Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **513,** 5517–5539 (2022) https://doi.org/10.1093/mnras/stac1160

Dark Energy Survey Year 3 results: calibration of lens sample redshift distributions using clustering redshifts with BOSS/eBOSS

R. Cawthon[1,2★] J. Elvin-Poole,[3,4] A. Porredon,[3,5,6] M. Crocce,[5,6] G. Giannini,[7] M. Gatti,[7] A. J. Ross,[3] E. S. Rykoff,[8,9] A. Carnero Rosell,[10,11] J. DeRose,[12,13] S. Lee,[14] M. Rodriguez-Monroy,[15] A. Amon,[8] K. Bechtol,[1] J. De Vicente,[15] D. Gruen,[8,9,16] R. Morgan,[1] E. Sanchez,[15] J. Sanchez,[17] I. Sevilla-Noarbe,[15] T. M. C. Abbott,[18] M. Aguena,[19,20] S. Allam,[17] J. Annis,[17] S. Avila,[21] D. Bacon,[22] E. Bertin,[23,24] D. Brooks,[26] D. L. Burke,[8,9] M. Carrasco Kind,[26,27] J. Carretero,[7] F. J. Castander,[5,6] A. Choi,[3] M. Costanzi,[28,29] L. N. da Costa,[20,30] M. E. S. Pereira,[31] K. Dawson,[32] S. Desai,[33] H. T. Diehl,[17] K. Eckert,[34] S. Everett,[13] I. Ferrero,[35] P. Fosalba,[5,6] J. Frieman,[17,36] J. García-Bellido,[21] E. Gaztanaga,[5,6] R. A. Gruendl,[26,27] J. Gschwend,[20,30] G. Gutierrez,[17] S. R. Hinton,[37] D. L. Hollowood,[13] K. Honscheid,[3,4] D. Huterer,[31] D. J. James,[38] A. G. Kim,[39] J.-P. Kneib,[40] K. Kuehn,[41,42] N. Kuropatkin,[17] O. Lahav,[25] M. Lima,[19,20] H. Lin,[17] M. A. G. Maia,[20,30] P. Melchior,[43] F. Menanteau,[26,27] R. Miquel,[7,44] J. J. Mohr,[45,46] J. Muir,[8] J. Myles,[8,9,16] A. Palmese,[17,36] S. Pandey,[34] F. Paz-Chinchón,[27,47] W. J. Percival,[48] A. A. Plazas,[43] A. Roodman,[8,9] G. Rossi,[49] V. Scarpine,[17] S. Serrano,[5,6] M. Smith,[50] M. Soares-Santos,[31] E. Suchyta,[51] M. E. C. Swanson,[27] G. Tarle,[31] C. To,[8,9,16] M. A. Troxel,[14] and R. D. Wilkinson[52] (DES Collaboration)

Affiliations are listed at the end of the paper

Accepted 2022 April 22. Received 2022 April 10; in original form 2021 January 21

ABSTRACT

We present clustering redshift measurements for Dark Energy Survey (DES) lens sample galaxies used in weak gravitational lensing and galaxy clustering studies. To perform these measurements, we cross-correlate with spectroscopic galaxies from the Baryon Acoustic Oscillation Survey (BOSS) and its extension, eBOSS. We validate our methodology in simulations, including a new technique to calibrate systematic errors that result from the galaxy clustering bias, and we find that our method is generally unbiased in calibrating the mean redshift. We apply our method to the data, and estimate the redshift distribution for 11 different photometrically selected bins. We find general agreement between clustering redshift and photometric redshift estimates, with differences on the inferred mean redshift found to be below $|\Delta z| = 0.01$ in most of the bins. We also test a method to calibrate a width parameter for redshift distributions, which we found necessary to use for some of our samples. Our typical uncertainties on the mean redshift ranged from 0.003 to 0.008, while our uncertainties on the width ranged from 4 to 9 per cent. We discuss how these results calibrate the photometric redshift distributions used in companion papers for DES Year 3 results.

Key words: surveys – galaxies: distances and redshifts – large-scale structure of Universe – cosmology: observations.

1 INTRODUCTION

Large galaxy imaging surveys have proven to be an effective tool for understanding the cosmos. Optical surveys, such as the Dark Energy Survey (DES; DES Collaboration 2005), the Kilo-Degree Survey (KiDS; de Jong et al. 2013) and the Hyper Suprime-Cam (HSC; Aihara et al. 2018), have shown the ability to catalogue millions of galaxies and to extrapolate cosmological information out to redshift, $z \sim 1$, probing the structure and dynamics of the Universe in the past $\sim$ six billion years (DES Collaboration 2018; Asgari et al. 2021). Accompanying this work, DES Collaboration et al. (2022) shows the latest analysis of the structure of the Universe using galaxy clustering and weak lensing measurements of more than 100 million galaxies. In the future, surveys such as the Vera Rubin Observatory Legacy Survey of Space and Time (LSST; Ivezić et al. 2019) and Euclid (Laureijs et al. 2011) will extend such analyses to include billions of galaxies further back in time.

A critical component of these imaging surveys is the estimation of galaxy redshifts. Accurate redshift information is necessary for precise cosmological measurements of the growth of structure across time. However, large imaging surveys tend not to have spectroscopic capabilities. Instead, spectral information tends to be limited to magnitude estimates in a few colour bands. In the DES, imaging data include the g, r, i, z and Y bands. There is a large body of literature that deals with the estimation of photometric redshifts (photo-z); see, for example, Hildebrandt et al. (2021) and Hoyle et al. (2018), and references therein. In these methods, a redshift estimate is extracted

★ E-mail: rcawthon28@gmail.com

from these few colour and magnitude measurements. These methods all require some form of testing on galaxies where photometric and spectroscopic measurements are taken. Despite much success with these methodologies, the best photometric redshift estimates in the DES for particularly suitable samples of galaxies are thought to have uncertainties around $\sigma_z \approx 0.02$ for individual galaxies, with many samples much more uncertain. These errors are orders of magnitude larger than typical spectroscopic redshift errors. One particular issue is that a systematic bias can emerge if the test samples of galaxies are not fully representative of the galaxies being studied (Rivera et al. 2018). This may happen, for example, because of a difference in the depth of the samples. Extrapolating from a few colour-band measurements to a precise redshift remains a difficult problem.

In recent years, an alternative and complementary method of estimating the redshifts of galaxies has developed. The approach, called 'clustering redshifts' or 'cross-correlation redshifts', computes an angular cross-correlation of the galaxy sample in question and a galaxy sample with known (spectroscopic) redshifts. This cross-correlation will contain a signal proportional to the redshift overlap of the two samples. The method is completely independent of photometry, not relying on the colour–magnitude information at all (other than for initially binning the galaxies). Instead, it relies on gravity. Because galaxies cluster, objects near each other in angular coordinates are more likely to be near each other in radial separation, and thus redshift. While this spatial information will not be significantly informative on a galaxy by galaxy basis, it is very useful probabilistic information when trying to estimate the redshift distribution of thousands or millions of galaxies.

The use of angular clustering to infer proximity in distance between two samples extends back to Seldner & Peebles (1979) and Phillipps & Shanks (1987). The modern method of using that information for a rigorous estimate of a redshift distribution traces back to Newman (2008). It has since been developed theoretically and implemented on data in a number of papers including Matthews & Newman (2010), McQuinn & White (2013), Ménard et al. (2013), Schmidt et al. (2013), Choi et al. (2016), Scottez et al. (2016), Johnson et al. (2017) Krolewski et al. (2020), Hildebrandt et al. (2021), and van den Busch et al. (2020). In the DES Year 1 cosmology analysis (DES Collaboration 2018), clustering redshifts of both lens and source galaxies were computed (Davis et al. 2017; Cawthon et al. 2018; Gatti et al. 2018).

In this work, we present the clustering redshift estimates for the DES 'lens' galaxies used in the 'Year 3' cosmological analyses (based on data from the first three years of DES observations). These galaxies are used as lenses for galaxy–galaxy lensing measurements, and for galaxy clustering measurements in the cosmology analysis in DES Collaboration et al. (2022). The lens galaxies and those measurements are analysed in more detail in several related DES Year 3 analyses (Pandey et al. 2021; Porredon et al. 2021a; Prat et al. 2021; Rodríguez-Monroy et al. 2022; Elvin-Poole et al., in preparation). There are two samples of DES lens galaxies presented in these works: the red-sequence matched-filter galaxy catalogue (redMaGiC, for short) and a magnitude-limited sample called MagLim. redMaGiC (Rozo et al. 2016) is an algorithm that finds luminous red galaxies (LRGs) by using the red sequence of galaxies (Gladders & Yee 2000; Rykoff et al. 2014). This type of selection has been shown to give fairly small photometric redshift errors for the sample. A similar sample was used in the DES Year 1 analysis (Cawthon et al. 2018; Elvin-Poole et al. 2018; DES Collaboration 2018). The MagLim sample is a denser sample that goes to slightly higher redshifts, as described in Porredon et al. (2021b), and it is expected to have more uncertainty in its photo-z estimates than redMaGiC. The redMaGiC and MagLim samples are split into five and six tomographic redshift bins respectively, selected by photo-z estimates.

To calibrate the redshift distributions of these two samples, in each of their redshift bins, we cross-correlate them with spectroscopic samples of galaxies. For these spectroscopic samples, we use galaxies observed by the Sloan Digital Sky Survey (SDSS; Gunn et al. 2006; Eisenstein et al. 2011; Smee et al. 2013; Blanton et al. 2017). Specifically, we use galaxies from the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), as used by Cawthon et al. (2018), as well as from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016). About 15 per cent of the DES Year 3 samples overlap with BOSS and eBOSS.

Much of the methodology in this work is similar to that of Cawthon et al. (2018). We briefly highlight the main differences in this analysis.

(i) We use significantly larger data sets for the DES and spectroscopic reference galaxies. In addition to the larger area of coverage for the DES, we calibrate two lens samples (redMaGiC and MagLim) while DES Year 1 results only used redMaGiC. For spectroscopic samples, we are able to use more of BOSS due to the wider area of the DES in Year 3. We are also able to use the eBOSS galaxy catalogue, which greatly improves the redshift coverage available, increasing the maximum redshift of our study from roughly $z = 0.7$ to $z = 1.15$. As a result of both area and redshift coverage, the overall number of DES redMaGiC galaxies and spectroscopic galaxies used in this work are each a factor of 10 larger than in Cawthon et al. (2018). In addition, the MagLim sample is about 3.5 times larger than the redMaGiC sample in the Year 3 studies.

(ii) While much of the methodology is the same, it is much more extensively tested in simulations. These tests were possible because the simulated spectroscopic samples are similar to the BOSS and eBOSS catalogues. These tests give a more thorough estimate of the errors and uncertainties in the method.

(iii) We introduce a novel step in correcting for the evolution of the galaxy clustering bias. The galaxy bias describes the relationship between the distribution of galaxies and of total matter. The change in this parameter with redshift within a single tomographic bin is known as a challenging systematic in the clustering redshifts method; see van den Busch et al. (2020) for a recent review of attempts to correct this effect. Autocorrelations of galaxies can in principle be used both for the photometric and spectroscopic samples as an estimate of the galaxy bias, which is then calibrated out of the clustering redshift estimate. Because the DES samples do not have spectroscopic redshifts, their autocorrelations are a function of not only the galaxy bias (and cosmology) but also the scatter in their true redshift distributions. In the DES Year 1 analysis in Cawthon et al. (2018), this photo-z scatter effect on the autocorrelations was calibrated from simulations. In our work, we calibrate this scatter effect with cross-correlations of the DES and spectroscopic samples on smaller redshift bins. The main advantage of this new step is that it is empirically driven, no longer assuming any information from simulations (although the step is tested along with all the others in simulations).

(iv) We test a few different 'two-parameter fits', which effectively constrain both the mean redshift and the width of a distribution. In detail, the fits solve for a shift and a stretch of a photo-z distribution to better match the clustering redshift data. This procedure is needed when the shapes of the two distributions mismatch, and a single shift parameter would not make them match well enough.

We also note that a clustering redshift measurement of the weak lensing 'source' galaxies in DES Collaboration et al. (2022) and companion papers is performed in Gatti et al. (2022), which has several similarities and differences in methodology compared with this work. One example regards constraining the galaxy bias evolution. For the lens sample, we have a generated random catalogue, which samples the survey selection function. This gives us a greater ability to measure the galaxy bias evolution effects with autocorrelations. Because the source galaxies do not have such a catalogue, Gatti et al. (2022) use a more agnostic model to account for galaxy bias evolution.

The structure of our paper is as follows. In Section 2, we discuss the data sets used in this work. In Section 3, we describe the simulated data sets used for validating our methodology. In Section 4, we present our methodology for performing a clustering redshift measurement and calibrating it to find a best-fitting shift, or shift and stretch parameters to be applied to a photometric estimate of the redshift distribution. In Section 5, we validate our methodology in simulations and derive systematic uncertainties for different parts of the method, as well as testing different methods for performing a two-parameter fit. In Section 6, we show our results, the clustering redshift measurements of each of the redshift bins of the two DES lens samples. In Section 7, we calculate a theory prediction for magnification effects in our measurements, showing they are likely to be insignificant. In Section 8, we summarize our work.

2 DATA SETS

In this section, we describe the data sets used for the spectroscopic reference galaxies and the photometric DES galaxies that we wish to calibrate. The redMaGiC and MagLim samples are derived from the 'Y3 Gold catalogue' (Sevilla-Noarbe et al. 2021), which contains galaxies found in the first three years of DES data. The Gold catalogue covers the full DES footprint of nearly 5000 deg^2, and contains around 388 million objects. The two samples are used for cosmological analyses in DES Collaboration et al. (2022), and are described in detail in Rodríguez-Monroy et al. (2022) and Porredon et al. (2021b). We repeat some of the main information about each sample here. This work only uses the part of the DES catalogues that overlaps the sky area of BOSS or eBOSS galaxies, about 860 deg^2, with slightly less overlap at higher redshifts. Masks are also derived from the Gold catalogue as well as random galaxy catalogues, which reflect the survey selection efficiency at different points. After masks are applied, the effective DES area in our study is 632 deg^2. In Appendix E, we describe and analyse a third sample, called the 'flux-limited' sample, which is not used in the cosmology analyses.

2.1 DES redMaGiC sample

To create the redMaGiC sample, the cluster-finding algorithm redMaPPer (Rykoff et al. 2014) is run on the Gold catalogue to calibrate the red sequence of galaxies (Gladders & Yee 2000). The redMaGiC algorithm (Rozo et al. 2016) then selects luminous red galaxies with colours that fit with the red-sequence template. This fitting also estimates a redshift probability distribution function for each LRG. The redMaGiC algorithm further tunes the colour selection threshold to produce a constant comoving density, which is expected for passively evolving red galaxies (Rozo et al. 2016). redMaGiC galaxy catalogues were similarly used for DES Year 1 analyses (Cawthon et al. 2018; DES Collaboration 2018; Elvin-Poole et al. 2018).

Table 1. redMaGiC galaxies used in this work.

Redshift bin	L/L_*	$n_{\rm gal}$ (arcmin^{-2})	$N_{\rm gal}$
1: $z_{\rm ph} \in [0.15, 0.35]$	0.5	0.027	61 586
2: $z_{\rm ph} \in [0.35, 0.5]$	0.5	0.049	110 586
3: $z_{\rm ph} \in [0.5, 0.65]$	0.5	0.075	170 102
4: $z_{\rm ph} \in [0.65, 0.8]$	1.0	0.038	86 767
5: $z_{\rm ph} \in [0.8, 0.9]$	1.0	0.032	72 833

The redMaGiC algorithm selects galaxies above a given luminosity threshold. For the Year 3 lens samples, thresholds of either 0.5 L_* or 1.0 L_* were used for the different redshift bins (see Table 1). The reference luminosity, L_*, comes from a model (Bruzual & Charlot 2003) for a single star-formation burst at $z = 3$, as described in Rykoff et al. (2014). For the reference luminosities 0.5 L_* and 1.0 L_*, the comoving densities produced by the redMaGiC algorithm are $\bar{n} = 10^{-3}$ and 4×10^{-4} galaxies $(h^{-1}{\rm Mpc})^{-3}$ respectively, where h is the reduced Hubble constant (Rodríguez-Monroy et al. 2022). The redMaGiC galaxies are split into tomographic bins by the mean redshift of each galaxy's redshift probability distribution function produced by the redMaGiC algorithm. We show the number of galaxies used in this analysis (covering the 632 deg^2 of overlap with BOSS) for each tomographic bin in Table 1.

We also apply weights to redMaGiC galaxies as described in Rodríguez-Monroy et al. (2022). These weights are selected based on survey properties, such as seeing and sky brightness, for each of the observed galaxies. The weights are chosen to minimize the effects of these survey properties on galaxy clustering measurements.

2.2 DES MagLim sample

The DES MagLim samples are described in Porredon et al. (2021b). They are created using a redshift-dependent magnitude cut, with the redshift estimate for each galaxy coming from the DNF (De Vicente, Sánchez & Sevilla-Noarbe 2016) photometric redshift algorithm. This redshift dependence tends to eliminate faint, low-redshift galaxies from entering the sample (as verified in Appendix E). The creation of this sample was motivated by a significantly larger number density than redMaGiC. However, photo-z error estimates were expected to be larger, making the calibration of photo-z biases in this work essential.

In Porredon et al. (2021b), a Fisher forecast is run to find the best magnitude cuts for the DES cosmology analyses, with different cuts trading off number density and larger photo-z scatter. From that work, the optimal redshift-dependent cut is selecting on i-band magnitude, $i < 4\,z_{\rm phot} + 18$. Bright galaxies with $i < 17.5$ are also removed. These MagLim galaxies are split into tomographic bins by the mean redshift of the redshift probably distribution function given by DNF. Notably, the MagLim sample extends to slightly higher redshifts than redMaGiC. The numbers of galaxies used in this work for each tomographic bin (again reflecting only the galaxies overlapping with BOSS) are shown in Table 2. Again, we also apply survey property weights as described in Rodríguez-Monroy et al. (2022).

2.3 BOSS galaxies (SDSS DR12)

Our first source of reference galaxies comes from catalogues created by the BOSS from SDSS Data Release 12 (DR12; Alam et al. 2015). We use their LOWZ and CMASS galaxy and random catalogues described in Reid et al. (2016). For $z > 0.6$, we use a joint catalogue of CMASS galaxies and eBOSS LRG galaxies created by eBOSS

Table 2. MagLim galaxies used in this work.

Redshift bin	$n_{\rm gal}$ (arcmin^{-2})	$N_{\rm gal}$
1: $z_{\rm ph} \in [0.2, 0.4]$	0.154	349 673
2: $z_{\rm ph} \in [0.4, 0.55]$	0.115	260 671
3: $z_{\rm ph} \in [0.55, 0.7]$	0.115	262 468
4: $z_{\rm ph} \in [0.7, 0.85]$	0.154	349 996
5: $z_{\rm ph} \in [0.85, 0.95]$	0.117	266 750
6: $z_{\rm ph} \in [0.95, 1.05]$	0.113	257 139

to prevent double counting of galaxies (see Ross et al. 2020). Our spectroscopic tracers at $z < 0.6$ are solely from the BOSS samples. These samples were also used for clustering redshifts in DES Year 1 cosmology (Cawthon et al. 2018; DES Collaboration 2018). These catalogues were optimized for clustering in order to measure the baryon acoustic oscillation (BAO) signal (Alam et al. 2017), but their wide, uniform coverage of the sky makes them one of the best spectroscopic data sets for clustering redshifts.

2.4 eBOSS (SDSS DR16)

We also use spectroscopic galaxies from the eBOSS. The galaxies are part of the SDSS Data Release 16 (DR16; Ahumada et al. 2020). We use the large-scale structure (LSS) catalogues of emission-line galaxies (ELGs), luminous red galaxies (LRGs) and quasi-stellar objects (QSOs). The creation of the ELG catalogues is described in Raichoor et al. (2021) and the LRG and QSO catalogues are described in Ross et al. (2020). The catalogues were provided to DES before being made public for clustering redshift usage by agreement between DES and eBOSS.

The target selection for the eBOSS ELG sample is described in Raichoor et al. (2017). The sample selection is based on observations from the Dark Energy Camera Legacy Survey (DECaLS; Dey et al. 2019) with colour and magnitude cuts to yield strong [O II] emitters in the redshift range of $0.6 < z < 1.1$. The LRG sample selection is described in Prakash et al. (2016). The LRGs were selected using colour and magnitude cuts on objects found in SDSS and the *Wide-Field Infrared Survey Explorer* (*WISE*; Wright et al. 2010) photometry. The LRG sample primarily spans the range $0.6 < z < 1.0$. The LRGs are combined with the BOSS CMASS sample because there are duplicate objects. The QSO sample selection is described in Myers et al. (2015). The QSO sample spans from $0.9 < z < 2.2$, although we only use up to $z = 1.18$ for clustering redshift measurements because of the low density of objects at higher redshifts for both DES and reference samples. The target selection used photometric observations from the SDSS as well as *WISE*.

The details of creating LSS data sets from these samples are described in Raichoor et al. (2021) and Ross et al. (2020). We use the weights (given by $w_{\rm tot}$ in Ross et al. 2020) and random points associated with these catalogues to account for the survey selection function. We also use the combined LRG catalogue using eBOSS galaxies as well as $z > 0.6$ BOSS CMASS galaxies, as described in Ross et al. (2020). We show the total number of reference galaxies used in this work by their catalogue source in Fig. 1 and Table 3. In our measurements, we combine all these catalogues into a single sample.

3 SIMULATED DATA SETS

Along with many of the other accompanying papers related to DES Collaboration et al. (2022), our paper makes use of the Buzzard simulations (DeRose et al. 2019, 2021). Buzzard simulates a Universe of dark matter only, which is then populated with galaxies by the ADDGALS algorithm (Wechsler et al., in preparation). ADDGALS is calibrated by a series of algorithms, many of which are fit empirically to galaxy distributions (in terms of luminosity, clustering, abundance, etc.) of SDSS galaxies (e.g. subhalo abundance matching fits from Lehmann et al. 2017). The resulting galaxy catalogues are then sampled similarly to how DES creates its cosmological data sets. For the samples used in this work, this specifically means running the redMaPPer and redMaGiC algorithms on Buzzard to create a simulated redMaGiC catalogue, and using the colour and magnitude cuts from the MagLim sample to create a simulated version of it. The Buzzard simulations used for DES Year 3 analyses are described in more detail in DeRose et al. (2021), where the simulated DES data sets are shown to replicate well the galaxy properties and cosmological measurements from data.

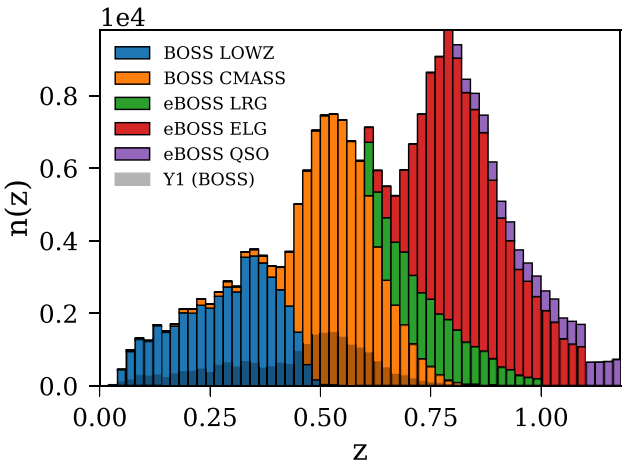

Figure 1. The BOSS/eBOSS $n(z)$ used in this work as reference samples. In shaded outline, we show the BOSS $n(z)$ used in the Year 1 analysis of Cawthon et al. (2018). This work uses about a factor of 10 more reference galaxies overall. The numbers of galaxies from each BOSS/eBOSS catalogue are shown in Table 3.

Table 3. Spectroscopic samples used as the reference galaxies for clustering redshifts in this work. We show the approximate redshift range of the BOSS samples used. In contrast, the eBOSS catalogues each have set redshift boundaries.

Name	Redshifts	$N_{\rm gal}$	Area
LOWZ (BOSS)	$z \sim [0.0, 0.5]$	45 671	$\sim$860 deg^2
CMASS (BOSS)	$z \sim [0.35, 0.8]$	74 186	$\sim$860 deg^2
LRG (eBOSS)	$z \in [0.6, 1.0]$	24 404	$\sim$700 deg^2
ELG (eBOSS)	$z \in [0.6, 1.1]$	89 967	$\sim$620 deg^2
QSO (eBOSS)	$z \in [0.8, 1.18]$	10 502	$\sim$700 deg^2

3.1 Simulated DES redMaGiC

As described in DeRose et al. (2021), the redMaGiC algorithm is run on Buzzard galaxies similar to the procedure on data. In particular, colour-dependent clustering was improved for Year 3 Buzzard to better match the redMaGiC-selected galaxies in data. The same redshift and L_* cuts applied on the data are applied to obtain the simulated samples.

3.2 Simulated DES MagLim sample

For MagLim galaxies in Buzzard, we use a similar redshift-dependent magnitude cut as is done on data. For the tests in this work, a slightly older version of the MagLim cuts was used to generate the sample. This version selected galaxies with i-band magnitude, $i < 4.28\, z_{\rm phot} + 18$. It also cut out bright galaxies with $i < 17.5$. The slight differences from the final MagLim cuts on data should not change the efficacy of our clustering redshift method, which is what the simulations are used to check. We do not use any information on, for example, galaxy bias or photo-z scatter from these simulated samples in our measurements. The redshift bins are selected in the same way as on data.

3.3 Simulated BOSS (CMASS) sample

To simulate the BOSS CMASS sample in Buzzard, we use the DMASS algorithm described in Lee et al. (2019). The goal of the DMASS algorithm was to create a CMASS-like sample of galaxies from the DES samples of galaxies. Because the properties of CMASS galaxies have been well characterized, a large CMASS-like sample in DES would be useful for several studies.

In Lee et al. (2019), the DMASS algorithm is trained on the overlapping area of DES and BOSS to derive a Bayesian model based on galaxy colours and magnitudes for any DES galaxy to be CMASS-like. For our work, this algorithm is used on the Buzzard-simulated DES galaxies. Each galaxy is given a CMASS-like probability. We then take one random draw based on these probabilities to define our simulated CMASS sample.

3.4 Simulated eBOSS (ELG) sample

To simulate the eBOSS ELG sample in Buzzard, we use the magnitude and colour cuts used for target selection in Raichoor et al. (2017) for the South Galactic Cap (SGC) sample. Since these targets were found by the Dark Energy Camera (DECam), the targeting magnitudes are in the DECam filter bands.

We show comparisons of the simulated BOSS/eBOSS samples to their real counterparts in Fig. 2. There is relative agreement in both redshift distribution and amount of clustering. We note that perfect agreement is not necessary, as the clustering and redshift distribution of the BOSS/eBOSS samples are well measured on the data, and there is no explicit reason why the method's accuracy should depend strongly on redshift or galaxy bias. The relative agreement should be sufficient to validate the methodology. We investigate the method's dependence on the number of galaxies in Appendix D.

4 METHODS

We now lay out our methodology for the clustering redshift measurement. A number of different redshift distributions, binning schemes and correlation functions are mentioned in this section. To aid the reader, a summary of various terms is shown in Table 4.

4.1 Unknown and reference correlation measurement

The clustering redshift methodology involves a cross-correlation of two samples: an 'unknown' sample with undetermined redshifts, and a 'reference' sample with known redshifts. For this work, the unknown samples will be the DES samples (redMaGiC and MagLim) and the reference sample will be the combined BOSS/eBOSS spectroscopic data set.

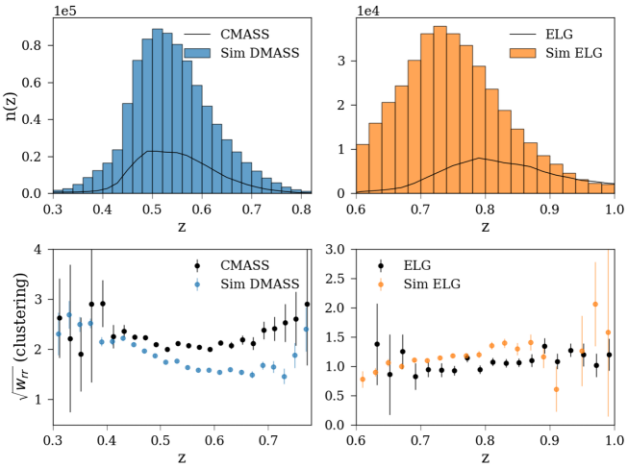

Figure 2. Comparison of simulated and real spectroscopic data sets. We show BOSS CMASS (South) and DMASS, an algorithm run on simulations, and also eBOSS ELG (South Galactic Cap) and a simulated ELG sample. The top row shows the redshift distributions. The bottom row shows the square root of a weighted autocorrelation (equation 5). The simulated samples go over the full DES 5000 deg^2, so they are larger than the real data sets.

Table 4. Definitions for various redshift distributions, binning schemes and correlation functions referred to in Section 4.

Redshift distributions	
$n_{{\rm u},i}(z)$	The true redshift distribution of an unknown sample in photometric redshift bin i
$n_{{\rm u},j}(z)$	The true redshift distribution of a sample binned by photometric redshift in micro-bin j
$n_{{\rm spec},j}(z)$	The true redshift distribution of a sample binned by spectroscopic (true) redshift in micro-bin j
$n^{{\rm pz},i}(z)$	The photometric redshift distribution of a photometric redshift bin i
Binning schemes	
Photometric bins, i	The main target bins used for cosmology, with sizes in the range $dz = 0.1-0.2$
Micro-bins, j	The reference sample bins of size $dz = 0.02$; unknown and reference samples use these bins for autocorrelations
Nano-bins	Bins of size $dz = 0.005$–0.01, which are used for computing the width of $n_j(z)$ in equation (10)
Correlation functions	
$\bar{w}_{\rm ur}$	Weighted cross-correlation between an unknown (photometric) and reference (spectroscopic) galaxy samples
$\bar{w}_{\rm rr}$	Weighted autocorrelation of a reference sample
$\bar{w}_{\rm uu,pz}$	Weighted autocorrelation of an unknown sample that is binned by photometric redshift
$\bar{w}_{\rm uu,spec}$	Weighted autocorrelation of an unknown sample that is binned by spectroscopic redshift (typically not possible with data)

We use a cross-correlation version of the Landy–Szalay estimator (Landy & Szalay 1993) over physical scales r,

$$w(r) = \frac{D_1D_2(r) - D_1R_2(r) - D_2R_1(r) + R_1R_2(r)}{R_1R_2(r)} \quad (1)$$

where $w(r)$ is the excess probability of finding a pair of galaxies a distance r away compared with a random sample, D signifies a data set of galaxies, R signifies a random distribution of galaxies and D_1D_2, for example, is the number of pairs between the two data sets separated by comoving length-scale r. The length-scale is set by $r = \theta\chi(z)$, where θ is the observed angle between the two galaxies and $\chi(z)$ is the comoving distance calculated using the Planck 2015 cosmology (Planck Collaboration XIII 2016). The redshifts are set by the centre of the reference sample bins. For all of our measurements, we use a weighted averaged estimate of $w(r)$ over a range of r values, $\bar{w}_{12}$:

$$\bar{w}_{12} = \int_{r_{\rm min}}^{r_{\rm max}} r^{-1} w(r)\, {\rm d}r. \quad (2)$$

Unless otherwise stated, we use eight bins between $r_{\rm min} = 0.5$ Mpc and $r_{\rm max} = 1.5$ Mpc. These parameters, as well as the weighting by r^{-1}, were first shown to be effective for clustering redshifts by Schmidt et al. (2013) and were used in the DES Year 1 analyses (Davis et al. 2017; Cawthon et al. 2018; Gatti et al. 2018). These comoving scales are smaller than the scales used for related cosmological galaxy clustering studies in DES Collaboration et al. (2022) and others to reduce covariance of the measurements. For all of the following weighted cross-correlations and autocorrelations given by equation (2), statistical errors are measured by 100 jackknife resamplings.

Our weighted cross-correlation, $\bar{w}$ of the unknown (u) and reference (r) samples should be written as

$$\bar{w}_{\rm ur} = \int_{z_{\rm min}}^{z_{\rm max}} n_{\rm u}(z) n_{\rm r}(z) b_{\rm u}(z) b_{\rm r}(z) \bar{w}_{\rm mm}(z)\, {\rm d}z, \quad (3)$$

where $n_{\rm u}$ and $n_{\rm r}$ are the normalized redshift distributions of the unknown and reference galaxy samples, $b_{\rm u}$ and $b_{\rm r}$ are the galaxy biases of the two samples and $\bar{w}_{\rm mm}$ is the weighted cross-correlation of the total (primarily dark) matter distribution.

We now introduce two redshift binning schemes that are important for our work. The goal is to derive correct mean redshifts for the photometric redshift-binned DES samples to be used in the DES cosmology analyses. There are five bins for redMaGiC, and six bins for the MagLim sample. We call these the photometric bins and they are signified by i. These bins are typically ${\rm d}z = 0.1{-}0.2$ in size. To obtain a measurement of $n_{{\rm u},i}(z)$ relevant for these photometric bins, we need to measure using thinner bin widths. These thinner bins will be of size ${\rm d}z = 0.02$. We call these the micro-bins and they are signified by j. Our spectroscopic reference samples will always be binned in these smaller ${\rm d}z = 0.02$ bins. We refer to the centres of these micro-bins as z_j. Our goal is to measure the photometric redshift sample's redshift distribution, $n_{{\rm u},i}$, in each of the micro-bins at z_j. The characters u and r will always refer to samples, and the characters i and j will always refer to bins. In one micro-bin, we know the exact number of galaxies in the reference sample (as it has spectroscopic redshifts). Going from equation (3), our estimate for $n_{{\rm u},i}$ at a micro-bin centred at z_j is

$$n_{{\rm u},i}(z_j) \propto \bar{w}_{\rm ur}(z_j) \frac{1}{b_{\rm u}(z_j)} \frac{1}{b_{\rm r}(z_j)} \frac{1}{\bar{w}_{\rm mm}(z_j)}, \quad (4)$$

where $n_{{\rm u},i}(z_j)$ is the desired quantity of the number of galaxies in unknown (photometric) sample i in the micro-bin centred at z_j, and $\bar{w}_{\rm ur}(z_j)$ is the weighted cross-correlation of the unknown sample in photometric bin i and reference sample in micro-bin j. As seen in the equation, we assume that within a micro-bin, the galaxy bias of each sample is constant.

It is easiest to note here that the key issue with clustering redshifts is not galaxy bias, but rather the galaxy bias evolution with redshift. In equation (4), if the galaxy biases are the same for all z_j (even if unknown), they will effectively cancel out when all $n_{{\rm u},i}(z_j)$ are combined, because the total number of galaxies in the ith bin is known. If the galaxy biases change with redshift within a single photometric bin though, they will not cancel out and will distort the estimated $n(z)$.

4.2 Correcting for galaxy bias

We can get closer to solving for $n_{{\rm u},i}(z_j)$ in equation (4) by using autocorrelations of each sample binned by the micro-bin j. The weighted autocorrelations for the samples, again assuming a single galaxy bias value for the micro-bin, are

$$\bar{w}_{\rm rr}(z_j) = b_{\rm r}(z_j)^2 \bar{w}_{\rm mm}(z_j) \int n_{{\rm r},j}(z)^2\, {\rm d}z, \quad (5)$$

$$\bar{w}_{{\rm uu},j}(z_j) = b_{\rm u}(z_j)^2 \bar{w}_{\rm mm}(z_j) \int n_{{\rm u},j}(z)^2\, {\rm d}z. \quad (6)$$

We note that in equation (6) we have introduced new quantities, $\bar{w}_{{\rm uu},j}$ and $n_{{\rm u},j}$. These are quantities related to an unknown sample (i.e. a DES sample) binned in a micro-bin j. Our previous equations had $n_{{\rm u},i}$, which relates to the unknown sample binned by photometric bin, i. These larger bins again correspond to the bins used by DES cosmology analyses, and thus what we ultimately want to figure out. However, as laid out in this section, we sometimes need to measure properties of this sample in smaller redshift slices (i.e. $n_{{\rm u},j}$) to ultimately work out $n_{{\rm u},i}$. We also note that these $n(z)$ refer to the true (spectroscopic) redshift distribution. Thus, for example, $n_{{\rm u},j}(z)$ is the true redshift distribution of galaxies binned into micro-bin j by photometric redshift. So, for example, $n_{{\rm u},j}(z)$ in micro-bin $z_{\rm pz} \in [0.2, 0.22]$ will extend beyond $z = 0.2$ and $z = 0.22$ in its true redshift distribution.

If spectroscopic redshifts are obtained, in the limit of a large number of galaxies, galaxy distributions tend to be fairly flat across the small redshift range of the micro-bins (${\rm d}z = 0.02$). In this limit, the normalized n^2 in the integrals of equations (5) and (6) is the same for all distributions and can be dropped. For spectroscopic (true) redshifts only, we can use equations (4)–(6) for an expression for $n_{{\rm u},i}(z_j)$ in terms of measurable correlation functions:

$$n_{{\rm u},i}(z_j) \propto \frac{\bar{w}_{\rm ur}(z_j)}{\sqrt{\bar{w}_{\rm rr}(z_j)\bar{w}_{{\rm uu,spec}}(z_j)}} \quad \text{(w/spec-}z\text{ only)}. \quad (7)$$

However, the assumption that $n^2_{{\rm u},j}$ is flat in equation (6) is almost certainly wrong as the unknown sample only has photometric redshifts. If the unknown sample is binned by photometric redshift into micro-bin j, then $n^2_{{\rm u},j}$ in equation (6) will span the entire true redshift range of that sample, which will extend beyond ${\rm d}z = 0.02$.

We can relate the theoretical autocorrelation of the unknown sample at z_j if it could be binned by spectroscopic redshift, to the measurable autocorrelation of the unknown sample binned by $z_{\rm ph}$ (photometric redshift) in micro-bin j:

$$\bar{w}_{{\rm uu,spec}}(z_j) \propto \bar{w}_{{\rm uu,pz}}(z_j) \frac{\int n_{{\rm spec},j}(z)^2\, {\rm d}z}{\int n_{{\rm u},j}(z)^2\, {\rm d}z}. \quad (8)$$

Here, $n_{{\rm spec},j}$ and $\bar{w}_{{\rm uu,spec}}$ are, respectively, the true redshift distribution and the theoretical autocorrelation of the unknown sample if it could

be binned by spectroscopic redshift into micro-bin j. Similarly, $n_{\mathrm{u},j}$ and $\bar{w}_{\mathrm{uu,pz}}$ are, respectively, the true redshift distribution and the measurable autocorrelation of the unknown sample when binned by photometric redshift into the micro-bin j.

This equation was used to solve for $n_{\mathrm{u},i}(z_j)$ in determining DES Year 1 clustering redshifts in Cawthon et al. (2018). Simulations were used to estimate both integrals. Again, for spectroscopic samples, $n(z)$ over a micro-bin tends to be flat and the upper integral can be dropped. The bottom integral, the true redshift distribution of the galaxies binned in micro-bin j by photo-z, is the main unknown. It essentially measures the photo-z scatter at redshift z_j, with more scatter producing a wider distribution, and smaller value of n^2.

In a change from Cawthon et al. (2018), we attempt to evaluate the photo-z scatter effect in equation (8) empirically by using clustering redshift measurements on the photometric galaxies binned in each micro-bin j by photometric redshift, as an estimate of $n_{\mathrm{u},j}(z)$. We assume in this narrower redshift range spanned by $n_{\mathrm{u},j}(z)$ that we can approximate $b_{\mathrm{u}}\bar{w}_{\mathrm{mm}}$ as constant. From that approximation and equations (4) and (5), we have

$$n_{\mathrm{u},j}(z) \propto \frac{\bar{w}_{\mathrm{ur}}(z)}{\sqrt{\bar{w}_{\mathrm{rr}}(z)}} \quad (\text{assume } b_{\mathrm{u}}\bar{w}_{\mathrm{mm}} = \text{const.}). \tag{9}$$

Measurements of $n_{\mathrm{u},j}(z)$ are noisier than evaluating $n_{\mathrm{u},i}(z)$ (equation 7). We are dividing up the DES photometric sample into smaller micro-bins j than the main photometric bins, i. Furthermore, as $n_{\mathrm{u},j}(z)$ is narrower than $n_{\mathrm{u},i}(z)$, we evaluate equation (9) on even smaller bins than the $\mathrm{d}z = 0.02$ micro-bins. These 'nano-bins' are either $\mathrm{d}z = 0.01$ or 0.005 depending on the signal-to-noise ratio. However, these further subdivisions make each measurement noisier. In order to reduce the computations and not propagate as many noisy data points, we make the approximation that the integral of $n^2(z)$ can be estimated by simply the inverse of the standard deviation of $n(z)$,

$$\int n_{\mathrm{u},j}(z)^2\,\mathrm{d}z \approx \frac{1}{\sigma_j}, \tag{10}$$

where σ_j is the standard deviation of the redshift distribution.

We note that photometric redshift scatter is often approximated as a Gaussian function (LSST Science Collaboration 2009; Cawthon 2020). If $n(z)$ is a Gaussian, that is,

$$n(z) \approx \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right],$$

where μ and σ are the mean and standard deviation, respectively, then the integral of n^2 directly evaluates to $\propto 1/\sigma$. Thus, the approximation of equation (10), particularly in trying to estimate photo-z scatter, seems appropriate.[1] We also note that if there is linear galaxy bias evolution across even the 'micro-bin' measurement (i.e. $b_{\mathrm{u}}\bar{w}_{\mathrm{mm}}$ varies across the small redshift range, just as we expect it to for the larger, main photometric bins), this would have far less impact on the standard deviation, σ, than it will on the mean redshift.

Using simulations, we validate the efficacy of the approximation in equation (10). In Fig. 3, we show an example of measuring σ_j, the standard deviation in one of the micro-bins, by carrying out clustering redshift measurements in the smaller 'nano-bins'. In Fig. 4, we show comparisons of the σ_j and $1/\int n^2$ correction terms measured in different ways in simulations and in data. We quantify the differences of each approximation (equations 7–10) in our tests in Section 5.

[1]Even in the case of a perfect Gaussian redshift distribution, equation (10) would not be exact because of the finite width of the micro-bins.

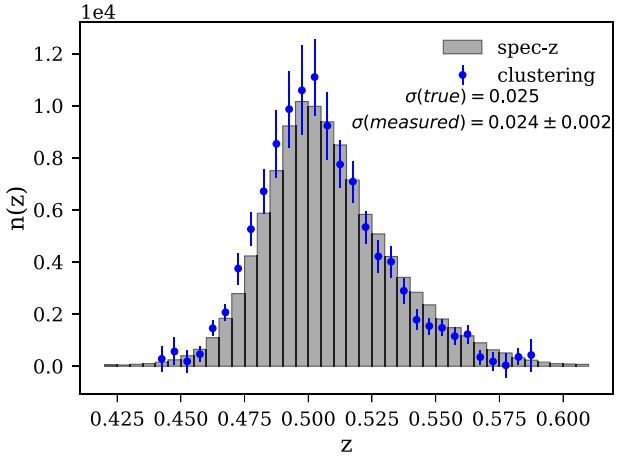

Figure 3. An example of clustering redshift measurements on a 'micro-bin' of size $\mathrm{d}z = 0.02$ in a simulation, showing the true redshift distribution and the one measured from clustering redshifts, on simulated redMaGiC galaxies with peak photo-z probability in the range $z = 0.49$–0.51. The goal of this particular measurement is to measure the standard deviation, σ, of $n(z)$. The standard deviation serves as a proxy for estimating how much an autocorrelation of this same micro-bin of redMaGiC will be reduced compared with the spectroscopic case – where the entire $n(z)$ would be entirely between 0.49 and 0.51. To perform this measurement, the spectroscopic sample (DMASS) is divided up into smaller 'nano-bins' of size $\mathrm{d}z = 0.005$ in order to obtain more data points for the measurement.

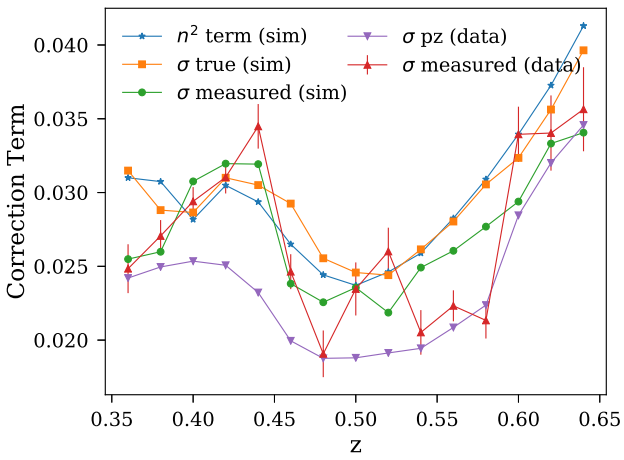

Figure 4. Various approximations of the 'correction term' needed to estimate the autocorrelation of a sample micro-binned by true (spectroscopic) redshift from the autocorrelation of the sample micro-binned by photometric redshift. The correction term is either σ, the standard deviation of an $n(z)$ estimate, or the 'n^2 term', $1/\int n^2_{\mathrm{u},j}$ from equation (8). The 'n^2 term' curve is normalized to the 'σ (true)' data points. Specifically, these σ estimates represent the standard deviation of the true redshift distribution when binned by photo-z. This value is either calculated exactly (simulations), measured by clustering redshifts on smaller 'nano-bins' (on either simulations or data) or calculated from the photometric redshift estimates (data). We quantify the agreement of some of these approaches in Section 5. The σ pz curve is the σ estimated from the photo-z algorithm itself, that is, from $n^{\mathrm{pz}}(z)$. The sample shown is the high-density redMaGiCsamples from $z_{\mathrm{ph}} = 0.35$–0.65 in both simulations and data.

Using equations (7)–(10), we find our estimate for $n_{\rm u}(z_j)$:

$$n_{\rm u}(z_j) \propto \frac{\bar{w}_{\rm ur}(z_j)}{\sqrt{\bar{w}_{\rm rr}(z_j)\bar{w}_{\rm uu,pz}(z_j)\sigma_j}}. \tag{11}$$

Because $\sqrt{\bar{w}_{\rm uu,pz}(z_j)\sigma_j}$ is a noisy approximation of $\sqrt{\bar{w}_{\rm uu,spec}(z_j)}$, we approximate it with a power law, as was done in Davis et al. (2018) and Cawthon et al. (2018):

$$\sqrt{\bar{w}_{\rm uu,spec}(z_j)} \approx \sqrt{\bar{w}_{\rm uu,pz}(z_j)\sigma_j} \propto (1+z)^{\gamma}. \tag{12}$$

We test the accuracy of our methodology in various steps in Section 5. Specifically, we test the method when using spectroscopic redshifts (equation 7), the power-law approximation when using spectroscopic redshifts (equation 12), and the approximate solutions when using photometric redshifts (equations 8 and 11).

4.3 Estimating photometric redshift bias (one-parameter fit)

Thus far, we have focused on the clustering redshift measurements themselves and their veracity. We now briefly discuss how specifically these measurements are used to calibrate the photometric redshift distributions used in DES Collaboration et al. (2022) and related papers.

Our general strategy is to use a calibrated photo-z distribution rather than clustering redshifts directly. This strategy is formed from a belief that the clustering redshifts are more accurate overall than the photometric estimate, but are not reliable in the tails due to noise and magnification effects. The clustering redshifts can also give unphysical, negative $n(z)$ measurements in the tails. A prediction-calibrated photo-z distribution, $n^{\rm pz}(z)$ – where we now use an upper index to distinguish from clustering or spectroscopic estimates of $n(z)$ – is thus preferable to using clustering redshifts directly.

For our fiducial plan, we assume that a single shift parameter, Δz, is enough to calibrate the photo-z distribution. This parameter is essentially the photo-z bias. Our final clustering distribution, $n_{\rm u}(z)$, comes from equations (11) and (12), estimated for each ith photometric bin, with data points in each micro-bin, j. Our final step is thus to find the shifted $n^{\rm pz}(z)$ that matches the mean of our $n_{\rm u}(z)$ from clustering redshifts. Specifically, we find the shift, Δz, that satisfies

$$\frac{\int_{z_{\rm min}}^{z_{\rm max}} z\, n^{\rm pz}(z-\Delta z)\,{\rm d}z}{\int_{z_{\rm min}}^{z_{\rm max}} n^{\rm pz}(z-\Delta z)\,{\rm d}z} = \frac{\int_{z_{\rm min}}^{z_{\rm max}} z\, n_{\rm u}(z)\,{\rm d}z}{\int_{z_{\rm min}}^{z_{\rm max}} n_{\rm u}(z)\,{\rm d}z}. \tag{13}$$

As in Cawthon et al. (2018) and Gatti et al. (2018), we set $z_{\rm min}$ and $z_{\rm max}$ to be at 2.5σ from the peak of the clustering redshift distribution. This cuts the tails of $n_{\rm u}(z)$ from being used in the calculation for the correct shift. The tails of a clustering redshift estimate can be noisy, with negative signals being difficult to calibrate. The tails can also be affected by magnification, as we discuss in more detail in Section 7.

4.4 Estimating photometric redshift bias + stretch (two-parameter fit)

In this work, we also use a two-parameter fit to calibrate the photo-z distribution with the clustering redshift measurements. In general, this fit will work in cases where a one-parameter shift of the photo-z distribution is a poor fit to the clustering data (i.e. the shapes of the distributions disagree).

For this fit, we use the Δz shift parameter from the one-parameter fit, as well as a 'stretch' parameter, s. The stretch parameter is included by shifting the photo-z distribution, such that its mean is centred at $z = 0$, then re-scaling the z-axis by a factor s, and finally shifting back to $z_{\rm mean}$. The functional form of this is given by

$$n_{\text{2-param}}(z) = \frac{1}{s} n_{\rm pz}\left(\frac{z - z_{\rm mean} - \Delta z}{s} + z_{\rm mean}\right), \tag{14}$$

where s is the new stretch parameter, equal to 1 if the width of the photo-z and clustering-z are the same, and Δz is the usual shift parameter. We refer to this as the two-parameter model. We apply a χ^2 least-squares fitting of s and Δz to the clustering redshift results. To account for the galaxy bias correction in a manner similar to our fiducial methods of Section 4, we propagate γ (equation 12) into the clustering redshift $n(z)$ and covariance when performing the χ^2 fit. In Section 5.7, we test in simulation the χ^2 fitting method and two other methods of fitting for Δz and s.

5 TESTING METHODOLOGY WITH SIMULATIONS

We validate various steps in our methodology, and estimate any associated biases or systematic uncertainties with those steps, with tests in simulations. As described in Section 2, we use the Buzzard simulations, and simulated samples of the DES redMaGiC and MagLim galaxies, BOSS CMASS galaxies and eBOSS ELG galaxies. We divide the simulated DES galaxies into six samples, corresponding to bins 2, 3 and 4 for redMaGiC, and bins 2, 3 and 4 for the MagLim sample. These samples were chosen for their redshift overlap with the two simulated BOSS/eBOSS samples.

We evaluate the different steps by testing our methodology on each of the six samples. From the six results, we then fit for bias and systematic uncertainty (on top of statistical uncertainties for the correlation functions). We describe the evaluation step in more detail at the end of this section.

5.1 Test 1: testing methodology with spectroscopic redshifts

Our first test evaluates the accuracy of equation (7), the solution for $n(z)$ when using all spectroscopic measurements. In this scenario, autocorrelations of both the unknown and reference samples can be used to calibrate the impact of galaxy bias evolution with redshift across a photo-z bin. Because the DES samples are photometric, we can only carry out this measurement on the simulations. This test is still useful to isolate performance of the method, including galaxy bias corrections from autocorrelations, before evaluating any of the effects associated with photometric redshifts.

5.2 Test 2: approximating galaxy bias correction with a power law

In equation (12), we approximate $\bar{w}_{\rm uu}(z)$ (or proxies of it) as a power law: $(1+z)^{\gamma}$. In Test 2, we test any biases in fitting the autocorrelation to a power law. To isolate the effects of this approximation from the effects of photometric redshifts, we do this test on the autocorrelation of the simulated DES samples, when using spectroscopic redshifts. We note that in the following three tests (Tests 3, 4 and 5), the proxies for the autocorrelation, $\bar{w}_{\rm uu}(z)$, are also approximated as a power law.

5.3 Test 3: galaxy bias correction using photometric redshifts and redshift scatter model

In Test 3, we test how well equation (8) corrects for effects of photo-z scatter in the simulations. The integrals in equation (8), which describe how the redshift distribution changes when binned by photometric or spectroscopic redshift, can only be evaluated in

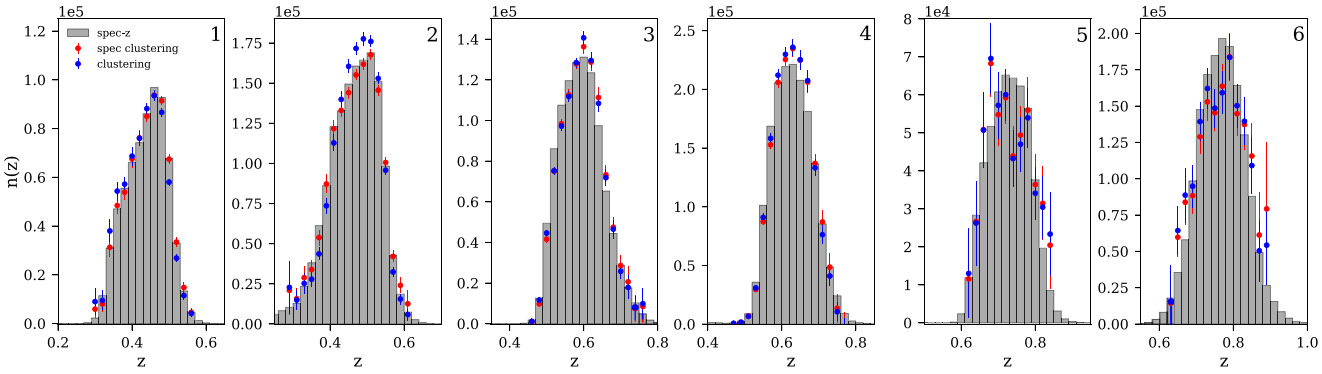

Figure 5. Clustering redshift estimates on simulated DES galaxy samples. The first, third and fifth samples are simulated redMaGiC redshift bins, and the others are simulated MagLim samples. The first four bins are cross-correlated with the simulated DMASS sample. The last two are correlated with the simulated ELG sample. Shown are the true (spec-z) redshift distributions, the clustering measurements that could be derived if both samples had spectroscopic redshifts ('spec clustering, Test 1'), and the clustering measurements described in 'Test 5', which is also the procedure done on the data. The simulated eBOSS sample runs out of galaxies around $z = 0.9$, limiting the higher-z bins shown here.

simulations with true redshift information. In the Year 1 DES results, Cawthon et al. (2018) used this calculation from simulations for the final correction to the clustering redshift results. We also fit equation (8) to a power law for this test.

5.4 Test 4: galaxy bias correction using standard deviation

In Test 4, we test the approximation of equations (10) and (11), using the standard deviation of a redshift distribution, σ_j, as an approximation in the photo-z correction of equation (8). In this test, we calculate σ_j for the DES samples exactly from their true redshifts in the simulation. Our approximation of $\bar{w}_{\mathrm{uu,spec}}(z_j)$ in this test is then $\bar{w}_{\mathrm{uu,pz}}(z_j)\sigma_j$. We again fit these new estimates of $\bar{w}_{\mathrm{uu,spec}}(z_j)$, for each micro-bin j within the photometric bin, i, to a power law.

5.5 Test 5: galaxy bias correction using standard deviation inferred from clustering

In Test 5, we test the last step in calculating a proxy for $\bar{w}_{\mathrm{uu}}(z_j)$. We again calculate σ_j, but this time from estimating the redshift distribution of the photo-z microbinned DES sample by using cross-correlations with BOSS/eBOSS samples on even smaller 'nano-bins' (Fig. 3). We estimate σ_j from these results and again test equations (10) and (11). Unlike Tests 3 and 4, Test 5 describes a measurement that can be done on the data, without true redshift information.

5.6 Summary of tests on photo-z bias

For each test in the simulations, we compute the clustering redshift measurements. We show in Fig. 5 for each of the six simulated samples: the true redshift distribution, the clustering estimate in Test 1 (using autocorrelations of DES samples using true redshifts), and the clustering estimate in Test 5, a procedure for calibrating the galaxy bias effects that can be done on the data.

We calculate the mean redshift in each case for the clustering redshifts. We summarize the accuracy of each test based on the measured mean redshift in Fig. 6. For each of our six simulated DES test samples, we plot the 'bias' of each step individually, each represented by a different test. For Test 1, the method, if spectroscopic redshifts are available for the unknown sample, the 'bias' is the true mean redshift of the photometric bin (calculated in the range where there are clustering estimates) minus the inferred mean redshift. For Tests 2–5, we plot as 'bias' the inferred mean redshift of Test 1 minus the inferred mean redshift of the test in question. We do this because Tests 2–5 all involve replacing w_{uu} in some way. Comparing these with Test 1 thus isolates the bias of the specific approximations of $w_{\mathrm{uu,\,spec}}$ compared with the case where $w_{\mathrm{uu,\,spec}}$ is actually available.

The method when spectroscopic redshifts are available (Test 1, 'Method') is shown to be generally accurate. Five of the six samples estimated mean redshifts that are consistent with the true mean redshift (i.e. zero bias) within the statistical error bar. The power-law approximation (Test 2) also shows little bias. Tests 3, 4 and 5, which all evaluate methods of modifying $w_{\mathrm{uu,\,pz}}$ to estimate $w_{\mathrm{uu,\,spec}}$, show a relatively small positive bias in all six samples, suggesting a true bias of this approximation. A positive bias means the inferred mean redshifts are too low. We also show the case of no correction, where no estimate of w_{uu} is used. We see that in each case, the attempted correction of Test 3, 4 or 5 does reduce some of the bias.

We now quantify these tests to add any necessary biases or systematic uncertainties of the method to our results later. A non-zero systematic uncertainty indicates that the overall uncertainty of the method is larger than the measured statistical uncertainty from the jackknife resamplings. For each test, we model its bias, b, and systematic uncertainty, ω with a two-parameter fit to the six data points in Fig. 6 and perform a chi-squared test to find the best-fitting parameters. Specifically, for each test, we calculate

$$\chi^2_{\mathrm{red}} = \sum_{x=1}^{6} \frac{(d_x - b)}{\sqrt{\sigma_x^2 + \omega^2}} / (\mathrm{dof} = 4) \tag{15}$$

where χ^2_{red} is the reduced chi-squared, x iterates over the six test samples, d_x is the measured bias on each sample for the given test, σ_x is the measured statistical error on those biases, and dof signifies our four degrees of freedom (six samples − two fit parameters). For a reduced chi-squared, a value of 1 represents a good fit, a value less than 1 indicates too good a fit (errors are overestimated) and a value greater than 1 indicates a poor fit (errors are underestimated).

We start by evaluating a large set of b parameters with $\omega = 0$. If any of the resulting reduced chi-square values are less than 1, this would indicate a good fit to a bias of the step, with no uncertainty needed to be added on top of the statistical uncertainties. For each of the five tests, no b values result in $\chi^2_{\mathrm{red}} < 1$ when $\omega = 0$. This indicates that each step tested has some systematic uncertainty

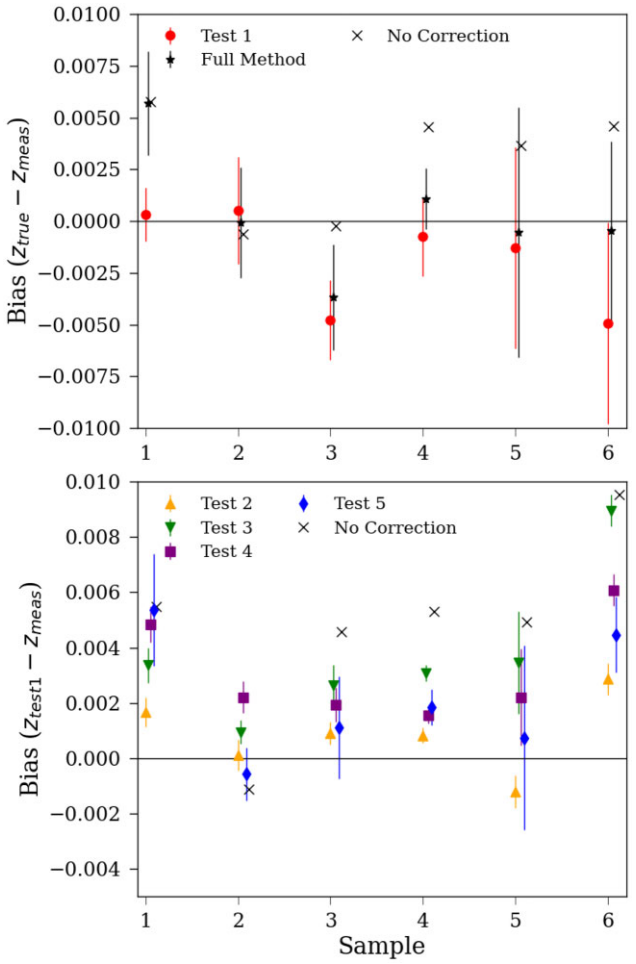

Figure 6. Results of testing different steps of the methodology on six different simulated samples. The tests are described throughout Section 5. Top: tests comparing the measured $z_{\rm mean}$ with the true value. Test 1 compares the method if spectroscopic redshifts were available for the photometric sample autocorrelations. The full method is the method we can do on the data. Bottom: Tests 2–5 focus on different correction steps or techniques for simulating the true (spectroscopic) photometric sample autocorrelation. The bias shown is with respect to the mean redshift from Test 1 in order to isolate bias due to the correction alone. Test 5 represents the full method that can be done on data. For each panel, the bias with 'no corrections' shows the results if no attempt is made to use or measure the photometric sample autocorrelation. The six test samples are in order of redshift. In order, they are simulated versions of: (1) redMaGiC, Bin 2; (2) MagLim, Bin 2; (3) redMaGiC, Bin 3; (4) MagLim, Bin 3; (5) redMaGiC, Bin 4; (6) MagLim, Bin 4. The simulated samples are described in Section 3.

beyond the statistical uncertainties from the autocorrelation and cross-correlation measurements. We then incrementally continue to calculate $\chi^2_{\rm red}$ for a range of ω and b values. We choose as our best fit the smallest value for ω, and the corresponding b that results in $\chi^2_{\rm red} < 1$. Our results are shown in Table 5. As suggested by Fig. 6, each of the biases are relatively small ($|b| \leq 0.0037$).

We choose to only use Tests 1 and 5 to estimate the bias and uncertainty we should add to our measurements of the data. Test 1 ('Method') essentially estimates biases across the full method presented in this work, modulo the complication of not having spectroscopic redshifts to estimate $w_{\rm uu,\,spec}$. Tests 3, 4 and 5 all measure the step of trying to estimate $w_{\rm uu,\,spec}$. Of these three, Test 5, which measures σ_j, the width of the true redshift distribution

Table 5. Analysis of the tests shown in Fig. 6, by fitting parameters in equation (15). The tests are described throughout Section 5. We note that the uncertainties represent uncertainty to be added to a step in the method. For example, Test 5 has more statistical uncertainty already in its step than Test 4, so the results do not imply Test 4 has more total uncertainty. As described in the text, we choose to incorporate the biases and uncertainties of Tests 1 and 5 so as not to double count uncertainties in any step of our methodology for solving for the redshift distribution, $n_{\rm u}(z)$. 'Combined errors' adds the counted biases and uncertainties (the latter in quadrature). The full method check is the bias and uncertainty found when not breaking up the analysis in different steps.

Name of test	Bias	Uncertainty	In error budget?
Test 1: method w/Spec-z	−0.0014	0.0013	Yes
Test 2: power-law approximation	0.0009	0.0015	No
Test 3: exact n^2 correction	0.0037	0.0030	No
Test 4: exact σ_j correction	0.0032	0.0020	No
Test 5: clustering-z σ_j correction	0.0021	0.0021	Yes
Combined errors	0.007	0.0025	–
Full method check	0.007	0.0023	–

of photometric galaxies in micro-bin j from clustering redshift measurements on 'nano-bins', is the only method that can be used solely from data. It happens to be that this method in the simulations (Test 5) is slightly more accurate than the other methods (Tests 3 and 4), which estimate the photometric correction either from an exact calculation of $\int n(z)^2$ or estimating σ_j exactly. Because Tests 3, 4 and 5 all approximate $w_{\rm uu,\,spec}$, adding all of their uncertainties would likely be redundant and overestimate our errors. Tests 3, 4 and 5 also involve a fit to a power law, so the small errors found in Test 2 are likely incorporated into the results of Test 5. Thus, deriving systematic biases and uncertainties from Tests 1 and 5 only incorporates each element of the measurements a single time.

From Table 5, we take the results from Tests 1 and 5. We add the biases linearly and the systematic uncertainties in quadrature. This results in adding a bias of +0.0007 and a systematic uncertainty of 0.0025 to each of our measurements. We see in Section 6 that these systematic errors are generally similar to or smaller than our inferred biases and statistical uncertainties on the mean redshift (or, equivalently, the Δz parameter).

We also show in Table 5 a calculation of a full method check. This calculation is simply taking our fiducial estimate of clustering redshifts (used in Test 5) and comparing with the true mean redshift, rather than the results of Test 1, which is what Test 5 does. In this case of the full method check, the intention is not to estimate a bias of a single step, but to estimate the bias of the entire method, end to end. We see in the table that the resulting bias and uncertainty are very similar to 'added' bias and uncertainty values of +0.0007 and 0.0025. We can thus conclude that whether we broke up the method into different steps or not would not have notably affected our derived bias and uncertainty.

In Appendix C, we briefly discuss a few alternative ways of evaluating these tests, as well as the assumption of the six samples each being independent tests of the method. We find very minor differences in bias and uncertainty in all cases explored.

5.7 Tests of two-parameter fits

In this section, we test the accuracy of the two-parameter fit (equation 14). The second parameter, the stretch, changes the width of the distribution. The galaxy bias correction, at least in the power-law form (equation 12), has very little impact on the stretch parameter, so our previous approach of breaking up the method into various steps is less well motivated. Therefore, we estimate the accuracy of

the two-parameter fit using just the 'full method' estimate, where the final Δz and s parameters from our fiducial clustering redshift estimate (equations 11 and 12) are compared to the true values. As seen in Table 5, this test produced nearly identical results for the one-parameter fit as when we evaluated different steps separately.

We test a few different ways of fitting the two-parameter model, each using the Δz and s parameters from equation (14). The first, as mentioned in Section 4.4, is a χ^2 fit to the clustering redshift data points, selecting Δz and s to change the photo-z distribution so as to be as close to the clustering redshifts as possible. The second method is a more natural extension of the one-parameter fit. It selects Δz in the same way as the one-parameter fit. Then, it selects the s parameter that makes the photo-z distribution have the same standard deviation, σ, as the clustering data. We call it the 'shift then stretch' (STS) method. This method fits to parameters, rather than fitting to all the points, like the one-parameter fit. It will also give the same Δz constraints as the one-parameter fit, while the χ^2 method may not. We also try a third method, which we call 'Mix'. This first computes the one-parameter shift (so also gives same results on Δz as STS). Then, with a fixed Δz, it does a χ^2 fit for the stretch.

To test the three methods of fitting, we compute Δz and s for each of the six simulated samples, given their simulated photo-z distributions. We also calculate the true Δz and s parameters that would make the photo-z distribution most closely match the true distribution. Given these estimated and true Δz and s parameters, we compute biases on each parameter, for each method on each sample. The results are shown in Fig. 7.

As for the one-parameter fit (photo-z bias only), we evaluate a 'method error' by comparing the estimated parameters with their true values. We again use equation (15), though we decide to use a more agnostic model of no bias, but just an added uncertainty (ω in equation 15). We chose this because, in Figs 5 and 7, there is some evidence of a directional bias to the stretch parameter based on the signal-to-noise ratio, with stretches (widths) being overestimated with noisy data (samples 5 and 6). In the first four bins, with better signal, there is a small preference for an opposite bias on the stretch, with a preference for narrower distributions than truth. Without more simulated samples to investigate these relationships, a no bias fit seemed most conservative, likely resulting in larger uncertainty than otherwise.

In Table 6, we show the estimated 'method uncertainty' of each parameter for each two-parameter method. These uncertainties are to be added to the statistical errors of a given method. As can also be inferred from Fig. 7, the STS (shift then stretch) method does better at finding the mean redshift, but is significantly less accurate than χ^2 in recovering the stretch parameter. The χ^2 method is the most accurate and has the smallest errors on the stretch parameter. The Mix method gives similar results to χ^2, but is still notably less accurate for the stretch parameter. It is more accurate than the STS method for the stretch though.

We note that for sample 6 (simulated MagLim bin 4), for only the two-parameter tests in Fig. 7, we used larger clustering scales, 0.5–4 Mpc, as we do for the noisier MagLim bins in the data. Without this, the stretch for sample 6 is significantly more negatively biased for each fit, and drives the method uncertainty for χ^2 to 0.07, which deviates significantly from the uncertainty of 0.035 when fitting the first five bins. This change in scales would affect all of the one-parameter results by less than 0.001.

In principle, any of these methods with the extra uncertainties from these tests should be unbiased. When we look at the overall error budget that the three methods give on measurements of data, we find that the STS and Mix methods have slightly smaller uncertainty on Δz, but notably larger uncertainty on the stretch (particularly STS). Based on these results, we proceed with the χ^2 method as our fiducial two-parameter fit method. We reiterate that each method is just a different way of fitting to the clustering redshift data, so the similar

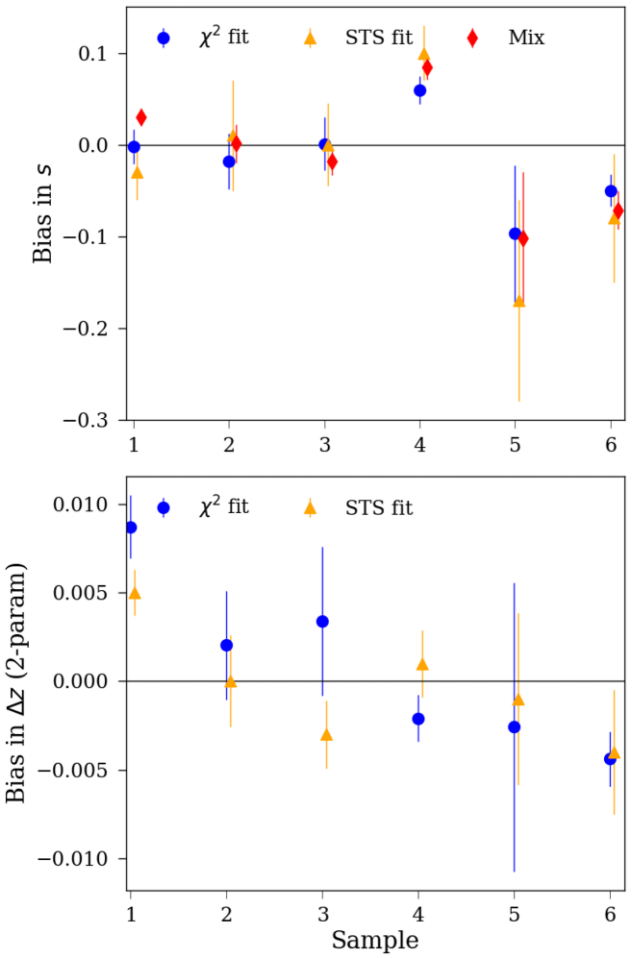

Figure 7. Results of testing the three different two-parameter fit methods on the six simulated samples: χ^2, STS (shift then stretch) and Mix (shift, then χ^2 fit for stretch). We show the biases in measured Δz and s parameters compared to the true values that make the photo-z distribution closest to the truth. As seen, the STS method has smaller errors and is more accurate in obtaining Δz correct, but has larger errors and is more biased on the stretch parameter, s. The χ^2 method is the most accurate and has the smallest errors on the stretch parameter. Mix gives similar results to χ^2 for the stretches, but is still slightly more inaccurate for that parameter. Mix has the same Δz estimate as STS. We fit for a 'method uncertainty' for each method, on each parameter, based on these results in Table 6.

Table 6. Analysis of the two-parameter methodology test results shown in Fig. 7. We compare the estimates of Δz and s with the true values in the simulation with equation (15) (setting $b = 0$ and just solving for an uncertainty, ω).

Method	Parameter	Uncertainty
χ^2	Δz	0.0044
χ^2	s	0.038
STS	Δz	0.0025
STS	s	0.060
Mix	s	0.052

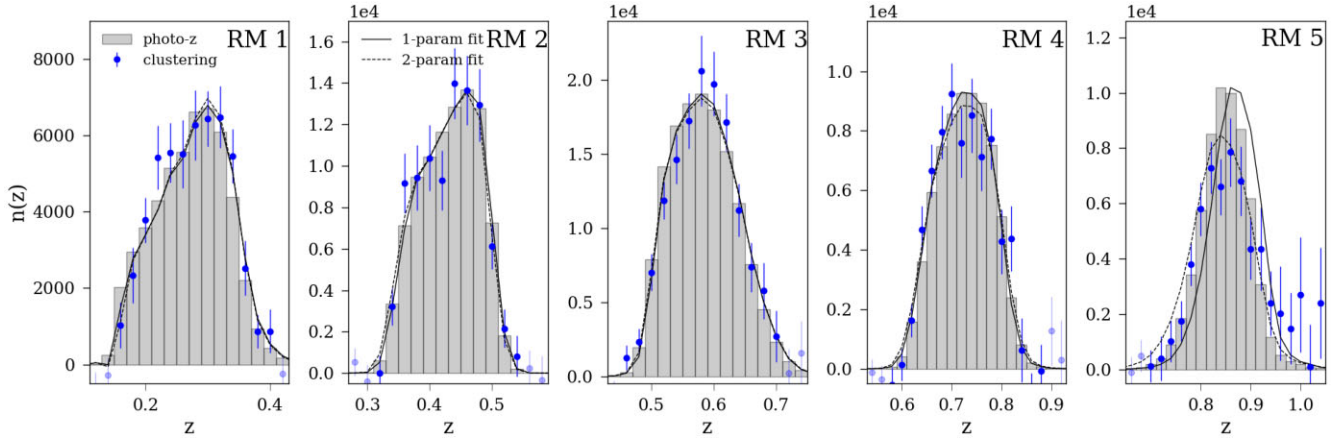

Figure 8. The clustering redshift measurements for the five redMaGiC bins. The photo-z prediction comes from the redMaGiC algorithm itself. The dark blue data points indicate the range where clustering and photo-z are compared in order to find the best-fitting shift, Δz. The grey points are outside this range and not used. Error bars only reflect statistical errors from the cross-correlation of DES and reference (BOSS/eBOSS) galaxies, and the autocorrelation of the reference galaxies. Best-fitting parameters are given in Tables 7 and 8.

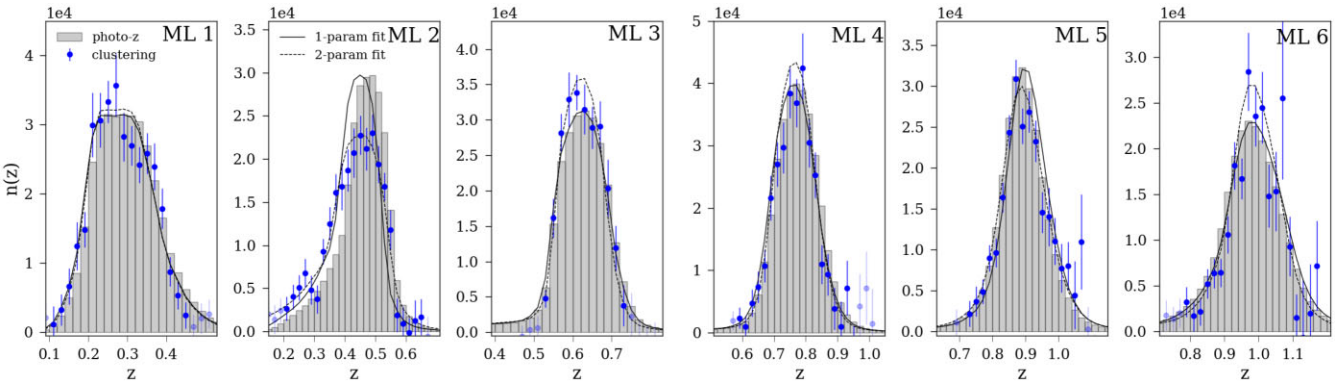

Figure 9. The clustering redshift measurements for the MagLim sample. The photo-z prediction comes from the summation of the entire DNF PDF for each galaxy. See Fig. 8 for more details about the data points. Best-fitting parameters are given in Tables 9 and 10.

biases in Fig. 7 are expected. We note that we also tried fitting for the parameters in Tables 5 and6 using a maximum likelihood formalism, and we found consistent results in each case.

In Appendix D, we describe tests to analyse possible density dependence of the systematic uncertainties derived from this section. We do find some indication of a density dependence that would result in the uncertainties presented here being underestimated for the data being used. We discuss these results and implications in Appendix D, finding, even in a pessimistic case, that the impact on cosmological results is minimal.

6 RESULTS

We show our clustering redshift estimates for the five redMaGiC and six MagLim redshift bins in Figs 8 and 9, where we also show the predicted redshift distributions from photometric redshift algorithms. For redMaGiC, this is provided by the redMaGiC algorithm itself. For MagLim, this is provided by the DNF photo-z algorithm, using its full probability distribution function (PDF). In each figure, the solid blue points indicate the 2.5σ range of the clustering signal which we use to calculate the mean redshift from clustering. The faded grey points are shown, but not used. For each bin, we compute the one-parameter fit, where we find Δz, a shift parameter to make the photometric and clustering distributions match in mean redshift

Table 7. redMaGiC clustering redshift results for a one-parameter fit. Δz is the shift that makes the photo-z prediction from redMaGiC match the mean of the clustering measurements. Statistical errors are from DES-reference cross-correlation and reference autocorrelations. Systematic errors are a combination of errors on the power-law fit to the DES autocorrelation (equation 12) and a 0.0025 method error from Section 5. The bias of +0.0007 from that section is also applied.

Redshift bin	Δz	χ^2 (points)
1: $z_{\rm ph} \in [0.15, 0.35]$	0.006 ± 0.004	6.81 (13)
2: $z_{\rm ph} \in [0.35, 0.5]$	0.001 ± 0.003	10.03 (11)
3: $z_{\rm ph} \in [0.5, 0.65]$	0.004 ± 0.003	7.32 (13)
4: $z_{\rm ph} \in [0.65, 0.8]$	-0.002 ± 0.005	19.92 (16)
5: $z_{\rm ph} \in [0.8, 0.9]$	0.020 ± 0.010	69.35 (18)

(equation 13). We also compute the two-parameter fit, where a χ^2 fit to the data simultaneously fits for a Δz and a stretch parameter, s. This two-parameter fit is also done on the 2.5σ range of the clustering signal.

The best one- and two-parameter fits and uncertainties are listed in Tables 7–10. Also shown is the χ^2 value between the fit and the clustering redshift data points. The listed uncertainties in the fits, as well as the error bars in the figures, include contributions from statistical and systematic uncertainties. The statistical uncertainties come from

Table 8. redMaGiC clustering redshift results for a two-parameter fit. Δz is the shift parameter and s is the stretch parameter. Each is fit by changing the photo-z distribution to match the clustering data points with a χ^2 fit. Uncertainty comes from statistical uncertainty of the fit, which includes contributions from the power-law fit uncertainty (equation 12), and the 'method error' in Section 5.6.

Redshift bin	Δz	s	χ^2 (points)
1: $z_{\rm ph} \in [0.15, 0.35]$	0.007 ± 0.005	0.975 ± 0.043	5.90 (13)
2: $z_{\rm ph} \in [0.35, 0.5]$	-0.002 ± 0.005	1.015 ± 0.045	7.14 (12)
3: $z_{\rm ph} \in [0.5, 0.65]$	0.003 ± 0.005	1.017 ± 0.048	6.99 (11)
4: $z_{\rm ph} \in [0.65, 0.8]$	-0.002 ± 0.006	1.051 ± 0.065	17.54 (16)
5: $z_{\rm ph} \in [0.8, 0.9]$	-0.007 ± 0.006	1.230 ± 0.066	16.74 (18)

Table 9. MagLim sample clustering redshift results for a one-parameter fit. Δz is the shift that makes the photo-z prediction from DNF match the mean of the clustering measurements. See Table 7 for further comments on uncertainty column sources.

Redshift bin	Δz	χ^2 (points)
1: $z_{\rm ph} \in [0.2, 0.4]$	-0.010 ± 0.004	23.96 (18)
2: $z_{\rm ph} \in [0.4, 0.55]$	-0.028 ± 0.006	91.91 (23)
3: $z_{\rm ph} \in [0.55, 0.7]$	-0.004 ± 0.004	12.64 (11)
4: $z_{\rm ph} \in [0.7, 0.85]$	-0.010 ± 0.005	22.61 (19)
5: $z_{\rm ph} \in [0.85, 0.95]$	0.013 ± 0.007	44.34 (18)
6: $z_{\rm ph} \in [0.95, 1.05]$	0.009 ± 0.016	36.61 (20)

Table 10. MagLim clustering redshift results for a two-parameter fit. Δz is the shift parameter, and s is the stretch parameter. See Table 8 for further comments on the fit and uncertainties. A typo in bin 6 from the original version of this paper has been corrected.

Redshift bin	Δz	s	χ^2 (points)
1: $z_{\rm ph} \in [0.2, 0.4]$	-0.009 ± 0.007	0.975 ± 0.062	24.27 (18)
2: $z_{\rm ph} \in [0.4, 0.55]$	-0.035 ± 0.011	1.306 ± 0.093	22.94 (23)
3: $z_{\rm ph} \in [0.55, 0.7]$	-0.005 ± 0.006	0.870 ± 0.054	6.55 (11)
4: $z_{\rm ph} \in [0.7, 0.85]$	-0.007 ± 0.006	0.918 ± 0.051	21.96 (19)
5: $z_{\rm ph} \in [0.85, 0.95]$	0.002 ± 0.007	1.080 ± 0.067	25.79 (18)
6: $z_{\rm ph} \in [0.95, 1.05]$	0.009 ± 0.008	0.845 ± 0.073	36.59 (20)

the cross-correlation of unknown and reference samples ($\bar{w}_{\rm ur}$) and autocorrelations of the reference sample ($\bar{w}_{\rm rr}$) in equation (11). For the one-parameter fits, the systematic uncertainty is calculated on the derived mean redshift specifically and has two sources. The 'method uncertainty' of 0.0025 derived from Section 5, and uncertainty in the calculations of the autocorrelations of the unknown (DES) sample and the calculation of the standard deviation parameter ($\bar{w}_{\rm uu,pz}\sigma_j$). This uncertainty is propagated into an uncertainty on γ in fitting that quantity to a power law (equation 12). The autocorrelations and the power law fits are shown in Fig. 10. These two sources of systematic uncertainty are added in quadrature. The total uncertainty comes from adding the systematic and statistical uncertainty in quadrature.

For the two-parameter fit, the uncertainty from the power law is propagated into the data points before the fit is performed, so the statistical error incorporates the power-law fit. The only remaining systematic uncertainty is the 'method' uncertainty from Table 6, which is added in quadrature to each parameter's statistical uncertainties from the χ^2 fit. The exact contributions of statistical and systematic uncertainty to each fit in each redMaGiC and MagLim bin are shown in Tables B1 and B2 in Appendix B.

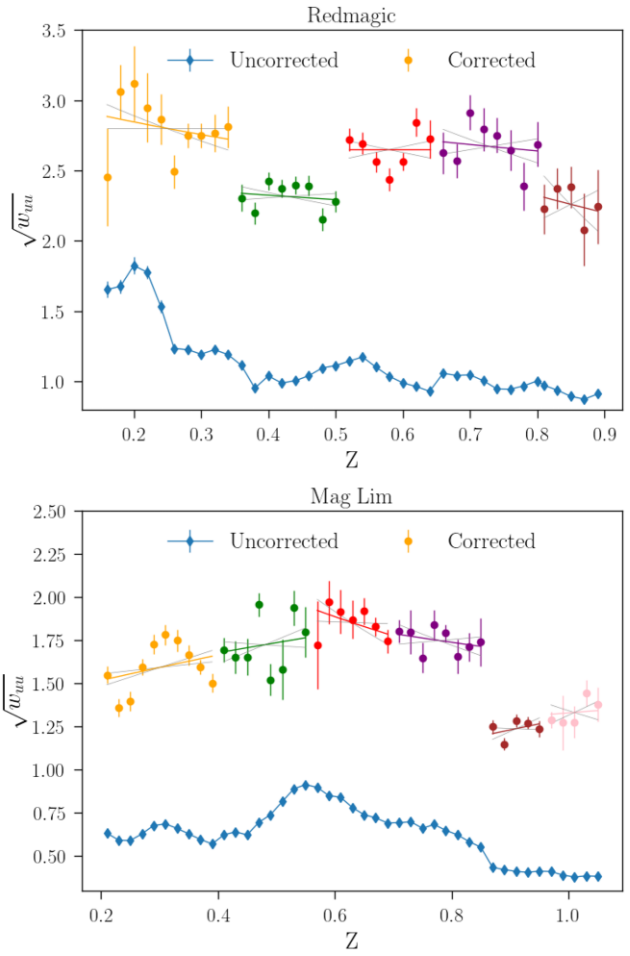

Figure 10. Autocorrelations of the DES lens samples, redMaGiC and MagLim. We show the square root of the autocorrelation, $\sqrt{\bar{w}_{\rm uu,pz}(z_j)}$, listed as uncorrected. The corrected data points are $\sqrt{\bar{w}_{\rm uu,pz}(z_j)\sigma_j}$, to undo the influence of photo-z scatter on the autocorrelations. The 11 different tomographic bins are differentiated by colour. The solid lines in colour are the best power-law fit to the points in a given bin (equation 12). The grey lines indicate the 1σ range of the power-law fit. We note that, because of the normalization of $n(z)$, the amplitudes of these measurements are not important, only the change across a single tomographic bin

6.1 redMaGiC results

Our clustering redshift estimates for the redMaGiC sample are shown in Fig. 8 and in Tables 7 and 8. In the first four bins of Fig. 8, we see a generally good agreement in the shapes and means of the clustering and photometric redshift distributions. The first four bins have relatively small biases ($\leq$0.006), and are within 1.5 standard deviations of zero bias for the one-parameter fit. The fifth bin has a more obvious difference in shape between the clustering and photo-z distributions, with a high-z tail in the clustering redshift distribution. Driven by this tail, a large shift parameter of 0.02 is needed to match the means of the distributions in the one-parameter fit.

There are different possible metrics in selecting whether the one- or two-parameter model should be used for the DES cosmology analysis. In a cosmology analysis, every extra parameter allowed to vary will typically reduce the constraining power of the experiment. Thus, there is a benefit to having a simpler model if it is accurate enough. One possible metric for deciding whether a one- or two-parameter model is sufficient is whether the two-parameter model

prefers $s \neq 1$. In Table 8, we see that the first four bins are all well fit by $s = 1$, suggesting that a two-parameter fit may be unnecessary. Bin 5, in contrast, prefers $s \neq 1$ at $>3\sigma$ confidence.

Another metric could be to use the goodness of fit, which can be assessed with the χ^2 values listed in Tables 7 and 8. We see that the redMaGiC one-parameter fits in the first four bins are all close to a reduced χ^2 of 1. As expected, the fifth bin exhibits a very poor fit. We do see lower χ^2 values in the two-parameter fits for the first four bins, but at the cost of a second parameter (s), and increased errors on Δz. In the fifth bin though, the two-parameter fit is much better, with a reduced χ^2 near 1. We note that though the fit may not look that good in Fig. 8, the off-diagonal terms of the covariance between the $n(z)$ data points have a strong effect on this particular fit.

For the cosmology analysis in DES Collaboration et al. (2022), the general strategy is to add complexity to models only if it is necessary not to bias the cosmological results. With this strategy in mind, we show another test in assessing whether the one- or two-parameter models are needed in Appendix A. There, we show Markov chain Monte Carlo (MCMC) chains from a simulated cosmology analysis approximating the analysis in DES Collaboration et al. (2022). Chains are run with fixed cosmology but different redshift inputs, to assess whether the redshift modelling is sufficient not to bias cosmological results.

In Fig. A1 in Appendix A, we show simulated results for Ω_m, σ_8 and the galaxy bias in each of the five redMaGiC bins for four different redshift inputs. The different inputs are: the clustering redshift results directly, a multi-Gaussian fit very closely matching the clustering redshifts but smoother, and the one- and two-parameter fits listed in Tables 7–10. We find in this test that the cosmological parameters are similar in all cases. However, the galaxy bias recovered in the fifth redMaGiC bin is offset from the more direct clustering fits. This suggests that the poor fit of the one-parameter model in the fifth bin will give biased results for the galaxy bias if the clustering redshift results are accurate.

Based on these two tests, the one-parameter fits for the first four REDMAGIC bins and the two-parameter fit for the fifth bin were chosen as the fiducial models for DES Collaboration et al. (2022). In order to be conservative with the fifth bin, the uncertainty on Δz was increased from 0.007 to 0.010, to match the uncertainty from the one-parameter fit. We investigate the fifth redMaGiC bin clustering results in more detail in Section 7, specifically checking whether the high-z tail could be explained by magnification. As shown there, we conclude it cannot be. We also note that we measured redMaGiC bin 5 at the larger clustering scale range, 0.5–4 Mpc, which we use for two of the MagLim bins. With these scales, this high-z tail for redMaGiC bin 5 remains, with marginally smaller error bars.

6.2 MagLim sample results

Our clustering redshift estimates for the MagLim sample are shown in Fig. 9 and in Tables 9 and 10. We can see that there are significantly different shapes of the clustering and photo-z distributions in multiple bins. It was expected that this larger sample of fainter galaxies would have larger photo-z biases than redMaGiC.

It is again important to analyse whether a one- or two-parameter fit will be more appropriate for the cosmology analyses. For the metric of checking whether the two-parameter fit is consistent with $s = 1$, we find in Table 10 only one bin where this is clearly the case (bin 1). In examining the χ^2 fits for the one- and two-parameter fits in Tables 9 and 10, we see a very strong preference for the two-parameter fit in bins 2 and 5, though it is less clear if the two-parameter fit is necessary for the other bins by this metric.

The final decision on whether one or two parameters are needed for the MagLim fits is determined by the procedures in DES Collaboration et al. (2022) and Krause et al. (2021), which focus on whether there will be biases in the cosmological analysis if the simpler model is used. These tests for the MagLim redshifts, analogous to the redMaGiC tests in Appendix A, are shown in Porredon et al. (2021a, fig. 6). There it is shown that the galaxy bias in MagLim bins 2–6 using the one-parameter fits of (this paper's) Table 9 are offset from what the estimates would be when inputting the clustering redshift measurements directly. This would mean that the one-parameter fits may bias the galaxy bias measurements. Thus, it was decided to use two-parameter fits for all of the MagLim bins.

We note that our initial measurements of the two highest redshift bins for the MagLim sample were very noisy, mainly due to the low number of eBOSS objects available to correlate with. Our final analysis for bins 5 and 6, shown in Fig. 9, used the clustering redshift method on scales 0.5–4.0 Mpc, while the rest of this work used 0.5–1.5 Mpc. At these noisier high redshifts, we found that the inclusion of more scales improved the signal significantly, reducing total uncertainty on Δz by about 50 per cent and 80 per cent for bins 5 and 6, respectively (for the one-parameter fit). We tested other bins at these scale ranges and found negligible differences in other cases. As we weight the smallest scales most, where more information is expected, this lack of change in most cases is unsurprising.

We also checked the $z = 1.06-1.08$ clustering data point, which stands out clearly in bin 5 and slightly in bin 6. An isolated peak such as this appears strange and perhaps anomalous. We tested several things to try to find an issue with the clustering data, but could find none. It was not found in redMaGiC when we extended the redshift range. It was not found in autocorrelations, or cross-correlations of the eBOSS ELG and QSO samples with each other. It was found in cross-correlations of MagLim with either the ELG or QSO alone. Changing the binning, changing the number of jackknife patches or splitting up the data into different large regions did not remove the signal. We also tried not using either MagLim or eBOSS weights, and also estimators that would not use one of the sample's randoms. In all cases, the signal persisted. The signal is also much too large for magnification, which should not produce a sharp change in redshift anyway. We conclude that the signal is real, either some statistical fluctuation or some interloping population of galaxies that entered the MagLim cuts.

6.3 Comments on redshift impacts to cosmological analysis

The redshift parameters and choices described in this section for both redMaGiC and MagLim were made before unblinding the cosmological results shown in DES Collaboration et al. (2022). In assessing those results post-unblinding, there were some redshift-related tests worth noting.

One of the noteworthy issues discussed in DES Collaboration et al. (2022) is the apparent disagreement of the redMaGiC galaxy clustering plus galaxy–galaxy lensing results and the cosmic shear results. This is discussed in detail in DES Collaboration et al. (2022). One of the first tests done was to see if using the two-parameter fits from Table 8 in all redMaGiC redshift bins could alleviate this issue. It did not. The resulting chains with the two-parameter redshift model had only slightly larger contours and there was still significant inconsistency in the data.

Another significant test for understanding the redMaGiC inconsistency was dropping different redshift bins. Most notably for this work, the fifth redMaGiC bin was shown to have little impact on the cosmological results and the inconsistency. Therefore, though the

fifth redMaGiC bin is clearly the one with the greatest redshift uncertainties, it cannot be driving the inconsistencies in the redMaGiC results.

For the MagLim sample, a similar issue of significant mismatch between the galaxy–galaxy lensing + galaxy clustering amplitude and cosmic shear amplitude was found in bins 5 and 6, as discussed in Porredon et al. (2021a). The measurements in the first four bins were internally consistent though and were used for the fiducial results in DES Collaboration et al. (2022). It is unclear at this time what the issue in these high-redshift bins is. In Porredon et al. (2021a, appendix A), tests seemed to indicate that the problem is more with galaxy–galaxy lensing than with galaxy clustering. Based on this, it is unlikely errors in the lens redshifts are driving this tension. The galaxy clustering measurements will depend much more on the lens redshifts, particularly the width parameter.

7 MAGNIFICATION

In this section, we calculate whether magnification may be affecting our results significantly. Clustering redshifts are known to be affected by magnification effects (Choi et al. 2016; Gatti et al. 2018), which become significant in the tails of the redshift distribution where the normal clustering signal is small. Our cut of the tails at 2.5σ from the peak of the clustering $n(z)$ should remove most of the redshift range where magnification effects are significant in each bin. However, some of our results do include a fairly large signal in the tails. Notably, the high-z tail in redMaGiC bin 5 is large enough to be within this cut and not removed.

We calculate a theory prediction for magnification in our clustering redshift measurements. Specifically, we calculate the theoretical signal in the redMaGiC bin 5 measurements to assess whether the high-z tail is magnification-induced, or is real evidence for a photo-z bias. We first calculate the strength of the galaxy clustering signal between the two samples (Choi et al. 2016):

$$w(\theta)^{\rm gg}_{\rm ur} = b_{\rm u} b_{\rm r} \int n_{\rm u}[z(\chi)] n_{\rm r}[z(\chi)] \frac{{\rm d}z}{{\rm d}\chi} \frac{{\rm d}z}{{\rm d}\chi} {\rm d}\chi \times \int \frac{k}{2\pi} P[k, z(\chi)] J_0(\chi k\theta)\, {\rm d}k. \quad (16)$$

Here, $\theta = r/\chi(z)$, as described in Section 4, b is the galaxy bias for the unknown (DES) and reference (eBOSS) samples, n is the redshift distribution of each sample, χ is the comoving distance, k is the wavenumber, $P(k)$ is the matter power spectrum and J_0 is the zeroth-order Bessel function.

We also calculate the strength of the magnification signal, specifically the signal from foreground redMaGiC galaxies lensing eBOSS galaxies (the magnification effect that could be found on the high-z end):

$$w(\theta)^{{\rm g}\mu}_{\rm ur} = b_{\rm u}(\alpha - 1) \int n_{\rm u}(z) K(\chi) \frac{{\rm d}z}{{\rm d}\chi} {\rm d}\chi \times \int \frac{k}{2\pi} P[k, z(\chi)] J_0(\chi k\theta)\, {\rm d}k. \quad (17)$$

$K(\chi)$ is the lensing kernel,

$$K(\chi) = \frac{3H_0^2\Omega_{\rm m}}{c^2} \frac{\chi}{a} \int_\chi^\infty n_{\rm r}(z) \frac{{\rm d}z}{{\rm d}\chi} \frac{\chi' - \chi}{\chi'} {\rm d}\chi', \quad (18)$$

where H_0 is the Hubble constant, $\Omega_{\rm m}$ is the matter density parameter, c is the speed of light and $a = 1/(1 + z)$ is the scalefactor of the Universe. In equation (17), α is the slope of the magnitude counts for, in this case, the eBOSS sample,

$$\alpha(m) = 2.5 \frac{{\rm d}\log_{10} n(> m)}{{\rm d}m}, \quad (19)$$

where m is the limiting magnitude of the galaxy sample (Choi et al. 2016). For eBOSS ELG (the main tracer used for the redshift range of redMaGiC bin 5), we calculate a value of $\alpha = 2.71$ at its limiting g-band magnitude, $m = 22.825$ mag (Raichoor et al. 2017).

We ignore the terms where eBOSS galaxies could magnify background redMaGiC galaxies. As the eBOSS galaxies have spectroscopic redshifts, we know the galaxies being correlated in the high-z tail are at $z \sim 1$. Thus, for eBOSS galaxies to be magnifying redMaGiC galaxies, there would need to be a significant number of very high redshift outliers in the redMaGiC population. However, redMaGiC galaxies around $z \sim 0.8$ in this fifth bin could plausibly magnify eBOSS galaxies at $z = 1$ where the excess is seen. We also ignore the magnification–magnification term, which should be negligible, particularly for two galaxy samples not widely separated in redshift space (for example, see Duncan et al. 2014). To do the calculations, we assume for $n_{\rm u}(z)$, the photometric redshift distribution estimate as an input, and we use *Planck* 2015 flat-ΛCDM cosmological parameters including external data (Planck Collaboration XIII 2016). These parameters include $H_0 = 67.74$ km s^{-1} Mpc^{-1} and $\Omega_{\rm m} = 0.3089$. The power spectrum $P(k)$ is calculated using the Boltzmann code in CAMB (Lewis, Challinor & Lasenby 2000, Howlett et al. 2012) with HALOFIT (Smith et al. 2003) used to calculate non-linear clustering effects. Minor deviations on the input redshift distribution or cosmology do not change the results qualitatively. Because the galaxy-clustering and galaxy-magnification equations both contain a factor of $b_{\rm u}$, it effectively cancels in a normalized clustering redshift calculation, so we set it to 1. We ignore the galaxy bias evolution, which will only produce small changes. We also set $b_{\rm r} = 1$, which is close to other studies of eBOSS ELGs that indicate $b_{\rm r} = 1.3$ (Guo et al. 2019).

We show the results of our theory calculations for the clustering–clustering and clustering–magnification terms for redMaGiC bin 5 as well as the measurements in Fig. 11. The results show that the magnification signal is far too small to explain the measured high-z excess in the bin. We show that the magnification term would need roughly an increase of a factor of 10 to match the data, so small errors in, for example, the α calculated in equation (19) could not explain the excess.

We perform similar tests on other bins and consistently see a theoretical magnification signal that is negligible to our results. This is mainly due to two factors. The first is the relatively narrow redshift bins used, reducing the amplitude of the lensing kernel (equation 18) between the DES galaxies and eBOSS galaxies. For example, with wider bins, one could have $z \sim 0.6$ DES galaxies lensing $z = 1$ eBOSS galaxies, which would have a larger lensing kernel than $z = 0.8$ DES galaxies. The second factor is our procedure of cutting the tails where the signal is small, as described in Section 4.3. This cuts out the regions with both highest magnification signal and lowest clustering signal.

We note that there is also other evidence that the high-z excess in redMaGiC bin 5 is a true photo-z bias. In the DES/eBOSS overlap area (mostly overlapping the region known as Stripe 82), about 3 per cent of the redMaGiC bin 5 galaxies have spectroscopic redshifts. In this subsample, a similar high-z excess in the spectroscopic $n(z)$ is seen compared with the photometric estimate. We extrapolate a prediction for $n(z)$ based on this spectroscopic subsample in Fig. 12. For this, we take the photometric $n(z)$ for the whole Stripe 82 region, and multiply that by a correction factor for each micro-bin (dz =

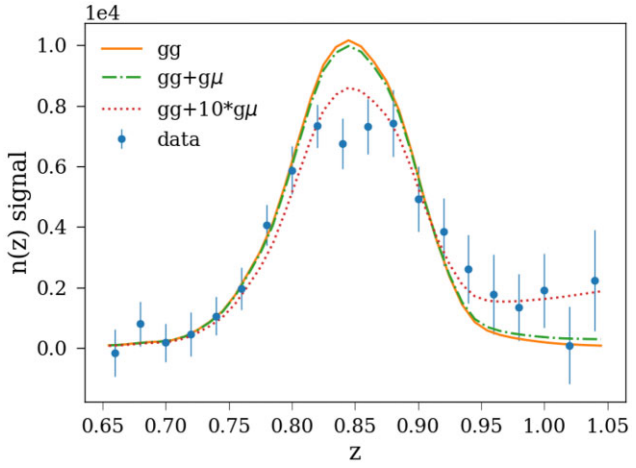

Figure 11. Estimates of the $n(z)$ signal for redMaGiC bin 5 ($z_{\rm ph} \in [0.8, 0.9]$). For the purpose of comparison, the clustering data points are only from the cross-correlation of redMaGiC and eBOSS, making them slightly different from Fig. 8. The solid (orange, gg) line shows the theoretical prediction for this cross-correlation due to clustering alone (equation 16). The dash-dotted (green, gg+$g\mu$) line is the theoretical prediction from galaxy clustering and correlations between redMaGiC galaxies and magnification effects on eBOSS galaxies (equations 16 and 17). The dotted (red, gg+10∗gμ) line is a theory prediction with galaxy clustering and a ten times larger amplitude prediction from the galaxy-magnification signal. Each of the theory predictions uses the photometric $n(z)$ prediction for the bin as input. Effects of magnification do not seem to be large enough in theory to account for the excess signal at high redshifts in this bin. There is no theoretical motivation for a factor of 10 error in magnification predictions.

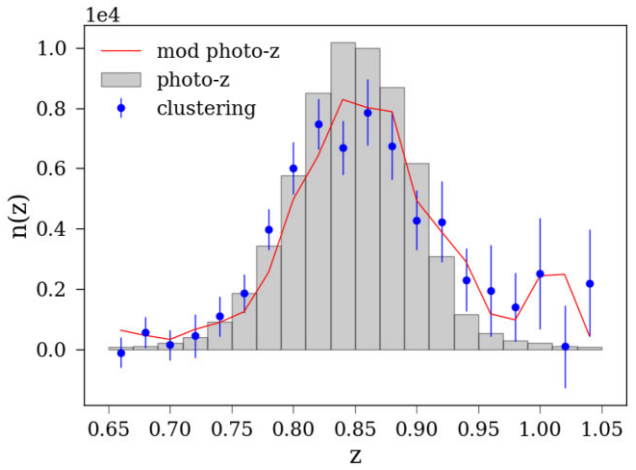

Figure 12. The solid (red) line shows the extrapolated $n(z)$ from spectroscopic redshifts in redMaGiC bin 5. The extrapolation is based on the ratio of galaxies in each d$z = 0.02$ bin by spec-z and by 'ZMC', a draw from the redMaGiC redshift distribution function in the subset of galaxies with spec-z measurements. This ratio is then applied to the full redMaGiC sample (the portion overlapping eBOSS) to give this extrapolated prediction. We see it matches the clustering results well, giving some evidence that the high-z tail is physical, rather than a clustering systematic.

0.02 in size) where the correction factor is $n_{\rm spec}/n_{\rm pz}$ as measured in the 3 per cent subsample with spectroscopic redshifts. We see that there is good agreement with the clustering redshift signal.

We caution that a comparison of photo-z errors on a sample with spectroscopic redshifts with a full sample is not generally reliable. Mismatches are seen when applying this procedure to other bins. This is generally because the subsamples with spectroscopic measurements tend to be brighter than the full samples, and will typically have smaller photo-z errors. As this particular error is seen in the brighter subsample of bin 5 though, it is more likely to be present in the full sample than the opposite case (i.e. assuming that a lack of error in the bright sample extrapolates well to the full sample). Between this minor evidence and the magnification calculations, we conclude that the high-z excess in the last redMaGiC bin is likely real. We also note that in early versions of the Year 3 redMaGiC catalogue, we saw more significant biases when the fifth bin extended to $z = 0.95$, also suggesting that the redMaGiC algorithm is encountering issues at high redshift.

8 SUMMARY

In this work, we present clustering redshift measurements of two DES lens samples, redMaGiC and MagLim (Figs 8 and 9). These measurements inform the redshift models for the analyses in DES Collaboration et al. (2022) and related papers. Our results are bolstered by the large number of spectroscopic galaxies available for this measurement from the BOSS and eBOSS galaxy clustering catalogues, and their several hundred deg^2 overlap with the DES. We generally find small biases ($|\Delta| < 0.01$) for the photometric redshift predictions of these samples (Tables 7–10). Our results suggest that the shapes of the redMaGiC photo-z distributions in particular are very accurate. The fainter, larger MagLim galaxy sample had more significant differences in shape when comparing the photo-z and clustering distributions, suggesting the need for a two-parameter fit for calibration. We were able to constrain the mean redshifts, in the form of the bias parameter Δz, of the different bins to a precision of typically around 0.005 when performing one-parameter fits. Our uncertainties on the mean redshifts were only marginally larger when performing two-parameter fits.

We tested our methodology in simulations (Section 5), including a new method of calibrating the galaxy bias systematic of the 'unknown' sample, the DES galaxies. This new method, involving cross-correlations on smaller redshift ranges, is made possible by the large number of spectroscopic tracers we had compared with previous work. We found the systematic errors on our method, beyond the statistical errors we account for in all of the autocorrelations and cross-correlations, to be small for calculating a single photo-z bias (Fig. 6 and Table 5). We also tested our ability to constrain the width of a redshift distribution. In this case, we found more significant 'method' errors not accounted for by the initial statistics. Despite this, our results suggest we can constrain the width to around 4–7 per cent for most samples. In Appendix D, we show some tests that suggest that the density dependence of the method may increase these uncertainties up to around 11 per cent, but this would have negligible impact on the DES cosmological results.

In one redMaGiC bin, we investigate specifically if a high-redshift tail in our results can be explained by magnification, and we find that it cannot. Our procedure of cutting the tails of the clustering redshift distribution, and the relatively narrow redshift bins for the DES lens samples in general, make magnification a negligible effect for our work.

For that particular redMaGiC bin in Appendix A, we also show that a two-parameter fit to the clustering data should not bias our cosmological results (while a one-parameter fit could at least bias the galaxy bias estimates). A similar analysis for MagLim is shown in Porredon et al. (2021a), where it is found that most of the MagLim bins need a two-parameter fit as well not biasing cosmological results.

These tests show the importance of using a multiparameter fit in calibrating photometric redshifts.

Our results provide important redshift constraints for our companion DES Year 3 results papers using redMaGiC and MagLim galaxies for galaxy clustering and galaxy–galaxy lensing (DES Collaboration et al. 2022; Pandey et al. 2021; Porredon et al. 2021a; Rodríguez-Monroy et al. 2022; Prat et al. 2021; Elvin-Poole et al., in preparation). This work also signifies some important steps in the progression of clustering redshift measurements, particularly in testing new methods to correct the galaxy bias systematic, and to constrain a width parameter of a redshift distribution. Future photometric and spectroscopic surveys will need to continue to develop the clustering redshift technique as they push to higher redshifts.

ACKNOWLEDGEMENTS

RC and KB acknowledge support from the US Department of Energy, Office of Science, Office of High Energy Physics, under Award Numbers DE-SC0020278 and DE-SC0017647.

Funding for the DES Projects has been provided by the US Department of Energy, the US 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 at NSF's NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed 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 program 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 Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the US Department of Energy, Office of Science, Office of High Energy Physics.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the US Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian (CfA), the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University and Yale University.

DATA AVAILABILITY

The full Year 3 redMaGiC and MagLim catalogues are available at https://des.ncsa.illinois.edu/releases. The subsets of these catalogues used in this paper will be made available upon reasonable request to the corresponding author.

REFERENCES

Ahumada R. et al., 2020, ApJS, 249, 3
Aihara H. et al., 2018, PASJ, 70, S4
Alam S. et al., 2015, ApJS, 219, 12
Alam S. et al., 2017, MNRAS, 470, 2617
Asgari M. et al., 2021, A&A, 645, A104
Blanton M. R. et al., 2017, AJ, 154, 28
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cawthon R., 2020, Phys. Rev. D, 101, 063509
Cawthon R. et al., 2018, MNRAS, 481, 2427
Choi A. et al., 2016, MNRAS, 463, 3737
Davis C. et al., 2017, preprint (arXiv:1710.02517)
Davis C. et al., 2018, MNRAS, 477, 2196
Dawson K. S. et al., 2013, AJ, 145, 10
Dawson K. S. et al., 2016, AJ, 151, 44
de Jong J. T. A., Verdoes Kleijn G. A., Kuijken K. H., Valentijn E. A., 2013, Experimental Astronomy, 35, 25

De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, MNRAS, 459, 3078
DeRose J. et al., 2019, preprint (arXiv:1901.02401)
DeRose J. et al., 2021, preprint (arXiv:2105.13547)
DES Collaboration, 2005, preprint (arXiv:astro-ph/0510346)
DES Collaboration, 2018, Phys. Rev. D, 98, 043526
DES Collaboration et al., 2022, Phys. Rev. D, 105, 023520
Dey A. et al., 2019, AJ, 157, 168
Duncan C. A. J., Joachimi B., Heavens A. F., Heymans C., Hildebrandt H., 2014, MNRAS, 437, 2471
Eisenstein D. J. et al., 2011, AJ, 142, 72
Elvin-Poole J. et al., 2018, Phys. Rev. D, 98, 042006
Gatti M. et al., 2018, MNRAS, 477, 1664
Gatti M. et al., 2022, MNRAS, 510, 1223
Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148
Gunn J. E. et al., 2006, AJ, 131, 2332
Guo H. et al., 2019, ApJ, 871, 147
Hildebrandt H. et al., 2021, A&A, 647, A124
Howlett C., Lewis A., Hall A., Challinor A., 2012, J. Cosmology Astropart. Phys., 2012, 027
Hoyle B. et al., 2018, MNRAS, 478, 592
Ivezić Ž. et al., 2019, ApJ, 873, 111
Johnson A. et al., 2017, MNRAS, 465, 4118
Krause E. et al., 2021, preprint (arXiv:2105.13548)
Krolewski A., Ferraro S., Schlafly E. F., White M., 2020, J. Cosmology Astropart. Phys., 2020, 047
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lee S. et al., 2019, MNRAS, 489, 2887
Lehmann B. V., Mao Y.-Y., Becker M. R., Skillman S. W., Wechsler R. H., 2017, ApJ, 834, 37
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
Matthews D. J., Newman J. A., 2010, ApJ, 721, 456
McQuinn M., White M., 2013, MNRAS, 433, 2857
Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint (arXiv:1303.4722)
Myers A. D. et al., 2015, ApJS, 221, 27
Newman J. A., 2008, ApJ, 684, 88
Pandey S. et al., 2021, preprint (arXiv:2105.13545)
Phillipps S., Shanks T., 1987, MNRAS, 227, 115
Planck Collaboration XIII, 2016, A&A, 594, A13
Porredon A. et al., 2021a, preprint (arXiv:2105.13546)
Porredon A. et al., 2021b, Phys. Rev. D, 103, 043503
Prakash A. et al., 2016, ApJS, 224, 34
Prat J. et al., 2022, Phys. Rev. D., 105, 083528
Raichoor A. et al., 2017, MNRAS, 471, 3955
Raichoor A. et al., 2021, MNRAS, 500, 3254
Reid B. et al., 2016, MNRAS, 455, 1553
Rivera J. D., Moraes B., Merson A. I., Jouvel S., Abdalla F. B., Abdalla M. C. B., 2018, MNRAS, 477, 4330
Rodríguez-Monroy M. et al., 2022, MNRAS, 511, 2665
Ross A. J. et al., 2020, MNRAS, 498, 2354
Rozo E. et al., 2016, MNRAS, 461, 1431
Rykoff E. S. et al., 2014, ApJ, 785, 104
Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, MNRAS, 431, 3307
Scottez V. et al., 2016, MNRAS, 462, 1683
Seldner M., Peebles P. J. E., 1979, ApJ, 227, 30
Sevilla-Noarbe I. et al., 2021, ApJS, 254, 24
Smee S. A. et al., 2013, AJ, 146, 32
Smith R. E. et al., 2003, MNRAS, 341, 1311
van den Busch J. L. et al., 2020, A&A, 642, A200
Wright E. L. et al., 2010, AJ, 140, 1868

APPENDIX A: VALIDATING THE PHOTOMETRIC REDSHIFT MODEL

Here, we show a test used to determine which photometric redshift model fits (one- or two-parameter) sufficiently match the clustering redshift measurements such that it would not bias cosmological results. We want a photo-z model that is flexible enough to agree with the clustering redshift points directly. We represent 'clustering points directly' in two ways, one with a multi-Gaussian fit to the clustering data, and one with a spline. The spline will it more exactly to the clustering data points, so we call it 'clustering direct' in Fig. A1. The multi-Gaussian fit may be more realistic as a result of the spline overfitting to noise in the clustering data points. For the fiducial cosmology used in Krause et al. (2021), we calculate a noiseless galaxy clustering and galaxy–galaxy lensing data vector given the different redshift distributions. We ignore cosmic shear (the third part of the '3 × 2' measurement) as it does not use the lens galaxies. We then run MCMC chains on these measurements to infer cosmological and galaxy bias parameters. In these chains, all cosmology and intrinsic alignment parameters are allowed to vary and other nuisance parameters (such as the weak lensing source redshifts) are fixed. We follow the model validation analysis choices outlined in Krause et al. (2021).

We ran chains with an input 'true redshift' distribution matching the unshifted, unstretched redMaGiC photo-z prediction. We then ran chains with four different redshift models: a one-parameter model using Δz in all bins to shift the photo-z predictions to match clustering, a two-parameter model that introduces a stretch parameter s in the fifth redMaGiC bin, a multi-Gaussian fit to the clustering data points, and a spline fit to the clustering data points. The results are shown in Fig. A1. In the first four redMaGiC bins, we find significant overlap for the constraints on σ_8, Ω_m and the galaxy bias, b, with our one-parameter model, and the redshift distributions that serve as a proxy for the clustering results. This signifies a good match between the clustering redshifts, and the one-parameter shifted photometric redshifts.

We do see that for the fifth redMaGiC bin, the galaxy bias is different when comparing the one-parameter model contours to the multi-Gaussian and direct clustering fits. We can infer that this means that a shifted redMaGiC photo- z distribution in this bin is still different enough from the clustering redshift distribution that they would produce statistically different amplitudes for galaxy clustering predictions in a given cosmology. The two-parameter contours show the constraints when the two-parameter fit is used in just the fifth bin. The addition of the stretch parameter, s, mitigates the discrepancy. With either the one- or two-parameter models, the cosmological constraints are unbiased compared with the clustering direct models. The specific σ differences between the two-parameter model and the multi-Gaussian fit for Ω_m, σ_8 and b^5 are 0.26, 0.39 and 0.31, respectively. The σ differences between the two-parameter model and the 'clustering direct' spline fit for Ω_m, σ_8 and b^5 are 0.11, 0.24 and 0.95, respectively. Because the 'clustering direct' and multi-Gaussian fits ignore the clustering uncertainties on shape, these values may effectively overestimate how much our model disagrees with the clustering data.

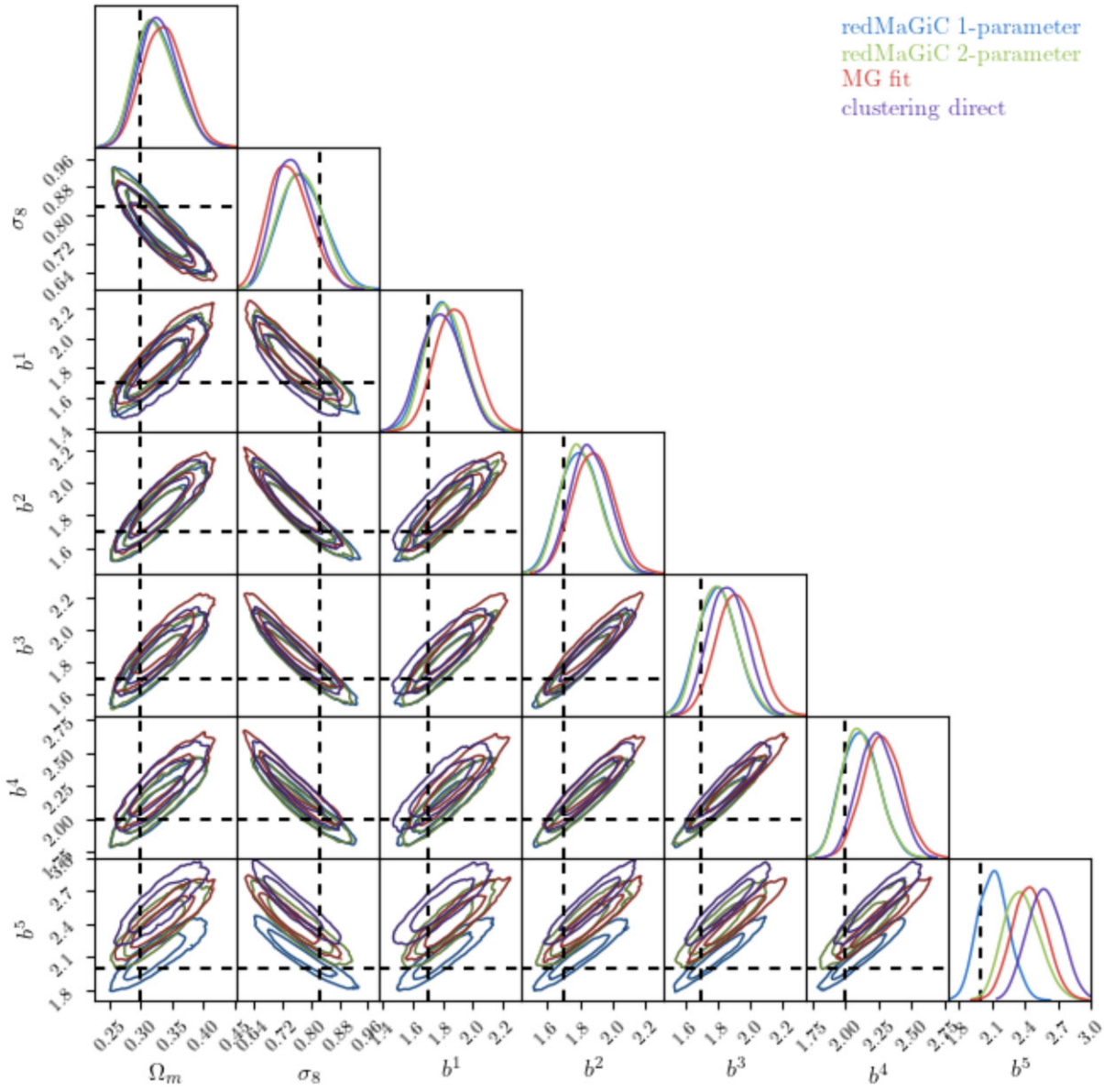

Figure A1. Simulated cosmological constraints from galaxy clustering and galaxy–galaxy lensing with four different redshift distribution models for the redMaGiC sample. The one-parameter contour uses the redMaGiC photometric redshift distribution with shift parameters in all redshift bins from the results in this work (Table 7). The two-parameter contour uses a two-parameter shift and stretch model for the fifth redMaGiC bin (Table 8). The MG fit contour uses a multi-Gaussian fit to the clustering redshift points. The clustering direct contour uses a spline fit to the clustering redshift points. The 'true redshift' data vector for the simulation is the unshifted redMaGiC photo- z prediction. We can see that the two-parameter model gives a better fit for the galaxy bias in bin 5 when compared with the predictions of the more exact fits to the clustering data, spline and multi-Gaussian.

APPENDIX B: BREAKDOWN OF UNCERTAINTY CONTRIBUTIONS TO MAIN RESULTS

Here, we show the uncertainty contributions for our main results in Tables 7–10. The calculations for overall uncertainty differ slightly between the one- and two-parameter cases. In the one-parameter case, there are three contributions of uncertainty. The first is the statistical errors from the cross-correlations of the unknown and reference samples and the autocorrelations of the reference samples. These give an uncertainty on the mean redshift in equation (7) with $w_{\rm uu}$ set to 1. The second contribution is from the power-law fit to the estimate of $w_{\rm uu}$ in equation (12). The uncertainty in the exponent γ is translated to an uncertainty contribution on the mean redshift of a photometric bin. Finally, the third contribution is the systematic uncertainty derived in Table 5.

For the two-parameter fit, the χ^2 fit for both Δz and s is done after the estimate and uncertainties of the power-law fit to $w_{\rm uu}$ (equation 12) are already applied. Thus, the uncertainties in this fit encompass the first two contributions of the one-parameter fit. We call this combination the statistical uncertainty. To this uncertainty, the systematic uncertainties from Table 6 are added. In each case of the one- and two-parameter uncertainty calculations, the contributions are added in quadrature. The contribution for each redshift bin is shown in Tables B1 and B2.

Table B1. One-parameter fit uncertainty contributions for our main one-parameter results in Tables 7 and 9.

Redshift bin	Total	Statistical	Power law	Systematic
redMaGiC bin 1	0.004	0.0026	0.0011	0.0025
redMaGiC bin 2	0.003	0.0016	0.0006	0.0025
redMaGiC bin 3	0.003	0.0018	0.0011	0.0025
redMaGiC bin 4	0.005	0.0038	0.0015	0.0025
redMaGiC bin 5	0.010	0.0089	0.0040	0.0025
MagLim bin 1	0.004	0.0028	0.0015	0.0025
MagLim bin 2	0.006	0.0036	0.0035	0.0025
MagLim bin 3	0.004	0.0027	0.0018	0.0025
MagLim bin 4	0.005	0.0036	0.0019	0.0025
MagLim bin 5	0.011	0.0097	0.0044	0.0025
MagLim bin 6	0.015	0.0127	0.0067	0.0025

Table B2. Two-parameter fit uncertainty contributions for our main two-parameter results in Tables 8 and 10.

Bin	Tot (Δz)	Stat (Δz)	Syst (Δz)	Tot (s)	Stat (s)	Syst (s)
redMaGiC bin 1	0.005	0.0025	0.0044	0.043	0.020	0.038
redMaGiC bin 2	0.005	0.0017	0.0044	0.045	0.023	0.038
redMaGiC bin 3	0.005	0.0029	0.0044	0.0025	0.029	0.038
redMaGiC bin 4	0.006	0.0035	0.0044	0.048	0.053	0.038
redMaGiC bin 5	0.006	0.0041	0.0044	0.065	0.052	0.038
MagLim bin 1	0.007	0.0048	0.0044	0.062	0.049	0.038
MagLim bin 2	0.011	0.0097	0.0044	0.093	0.085	0.038
MagLim bin 3	0.006	0.0036	0.0044	0.054	0.039	0.038
MagLim bin 4	0.006	0.0033	0.0044	0.051	0.033	0.038
MagLim bin 5	0.007	0.0049	0.0044	0.067	0.056	0.038
MagLim bin 6	0.008	0.0064	0.0044	0.073	0.062	0.038

APPENDIX C: ALTERNATIVE ASSESSMENTS OF SIMULATION TESTS

In this appendix, we describe a few variations on the evaluation of the tests on simulations in Section 5, focusing on the one-parameter fit results in Table 5. One alternative would be to remove a model having 'biased' tests at all, and only fit for additional uncertainty (i.e. fit for ω in equation 15, and set $b = 0$). We do evaluate the two-parameter fit in this manner in Section 5.6. In this case of no bias allowed, we would find uncertainty for Test 1, $\omega = 0.0016$, and for Test 5, $\omega = 0.0029$, for a combined uncertainty of 0.0033. For these calculations, we modify the degrees of freedom to be 5 instead of 4 in equation (15). This test would result in a very small increase in method systematic uncertainty from the fiducial choice, 0.0025.

We also address the question of whether it is appropriate that the six samples are treated as independent tests. While the six DES samples are different, there is significant redshift overlap between the redMaGiC and MagLim samples, so the same simulated reference BOSS/eBOSS galaxies are used for multiple samples. With only six simulated samples, it is difficult to definitively prove correlation or lack thereof. Ideally, a large set of simulated galaxy samples, in different regions of the sky, could test this. Within our limited samples, we can carry out a few tests though.

If there are correlations in the measured biases between samples, we would expect them to be on the bins with the most overlap in redshift space. The pairs of samples with significant redshift overlap in Figs 5 and 6 are [1, 2], [3, 4] and [5, 6]. Analysing only (mostly) non-overlapping redshift bins would likely have no correlations. We test this by performing again the calculations in Table 5 using separately just the three redMaGiC bins (samples 1, 3 and 5) and MagLim bins (samples 2, 4 and 6). The results for Tests 1 and 5 are shown in Table C1. For redMaGiC alone, combining the Test 1 and Test 5 results results in an overall bias of $b = 0.0007$ and systematic uncertainty $\omega = 0.0044$. For MagLim alone, the combined results give $b = 0.0011$ and $\omega = 0.0034$. These results are very similar to our fiducial values of $b = 0.0007$ and $\omega = 0.0025$. In the most pessimistic case of total correlation between overlapping samples, using just the three MagLim samples changes our bias and uncertainty measurements by less than 0.001 each.

Table C1. Re-analysis of the tests shown in Fig. 6 and Table 5. In this case, we derive separate bias and uncertainty values when analysing the three simulated redMaGiC and three simulated MagLimsamples separately.

Name of test	Bias	Uncertainty
redMaGiC alone		
Test 1: Method w/Spec-z	−0.0019	0.0033
Test 5: Clustering-z σ_j correction	0.0026	0.0029
MagLim alone		
Test 1: Method w/Spec-z	−0.0007	0.0000
Test 5: Clustering-z σ_j correction	0.0018	0.0034

Although we have limited information in testing the hypothesis that the pairs of overlapping samples give correlated results, we can look at Fig. 6. If we naively looked for correlations in, for example, the Test 1 and Test 5 results between samples, we might conclude that Test 1 has correlations in samples 3 and 6, and Test 5 has correlations in samples 1 and 6. These samples are very unlikely to be correlated though, as they cover different redshift ranges. This provides a little more evidence that the tests are likely to be uncorrelated.

A simple $w(\theta)$ measurement of overlapping galaxy samples should certainly be correlated. However, it is not clear that the many $w(\theta)$ measurements in the full clustering redshift methodology, including the different angular bins in equation (2), and different redshift bins in, for example, equation (7) should be correlated in their measurements of mean redshift. We leave a more explicit test of this for future work with a larger number of simulated samples.

We note that an evaluation of the two-parameter fits in Fig. 7 potentially indicates correlation in biases on the stretch parameter, unlike the one-parameter results. This is seen in bins 5 and 6, which do overlap in redshift, with each having a negative bias compared with the true width. However, we believe that if there is a correlation here, it is likely not a result of the covariance of the redshift bins, but a general bias for noisier clustering redshift measurements to predict too wide a redshift distribution (negative bias on Fig. 7). We did see in tests that adding scales to the measurements, which would decrease the uncertainties notably for higher redshift bins, also reduced this bias on the stretch parameter. We leave further study of this correlation of noise and bias in measured width to future work.

APPENDIX D: TESTS OF METHODOLOGY'S DEPENDENCE ON DENSITY

For any clustering statistics, uncertainty will depend significantly on the number of objects. In the pair counting of $w(r)$ in equation (1), a lower density of objects will result in larger uncertainty, which in this work we call the statistical uncertainty. We briefly investigate if there is also density dependence for the systematic errors determined in this section.

To do this, we take our six simulated samples shown in Fig. 5, as well as the simulated spectroscopic samples, and cut their regions to

Table D1. Results from recomputing the systematic uncertainty parameters of Tables 5 and 6 when using simulated samples that are one-fourth of the size of the original ones used in Section 4.

Method	Parameter	New value	Previous value
One-parameter test 1	b	−0.0014	−0.0014
One-parameter test 5	ω	0.0022	0.0013
One-parameter test 1	b	0.0014	0.0021
One-parameter test 5	ω	0.0038	0.0021
Two-parameter χ^2	$\delta\Delta z$	0.0044	0.0057
Two-parameter χ^2	δs	0.038	0.0067

approximately one-fourth of their original size. We then recompute each step of our methodology: taking clustering-z measurements with cross-correlations and autocorrelations of the sample. We repeat 'Test 1' by taking autocorrelations of the unknown (photometric) samples using their true redshifts. We also repeat 'Test 5', the method we will use on the data, where we take the photometrically binned (micro-binned) unknown sample autocorrelations, and use cross-correlations on nano-bins to calibrate those autocorrelations (equations 10 and 11). We finally recompute the two-parameter tests for the χ^2 method as well. In each case, we re-derive the systematic uncertainty or bias and compare with the values from Tables 5 and 6.

The new fits for systematic uncertainty in this one-fourth simulated data case are shown in Table D1. We see that the measures of bias in Tests 1 and 5 of Section 5 are the same or reduced, giving evidence that perhaps there is no overall bias of the method. We also see that the systematic uncertainty in both one-parameter (ω) and two-parameter tests is increased.

From this result, it is of interest to extrapolate how systematic uncertainties on the data may be underestimated. In the test described here, both the reference and unknown sample were cut into one-fourth of the size. The uncertainties measured in Table D1 increase by a factor of 1.3–1.8. As a toy model, we will say the uncertainties increased by roughly $4^{1/3} = 1.587$. This toy model would suggest systematic uncertainty scales as $\sqrt{N_r N_u}^{1/3}$. The data redshift bins have a range of number density contrasts with the full simulated samples. In this toy model, the systematic uncertainty in different redshift bins for MagLim and redMaGiC would range from two to three times the uncertainties derived in Section 5. As the systematic uncertainty is added in quadrature with the statistical uncertainty, the overall uncertainty would increase by less than a factor of 2 in most bins.

There are many caveats to this extrapolation. Most notably, if there is significant density dependence to the systematic uncertainty, then averaging the results of different simulated samples of different densities is likely not the optimal analysis. The cosmological analysis in DES Collaboration et al. (2022) ended up using the four most dense redshift bins (MagLim bins 1–4). It is likely that, in these bins, any increase in systematic uncertainty is at the lower end of this projection. If averaging different densities across the six simulated samples has a large effect, then the increases for these bins may be significantly overestimated here.

We tested the impact of having increased redshift uncertainty on the cosmology analyses in DES Collaboration et al. (2022) and Porredon et al. (2021a). We ran chains where the systematic redshift uncertainties for the two-parameter χ^2 fit from Table 6 were increased to Unc.$(\Delta z) = 0.01$ and Unc.$(s) = 0.1$, which increased the overall uncertainties in Table 10 by a factor of 2 or less. The main cosmological results used MagLim bins 1–4, so we did not need to consider the one-parameter fits, which were only used for redMaGiC. We ran chains for the '3 × 2' analysis of DES Collaboration et al. (2022) and the '2 × 2' analysis of Porredon et al. (2021a). In checking the impact on cosmological parameters σ_8, S_8 and Ω_m, we found in all cases parameter shifts of less than 0.17σ and contour increases of less than 9 per cent.

We note that in this appendix we are discussing systematic uncertainties beyond the statistical uncertainties, which also change with density. Interestingly, the statistical uncertainties have lower dependence on density. They are within a factor of 1.5 in comparing the simulated samples to data. This is less than the factor of 2–3 from the extrapolation of the systematic uncertainty. This may also suggest that more tests are needed to verify the magnitude of systematic uncertainty dependence on density.

The test in this appendix suggests that the accuracy and precision of clustering redshift methods may be dependent on density, beyond the usual counting statistics. More precise work with a larger suite of simulated samples at a range of densities will be needed to understand these effects in detail. However, the upper bounds of the increase in uncertainty from this section only increase the overall uncertainty by up to a factor of 2, and this has minimal impact on the cosmological results.

Several other parameters not explored here may also affect the precision of the clustering redshift measurements, including the exact shape of $n(z)$, the shape of the photo-z estimate, the width of the redshift distribution, the shape and strength of galaxy bias evolution with redshift, and the number of objects. Of these, the most likely correlation is with the number of objects, so this is where we investigate further. Future studies will attempt to quantify to a higher degree the dependence of the clustering redshift methodology on these several factors, as well as choices such as bin size and scales of measurement.

APPENDIX E: DES FLUX-LIMITED SAMPLE

In this section, we describe clustering redshift measurements for the DES 'flux-limited' sample described in Porredon et al. (2021b). This sample is not used in the DES Year 3 cosmology analyses (DES Collaboration et al. 2022). In Porredon et al. (2021b), this sample's cosmological constraining power is compared to the redMaGiC and MagLim samples.

We carry out the same full clustering redshift analysis as was done for the redMaGiC and MagLim samples in Section 6. The results are shown in Fig. E1 and Table E1. We see that this sample shows considerable photo-z biases of several σ, in sharp contrast to the results for the other two DES lens samples. Most notable are large excesses of low-z galaxies measured by the clustering redshifts in the first three tomographic bins ($z \in [0.2, 0.65]$) compared with the photo-z predictions.

Table E1. Clustering redshift results for the flux-limited sample (one-parameter). This sample is not used in the DES Year 3 cosmology analyses. The systematic uncertainties listed include the power-law uncertainty and the 0.0025 method uncertainty from Section 5.

Redshift bin	Δz	$\delta\Delta z$(syst.)	$\delta\Delta z$(stat.)
1: $z_{ph} \in [0.2, 0.4]$	-0.041 ± 0.007	0.005	0.005
2: $z_{ph} \in [0.4, 0.5]$	-0.058 ± 0.007	0.006	0.004
3: $z_{ph} \in [0.5, 0.65]$	-0.056 ± 0.007	0.005	0.005
4: $z_{ph} \in [0.65, 0.8]$	-0.026 ± 0.006	0.005	0.003
5: $z_{ph} \in [0.8, 1.05]$	0.008 ± 0.015	0.004	0.014

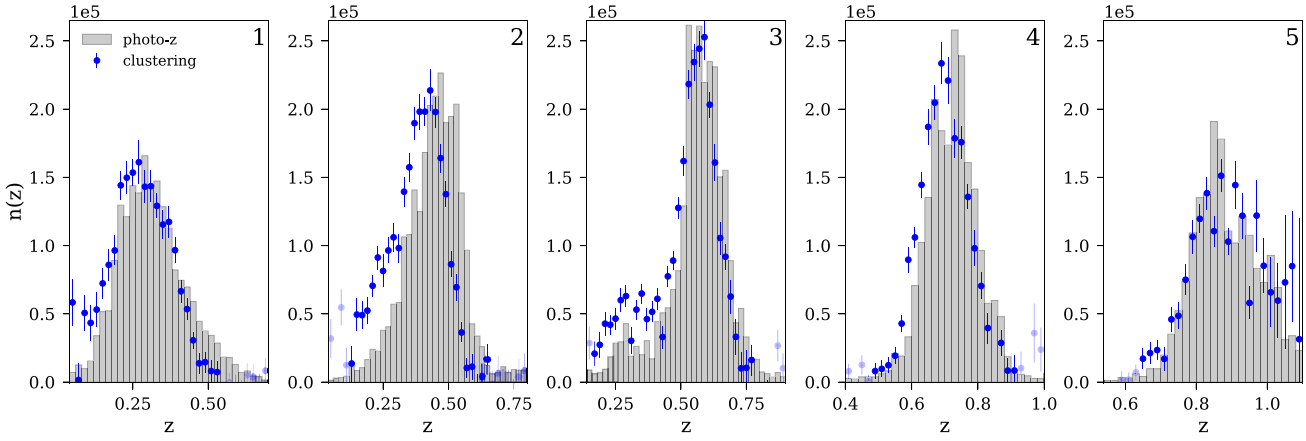

Figure E1. The clustering redshift measurements of the flux-limited DES sample. This sample is not used in DES Year 3 cosmology analyses, but has been studied in Porredon et al. (2021b).

[1]*Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, Madison, WI 53706, USA*
[2]*Physics Department, William Jewell College, Liberty, MO 64068, USA*
[3]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[4]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[5]*Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain*
[6]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*
[7]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain*
[8]*Kavli Institute for Particle Astrophysics and Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA*
[9]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[10]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[11]*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[12]*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*
[13]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[14]*Department of Physics, Duke University Durham, NC 27708, USA*
[15]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
[16]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[17]*Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[18]*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
[19]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP, 05314-970, Brazil*
[20]*Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[21]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain*
[22]*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*
[23]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[24]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[25]*Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[26]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*
[27]*National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[28]*INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[29]*Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy*
[30]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[31]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[32]*Department of Physics and Astronomy, University of Utah, 115 S. 1400 E., Salt Lake City, UT 84112, USA*
[33]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
[34]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[35]*Institute of Theoretical Astrophysics, University of Oslo, P.O. Box 1029 Blindern, NO-0315 Oslo, Norway*
[36]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[37]*School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[38]*Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[39]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[40]*Institute of Physics, Laboratory of Astrophysics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland*
[41]*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
[42]*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
[43]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[44]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[45]*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany*
[46]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany*
[47]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[48]*Waterloo Centre for Astrophysics, Dept. of Physics and Astronomy, University of Waterloo, 200 University Ave. W., Waterloo ON N2L 3G1, Canada*

[49]*Department of Physics and Astronomy, Sejong University, Seoul 143-747, Korea*
[50]*School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*
[51]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[52]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, UK*

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