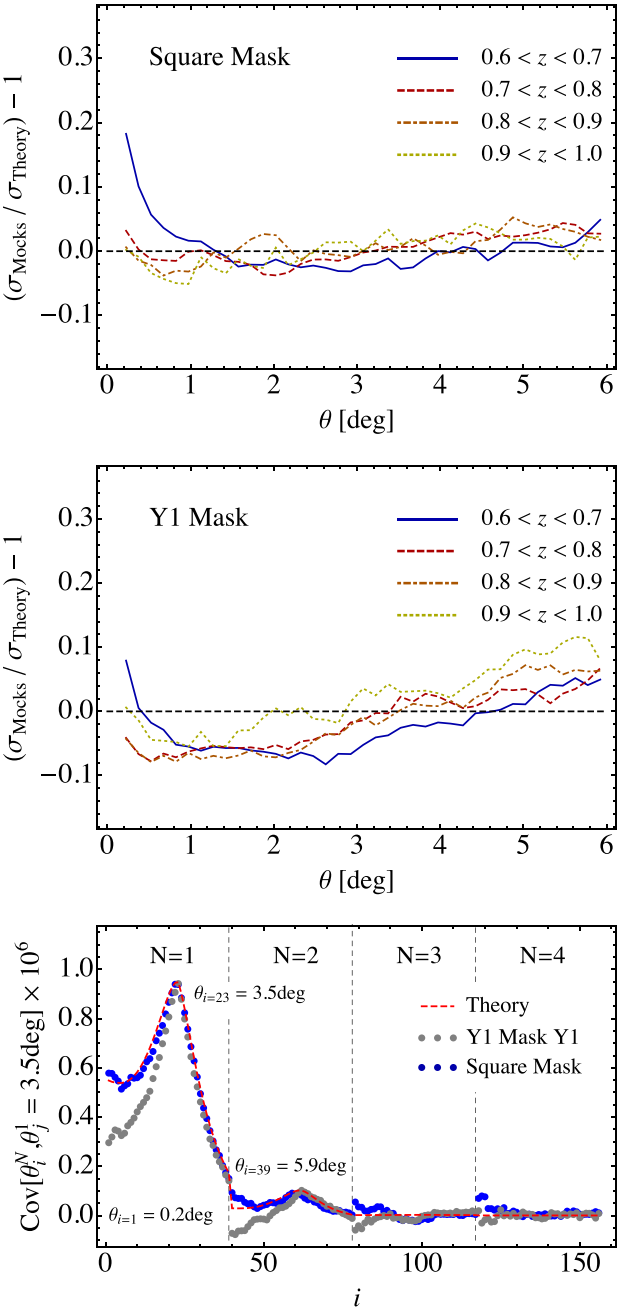

Figure 13. Covariance Matrices of the angular correlation function. Top: diagonal of the covariance matrices, theory versus 'Square' mask mocks. Middle: diagonal of the covariance matrices, theory versus 'Y1' mask mocks. Bottom: column of the covariance matrix $C[\theta_i^N, \theta_j^M]$ with fixed $\theta_j^1 = 3.5$ deg as a function of θ_i^N. The i, j subscripts represent the index of the data vector (from 0 to 159, with θ ranging in [0, 6] deg in $\Delta\theta = 0.15$ deg steps and 4 redshift bins), and the N, M superscripts represent the index of the redshift bin. We compare theoretical predictions against Y1 mocks with 'Square' and 'Y1' masks. In both the diagonal and off-diagonal cases, the theoretical prediction lies closer to the 'Square' mask mocks covariance.

of the covariance from the mocks, and the χ^2 were computed with the covariance from the mocks.

From the theoretical point of view, we can project $\xi(s, \mu)$ from equation (18) using equation (22). For this we use a Gaussian uncertainty approximation, following the procedure in DES-BAO-$s_\perp$-METHOD. In this case we re-fit $\Sigma_{\rm NL}$, expecting this to capture part of the effect of the non-Gaussian tails of the redshift uncertainties. We find a best-fitting value of $\Sigma_{\rm NL} = 8 M{\rm pc}\, h^{-1}$. We compute the bias of the mocks $b_{\rm mocks}$ as the weighted average of the values obtained in the previous subsection (Section 6.2). We plot the theoretical prediction for this bias, finding a good agreement that confirms the consistency of the mocks and theoretical framework we are working on.

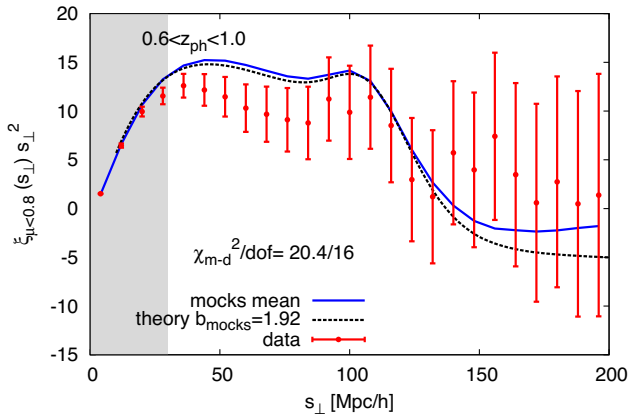

Figure 14. 3D correlation function using the estimator in equation (22). The mean of the mocks is shown as a solid line, whereas the data is represented as points attached error bars coming from the diagonal of the mock covariance.

6.4 Clustering in angular harmonic space: C_l

In spectroscopic surveys, it has been shown that even though theoretically $\xi(r)$ and $P(k)$ carry the same information in different spaces, in reality, when dealing with finite volume, they can provide complementary information (Ata et al. 2018). In photometric surveys, one usually deals with projected quantities. In this case, the complementary observables are the angular correlation $w(\theta)$ and its equivalent in harmonic space, the angular power spectrum C_ℓ.

The galaxy number density contrast in a given redshift bin $\delta_{{\rm gal},i}(\hat{\boldsymbol{n}})$ can be decomposed into spherical harmonics $Y_{\ell m}$ as

$$\delta_{{\rm gal},i}(\hat{\boldsymbol{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m,i} Y_{\ell m}(\hat{\boldsymbol{n}}) \,, \,, \tag{25}$$

where $a_{\ell m}$ are the harmonic coefficients. The angular power spectrum $C_{\ell,\,i}$ is then defined via

$$\langle a_{\ell m,i} a^*_{\ell' m',i} \rangle \equiv \delta_{\ell\ell'} \delta_{mm'} C_{\ell,i} \,. \tag{26}$$

For data collected over the whole sky, an unbiased estimator of the angular power spectrum is simply the average of the $a_{\ell m}$ coefficients over all m values:

$$\hat{C}_\ell = \frac{1}{2\ell+1} \sum_{m=-\ell}^{m=\ell} |a_{\ell m}|^2 \,. \tag{27}$$

When performing full-sky estimations, we compute the coefficients $a_{\ell m}$ from the pixelized density contrast maps using the `anafast` routine within HEALPIX.

In the case of partial sky coverage, the pseudo-C_ℓ method (Hivon et al. 2002) is used to measure the C_ℓs. The measurement is performed with a binning of $\Delta\ell = 20$ using a resolution of $N_{\rm side} = 1024$. A more detailed description of the methodology is found in DES-BAO-ℓ-METHOD.

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

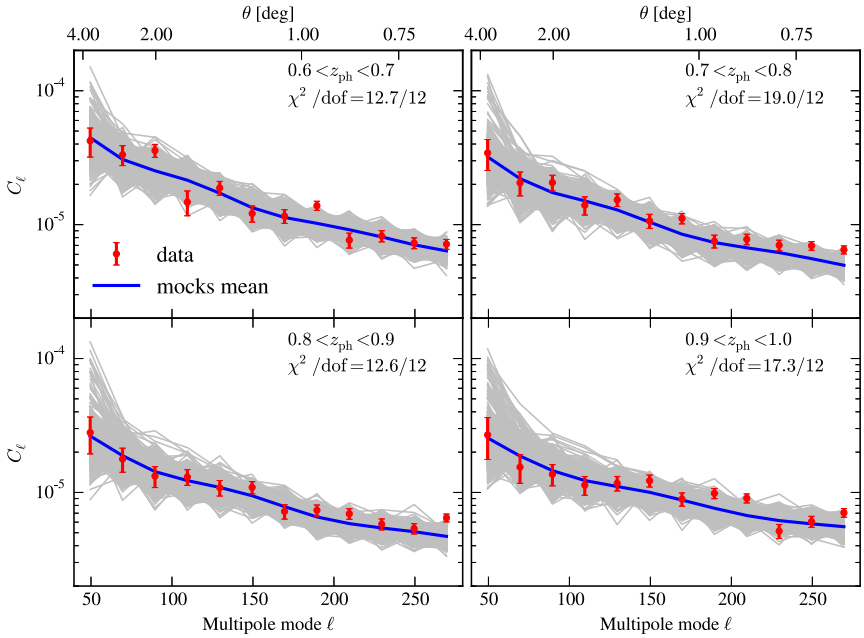

Figure 15. Angular power spectrum for four $z_{\rm ph}$ bins for the 1800 mocks. In blue solid lines, we show the mean of the mock catalogues, whereas the data are shown as points with error bars coming from the mock covariance. The results from all the individual realization of the mocks are also shown in grey (forming nearly a shaded band).

Following the procedure of previous subsections, we compute the average and the covariance matrix for the angular power spectra in each redshift bin using the 1800 mocks. The measurement is also performed on Y1 data, with error bars estimated from the covariance matrix. The results in Fig. 15 show that the angular power spectrum measured from the mocks are consistent with the measurements from data. For simplicity, we do not include the theoretical prediction here, but refer to DES-BAO-ℓ-METHOD, where we confirm that the bias values measured in Section 6.2 fit well the C_ℓ from the mocks.

7 SUMMARY AND CONCLUSIONS

In this paper, we have designed and analysed a set of 1800 mock catalogues able to reproduce statistically the properties of the Dark Energy Survey Year-1 BAO sample.

The three main properties reproduced are as follows:

(i) Sampled observational volume: $\{$RA, Dec., $z_{\rm ph}\}$ (Fig. 3).
(ii) Abundance of galaxies, redshift distribution, and redshift uncertainty: $n(z_{\rm ph}, z_{\rm sp})$ (Figs 7 and 8).
(iii) Clustering as a function of redshift: $b(z_{\rm ph})$, $w_i(\theta)$, $\xi_{\mu<0.8}(s_\perp)$ and C_l^i (Figs 11, 12, 14, and15).

Matching the properties listed above guarantees that our mock catalogues can correctly reproduce the clustering covariance of the data.

For the first time, we have presented a set of mock catalogues capable to reproduce simultaneously the clustering of a photometric sample together with an accurate description of redshift distribution and uncertainties.

Throughout the paper, we described in detail the way we design the galaxy catalogues, dividing the sections by the different processes/effects modelled (following the three bullet points above) and, in each section, building upon the modelling fixed in the sections before.

Once all the parameters and models have been fixed, catalogues are created following sequentially these steps:

(i) Generate halo catalogues with HALOGEN at fixed redshifts in cubic volumes (Section 3.1).
(ii) Compose a full sky light-cone by superposing snapshots in redshift shells and using periodic conditions (Section 3.2).
(iii) Add one central galaxy and $N_{\rm sat}$ satellite galaxies to each halo using equation (11) (Section 5).
(iv) Assign a luminosity l_p to each galaxy, based on the halo mass, using equation (14). Then, select only the galaxies with $l_p > l_{\rm th}(z_{\rm sp})$ (Section 5).
(v) Draw a photometric redshift $z_{\rm ph}$ for each galaxy from the distribution $P(z_{\rm ph}|z_{\rm sp})$ in equation (7). Then, select only galaxies in the range of interest $0.6 < z_{\rm ph} < 1$ (Section 4).
(vi) Apply the survey mask (Section 3.3).

Some of the steps listed above involved parameters that had to be adjusted to the data or to other simulations before starting producing the final batch of mock catalogues. The main fitted ingredients of the modelling are as follows:

(i) HALOGEN parameters $(\alpha(M_h), f_{\rm vel}(M_h))$ were tuned to reproduce the halo clustering and velocity distributions of the reference N-Body simulation MICE as a function of mass and redshift. The input HMF was also partially tuned to the simulation.
(ii) We modelled $n(z_{\rm ph}, z_{\rm sp})$ by fitting a *double-skewed* Gaussian to $\left.\frac{\partial n}{\partial z_{\rm ph}}\right|_{z_{\rm sp}}$ (equation 7) from the data for all $z_{\rm sp}$.
(iii) We explored the HOD parameter space $\{M_1(z_{\rm sp}), \Delta_{LM}(z_{\rm sp})\}$ to reproduce the bias evolution $b(z_{\rm sp})$ from the data. Simultaneously, the luminosity thresholds $l_{\rm th}(z_{\rm sp})$ need to be readjusted for each run in order to match the amplitude of $\left.\frac{\partial n}{\partial z_{\rm ph}}\right|_{z_{\rm sp}}$.

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

This work is particularly relevant for the first BAO measurement with DES data presented in DES-BAO-MAIN, where these mocks have been employed to run robustness tests, optimize the methodology, compute covariance matrices, and from them compute the best fit and uncertainty for angular scale of the BAO angular distance measurement D_A. It has also played key roles in other companion papers, e.g.: covariance for bias fits (DES-BAO-SAMPLE), study of methodology (DES-BAO-θ-METHOD, DES-BAO-$s_\perp$-METHOD, DES-BAO-ℓ-METHOD), and photo-z validation (DES-BAO-PHOTOZ). This will set up a framework for the DES mock catalogue designs of the coming data releases, and can also valuable for many other photometric surveys studying the LSS of the Universe.

ACKNOWLEDGEMENTS

SA and WJP acknowledge support from the UK Space Agency through grant ST/K00283X/1, and WJP acknowledges support from the European Research Council through grant *Darksurvey*, and the UK Science & Technology Facilities Council through the consolidated grant ST/K0090X/1. SA thanks IFT computing facilities and support that were necessary to run the simulations and fits. SA thanks Alexander Knebe for the support during his PhD, when this project was started. NB acknowledges the use of University of Florida's supercomputer HiPerGator 2.0 as well as thanks the University of Florida's Research Computing staff. HC is supported by CNPq. ML is partially supported by FAPESP and CNPq. NK was financed by a FAPESP fellowship. We thank the support of the Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015-71825, ESP2015-66861, FPA2015-68048, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA programme of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through project number CE110001020.

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

REFERENCES

Alonso D., 2015, Astrophysics Source Code Library, record ascl:1506.016
Ata M. et al., 2018, MNRAS, 473, 4773
Avila S., Murray S. G., Knebe A., Power C., Robotham A. S. G., Garcia-Bellido J., 2015, MNRAS, 450, 1856
Bautista J. E. et al., 2017, preprint (arXiv:1712.08064)
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Benítez N., 2000, ApJ, 536, 571
Benítez N. et al., 2004, ApJS, 150, 1
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Bertin E. Arnouts S., 1996, A&AS, 117, 393
Beutler F. et al., 2011, MNRAS, 416, 3017
Blake C. et al., 2011, MNRAS, 418, 1707
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bouchet F. R., Colombi S., Hivon E., Juszkiewicz R., 1995, A&A, 296, 575
Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, MNRAS, 447, 646
Chan K. C. et al., 2018, preprint (arXiv:1801.04390)
Chuang C.-H., Kitaura F.-S., Prada F., Zhao C., Yepes G., 2015a, MNRAS, 446, 2621
Chuang C.-H. et al., 2015b, MNRAS, 452, 686
Cole S. et al., 2005, MNRAS, 362, 505
Coles P., Jones B., 1991, MNRAS, 248, 1
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS, 414, 329
Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, MNRAS, 453, 1513
Crocce M. et al., 2017, preprint (arXiv:1712.06211)
De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, MNRAS, 459, 3078

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

DES Collaboration, 2017, preprint (arXiv:1712.06209)
Diehl H. T. et al., 2014, in Peck A. B., Benn C. R., Seaman R.L., eds, Proc. SPIE Conf. Ser., Vol. 9149, Observatory Operations: Strategies, Processes, and Systems V. SPIE, Bellingham, p. 91490V
Dodelson S., Schneider M. D., 2013, Phys. Rev. D, 88, 063537
Drlica-Wagner A. et al., 2018, ApJS, 235, 33
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eisenstein D. J. et al., 2005, ApJ, 633, 560
Favole G. et al., 2015, MNRAS, 461, 3421
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Fosalba P., Gaztañaga E., Castander F. J., Manera M., 2008, MNRAS, 391, 435
Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015, MNRAS, 448, 2987
Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015, MNRAS, 447, 1319
Friedrich O., Eifler T., 2018, MNRAS, 473, 4150
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Guo H., Zehavi I., Zheng Z., 2012, ApJ, 756, 127
Guo H. et al., 2016, ApJ, 831, 3G
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Hinshaw G. et al., 2007, ApJS, 170, 288
Hivon E., Gorski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2
Hoyle B. et al., 2018, MNRAS, 478, 592
Jing Y. P., Mo H. J., Börner G., 1998, ApJ, 494, 1
Kaiser N., 1984, ApJ, 284, L9
Kaiser N., 1987, MNRAS, 227, 1
Kitaura F.-S. et al., 2016, MNRAS, 456, 4156
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Manera M. et al., 2013, MNRAS, 428, 1036
Monaco P., 2016, Galaxies, 4, 53
Monaco P., Sefusatti E., Borgani S., Crocce M., Fosalba P., Sheth R. K., Theuns T., 2013, MNRAS, 433, 2389
Moutarde F., Alimi J.-M., Bouchet F. R., Pellat R., Ramani A., 1991, ApJ, 382, 377
Murray S. G., Power C., Robotham A. S. G., 2013, Astron. Comput., 3, 23
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nuza S. E. et al., 2013, MNRAS, 432, 743
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Peebles P. J. E., Yu J. T., 1970, ApJ, 162, 815
Planck Collaboration XIII, 2016, A&A, 594, A13
Pope A. C., Szapudi I., 2008, MNRAS, 389, 766
Press W. H., Schechter P., 1974, ApJ, 187, 425
Rodríguez-Torres S. A. et al., 2016, MNRAS, 460, 1173
Ross A. J. et al., 2017a, MNRAS, 464, 1168
Ross A. J. et al., 2017b, MNRAS, 472, 4456
Rozo E. et al., 2016, MNRAS, 461, 1431
Schneider M. D., Cole S., Frenk C. S., Szapudi I., 2011, ApJ, 737, 11
Scoccimarro R., 2000, ApJ, 544, 597
Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629
Seo H.-J., Eisenstein D. J., 2007, ApJ, 665, 14
Skibba R. A., Sheth R. K., 2009, MNRAS, 392, 1080
Sunyaev R. A., Zeldovich Y. B., 1970, Ap&SS, 7, 3
Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, J. Cosmol. Astropart. Phys., 6, 36
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16
Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230
White M., Tinker J. L., McBride C. K., 2014, MNRAS, 437, 2594
Zehavi I. et al., 2011, ApJ, 736, 59
Zel'dovich Y. B., 1970, A&A, 5, 84
Zheng Z. et al., 2005, ApJ, 633, 791

[1]*Institute of Cosmology and Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth PO1 3FX, UK*
[2]*Departamento de Física Teórica, Módulo C-15, Facultad de Ciencias, Universidad Autónoma de Madrid, E-28049 Cantoblanco, Madrid, Spain*
[3]*Instituto de Física Teórica, UAM-CSIC, Universidad Autonoma de Madrid, E-28049 Cantoblanco, Madrid, Spain*
[4]*Institut de Ciències de l'Espai, IEEC-CSIC, Campus UAB, Facultat de Ciències, Torre C5 par-2, E-08193 Barcelona, Spain*
[5]*Center for Cosmology and AstroParticle Physics, The Ohio State University, Columbus, OH 43210, USA*
[6]*Department of Physics, University of Florida, Gainesville, FL 32611, USA*
[7]*Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*
[8]*GRAPPA, Institute of Theoretical Physics, University of Amsterdam, Science Park 904, NL-1090 GL Amsterdam, the Netherlands*
[9]*Lorentz Institute, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*
[10]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo,CP 66318, São Paulo, SP, 05314-970, Brazil*
[11]*Laboratório Interinstitucional de e-Astronomia, Rua General José Cristino, 77, São Cristóvão, Rio de Janeiro, RJ, 20921-400, Brazil*
[12]*ICTP South American Institute for Fundamental Research & Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil*
[13]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[14]*School of Physics and Astronomy, Sun Yat-Sen University, Guangzhou 510275, China*
[15]*Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil*
[16]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[17]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, D-81679München, Germany*
[18]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[19]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa*
[20]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[21]*Institut d'Astrophysique de Paris, Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, F-75014 Paris, France*
[22]*Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA*
[23]*National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[24]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona) Spain*
[25]*Kavli Institute for Particle Astrophysics and Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA*
[26]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ, 20921-400, Brazil*
[27]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
[28]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[29]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[30]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[31]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[32]*Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland*
[33]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[34]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[35]*Astronomy Department, University of Washington, Box 351580, Seattle, WA 98195, USA*
[36]*Australian Astronomical Observatory, North Ryde, NSW 2113, Australia*

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018

[37] *Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[38] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*
[39] *School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[40] *Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, SP, Campinas, Brazil*
[41] *Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[42] *Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile*

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

Downloaded from https://academic.oup.com/mnras/article-abstract/479/1/94/5017483
by University College London user
on 03 September 2018