Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **525,** 3896–3922 (2023) https://doi.org/10.1093/mnras/stad2402
Advance Access publication 2023 August 17

# The Dark Energy Survey Year 3 high-redshift sample: selection, characterization, and analysis of galaxy clustering

C. Sánchez [1,2,3,4★] A. Alarcon [5★] G. M. Bernstein [4] J. Sanchez,[6] S. Pandey,[4] M. Raveri,[7] J. Prat,[8,9] N. Weaverdyck,[10,11] I. Sevilla-Noarbe,[12] C. Chang,[8,9] E. Baxter,[13] Y. Omori,[8,9,14,15] B. Jain,[4] O. Alves,[10] A. Amon,[16,17] K. Bechtol,[18] M. R. Becker,[5] J. Blazek,[19] A. Choi,[20] A. Campos,[21] A. Carnero Rosell,[22,23,24] M. Carrasco Kind,[25,26] M. Crocce,[3,27] D. Cross,[3] J. DeRose,[11] H. T. Diehl,[28] S. Dodelson,[21,29] A. Drlica-Wagner,[8,9,28] K. Eckert,[4] T. F. Eifler,[30,31] J. Elvin-Poole,[32] S. Everett,[31] X. Fang,[33,30] P. Fosalba,[3,27] D. Gruen,[34] R. A. Gruendl,[25,26] I. Harrison,[35] W. G. Hartley,[36] H. Huang,[30,37] E. M. Huff,[31] N. Kuropatkin,[28] N. MacCrann,[38] J. McCullough,[15] J. Myles,[14,15,39] E. Krause,[30] A. Porredon,[40,41,42] M. Rodriguez-Monroy,[12] E. S. Rykoff,[15,39] L. F. Secco,[9] E. Sheldon,[43] M. A. Troxel,[44] B. Yanny,[28] B. Yin,[21] Y. Zhang,[45] J. Zuntz,[42] T. M. C. Abbott,[45] M. Aguena,[23] S. Allam,[28] F. Andrade-Oliveira,[10] E. Bertin,[46,47] S. Bocquet,[34] D. Brooks,[48] D. L. Burke,[15,39] J. Carretero,[2] F. J. Castander,[3,27] R. Cawthon,[49] C. Conselice,[50,51] M. Costanzi,[52,53,54] M. E. S. Pereira,[55] S. Desai,[56] P. Doel,[48] C. Doux,[4,57] I. Ferrero,[58] B. Flaugher,[28] J. Frieman,[9,28] J. García-Bellido,[59] G. Gutierrez,[28] K. Herner,[28] S. R. Hinton,[60] D. L. Hollowood,[61] K. Honscheid,[40,41] D. J. James,[62] K. Kuehn,[63,64] J. L. Marshall,[65] J. Mena-Fernández,[12] F. Menanteau,[25,26] R. Miquel,[2,66] R. L. C. Ogando,[67] A. Palmese,[33] F. Paz-Chinchón,[16,25] A. Pieres,[23,67] A. A. Plazas Malagón,[68] E. Sanchez,[12] V. Scarpine,[28] M. Schubnell,[10] M. Smith,[69] E. Suchyta,[70] G. Tarle,[10] D. Thomas,[71] C. To[40] and (DES Collaboration)

*Affiliations are listed at the end of the paper*

Accepted 2023 August 3. Received 2023 July 14; in original form 2022 December 13

## ABSTRACT

The fiducial cosmological analyses of imaging surveys like DES typically probe the Universe at redshifts $z < 1$. We present the selection and characterization of high-redshift galaxy samples using DES Year 3 data, and the analysis of their galaxy clustering measurements. In particular, we use galaxies that are fainter than those used in the previous DES Year 3 analyses and a Bayesian redshift scheme to define three tomographic bins with mean redshifts around $z \sim 0.9$, 1.2, and 1.5, which extend the redshift coverage of the fiducial DES Year 3 analysis. These samples contain a total of about 9 million galaxies, and their galaxy density is more than 2 times higher than those in the DES Year 3 fiducial case. We characterize the redshift uncertainties of the samples, including the usage of various spectroscopic and high-quality redshift samples, and we develop a machine-learning method to correct for correlations between galaxy density and survey observing conditions. The analysis of galaxy clustering measurements, with a total signal to noise S/N $\sim 70$ after scale cuts, yields robust cosmological constraints on a combination of the fraction of matter in the Universe $\Omega_{\rm m}$ and the Hubble parameter $h$, $\Omega_m h = 0.195^{+0.023}_{-0.018}$, and 2–3 per cent measurements of the amplitude of the galaxy clustering signals, probing galaxy bias and the amplitude of matter fluctuations, $b\sigma_8$. A companion paper (in preparation) will present the cross-correlations of these high-$z$ samples with cosmic microwave background lensing from *Planck* and South Pole Telescope, and the cosmological analysis of those measurements in combination with the galaxy clustering presented in this work.

**Key words:** galaxies: high-redshift – cosmological parameters – large-scale structure of Universe.

## 1 INTRODUCTION

The combination of large-scale structure (LSS) and weak gravitational lensing (WL) constitutes one of the main avenues to study cosmology and to stress test the standard cosmological model. In recent years, several imaging surveys such as the Hyper Suprime-Cam (HSC[1]), the Kilo-Degree Survey (KiDS[2]), and the Dark Energy

★ E-mail: carles.sanchez.alonso@gmail.com (CS); alexalarcongonzalez@gmail.com (AA)

[1] hsc.mtk.nao.ac.jp/ssp/
[2] kids.strw.leidenuniv.nl

Survey (DES[3]), analysing data from more than 100 million galaxies, have used galaxy weak lensing to produce cosmological constraints that rival in precision those from cosmic microwave background (CMB) experiments like Planck (see Hikage et al. 2019; Heymans et al. 2021; DES Collaboration 2022; and references therein). These analyses have reported tensions between the amplitude of structures at late time and the predictions from the CMB (the so-called $S_8$ tension). However, the majority of these analyses probe the Universe at low redshifts, $z < 1$. There exist at least three reasons for this. First, due to the faint nature of high-redshift galaxies, it is difficult for imaging surveys to characterize such populations, both in terms of redshift distributions and also in terms of mapping the effect of spatially varying observing conditions on the selection function. Secondly, it is challenging to measure shapes of high-redshift sources for galaxy lensing at sufficient signal to noise. And third, even if those galaxy sources could be defined, their lensing signals are still most sensitive to mass structure at $z < 1$. On the other hand, if one can get around the first of these issues and characterize high-redshift lens galaxy samples, then the use of CMB lensing will provide a solution for the second and third problems.

The definition and characterization of galaxy samples at higher redshifts would enable a more optimal combination with CMB lensing, whose sensitivity peaks around $z = 2$ and drops significantly at redshifts $z < 1$. In this way, a combination of galaxy clustering and CMB lensing at high redshift would be key to cosmology in several ways. On the one hand, the regime at redshifts $z \geq 1.5$ remains largely unexplored by galaxy surveys in the context of the $S_8$ tension, and various alternative dark energy models predict deviations from the standard model at high redshifts (Bull, White & Slosar 2021), which could be tested in this way. On the other hand, being able to make this measurement is important to constrain large-scale observables like primordial non-Gaussianity, which would open the window to the physics of the early inflationary period sourcing the LSSs we see in the Universe today (Schmittfull & Seljak 2018). Furthermore, CMB lensing is subject to different systematic errors than galaxy lensing – the former measurement is not affected by intrinsic alignments or galaxy blending, and the redshift of the CMB is well known as opposed to the case of galaxy sources.

There exist numerous previous analyses that have explored the combination of galaxy clustering and CMB lensing to probe cosmology at redshifts $z < 1$ (Abbott et al. 2019; Marques & Bernui 2020; Alonso et al. 2021; Hang et al. 2021; Chang et al. 2023). Some analyses have also used the combination to probe cosmology at higher redshifts. In particular, the analysis of the unWISE sample (Schlafly, Meisner & Green 2019; Krolewski et al. 2020; Krolewski, Ferraro & White 2021) provided such measurements in three broad redshift bins, the last one with a median redshift around $z = 1.5$. Also, the HSC survey has explored much higher redshift regimes using dropout galaxies over smaller areas (Harikane et al. 2018; Ono et al. 2018), probing the Universe at the $4 < z < 7$ regime (Miyatake et al. 2022).

For the particular case of the DES, the analysis of Year 3 (Y3) data has so far used two different lens galaxy samples, MAGLIM and REDMAGIC (Porredon et al. 2022; DES Collaboration 2022; Pandey et al. 2022). The MAGLIM sample is a magnitude-limited galaxy selection, split into six redshift bins using the Directional Neighbourhood Fitting algorithm (De Vicente, Sánchez & Sevilla-Noarbe 2016), and the first four bins of the sample, covering an approximate redshift range $0 < z < 1$, were used as the fiducial lens sample in the Dark Energy Survey Year 3 (DES Y3) analysis. The REDMAGIC (Rozo et al. 2016) is a sample of bright Luminous Red Galaxies (LRGs), covering a similar redshift range in five redshift bins, and was used in Y3 as an alternative lens sample.

In this work, we push the limits of the DES Y3 data to explore the regime at redshift $z > 1$. To this end, we select and characterize 'high-redshift' (high-$z$) samples of galaxies in the DES wide-field survey. This includes the estimation of the redshift distributions of the samples and their uncertainties, corrections for variations in completeness across the survey footprint due to varying observing conditions, and characterization of the lens magnification coefficients of the samples. The definition and characterization of these high-$z$ samples differs from the process used for the fiducial DES Y3 lens samples (Porredon et al. 2022; Pandey et al. 2022) in several ways:

(i) We start from a fainter galaxy selection, already excluding all lens galaxies used in the DES Y3 fiducial analysis.
(ii) Both the selection and redshift characterization of the samples are based on a Bayesian scheme using *Self-Organizing Maps* (SOMs), and we use a new SOM algorithm, better suited for lower S/N galaxies (different than that used in Myles et al. 2021).
(iii) We use a different redshift marginalization scheme, explicitly accounting for uncertainties in low-redshift tails of the redshift distributions.
(iv) We use a non-linear, machine-learning-based approach to account for correlations in the galaxy number density with survey observing properties.

Steps (i) and (ii) are the ones responsible for the selection of high-redshift galaxies, while steps (iii) and (iv) are necessary because of that faint, high-redshift selection. The definition and characterization of the high-$z$ sample in this work is followed by the analysis of the clustering measurements of the galaxies in the sample. The clustering measurements are used to place constraints on the cosmological model, in particular as the shape of the clustering signal is sensitive to the scale of matter–radiation equality in the mass power spectrum, which in turn depends on a combination of the matter density $\Omega_m$ and the Hubble constant $h$, close to the direction $\Omega_m h$ (see e.g. Philcox et al. 2021). The high-$z$ samples defined in this work, given their redshift range and sky density, will make excellent lens galaxy samples for CMB lensing. In this way, this paper will be followed by a companion paper (in preparation) that will present the cross-correlations between these high-$z$ samples and CMB lensing from Planck (Planck Collaboration VIII 2020) and the South Pole Telescope (SPT; Carlstrom et al. 2011), and use the combination of clustering and CMB lensing to place constraints on the cosmological model using information from high redshift.

This paper is organized as follows. Section 2 describes the different data products used for the analysis. Section 3 describes the redshift inference scheme and the method to select tomographic bins. Section 4 describes the way we correct for correlations between galaxy density and survey observing properties. Section 5 presents the characterization of redshift uncertainties, and the parametrization we use to marginalize over them in the clustering analysis. Section 6 describes the characterization of lens magnification for the high-$z$ samples. Finally, Section 7 presents the measurements and analysis of galaxy clustering, and we conclude in Section 8.

## 2 DATA

In this section, we describe and motivate the different data samples to be used in this work. We begin with the DES Y3 wide-field data, which will contain our high-$z$ samples, and then describe other data

[3] darkenergysurvey.org

sets needed for the characterization of those samples: the DES deep-field data, and the external data used for redshift characterization.

### 2.1 DES wide-field data

The high-$z$ samples are subsets of the DES Year 3 Gold catalogue of photometric objects (Sevilla-Noarbe et al. 2021), which has a total of nearly 400 million objects in about 5000 square degrees of area, covering the entire DES footprint. After removing stars and applying quality cuts (following Sevilla-Noarbe et al. 2021), the catalogue consists of ∼227 million galaxies. For these objects, we use single-object-fitting (SOF) photometry in the *griz* bands, which have magnitude limit (defined as the average SOF magnitude at S/N = 10) of 23.8, 23.6, 23.0, and 22.4, respectively. We apply an initial $i$-band magnitude 'pre-selection' of $22 < i < 23.5$. The lower limit of this cut removes bright galaxies that are unlikely to be at redshifts $z > 1$, and the faint limit excludes the region of magnitude space where the DES Y3 Gold catalogue becomes highly incomplete. Please note that, even with the $i < 23.5$ cut, this selection includes galaxies measured with S/N < 10 in the $i$ band, pushing the limits of the DES Y3 sample, and therefore the completeness of the sample has significant spatial variations. The characterization of that spatial completeness is a key aspect of this work, and is described in Section 4.

For the pre-selected sample, we apply the standard DES Y3 mask, which includes masking of astrophysical foregrounds (e.g. bright stars and large nearby galaxies) and of regions with recognized data-processing issues, as described in Sevilla-Noarbe et al. (2021). Given that we are pushing the limits of DES Y3 photometry, we apply some additional conservative cuts on the mask to avoid regions where our completeness corrections would be less reliable: we remove the 3 per cent of the footprint area with the highest stellar density, the 3 per cent with the highest (worst) $g$-band seeing, and then we remove the worst 10 per cent area in photometric depth, exposure times, and sky brightness in each of the *griz* bands, some of which are correlated. After applying this mask, the $22 < i < 23.5$ pre-selected galaxy sample has a total of 77 million galaxies in 2621 square degrees of area. For comparison, the fiducial DES Y3 analysis uses 4143 square degrees of total area.

The analysis presented here will be followed by a companion paper (*in preparation*) that will combine the clustering measurements shown here with CMB lensing measurements from the *Planck* satellite and the SPT. Due to SPT data being available only in the south region of the DES Y3 footprint, we will split the sample in this work into two independent regions, 'North' (Dec. $> -39°$) and 'South' (Dec. $< -40°$), and test for the consistency of the two. For that test, we choose to leave a separation of 1 degree between the two regions, which corresponds to the maximum angular separation used later on in the galaxy clustering measurements. A similar separation of the DES footprint was made in the analyses studying CMB lensing for the fiducial DES Y3 sample (Abbott et al. 2019; Baxter et al. 2019; Omori et al. 2019a, b).

### 2.2 DES deep-field data and artificial wide-field data

The scheme for redshift selection and characterization, described in detail in Section 3, makes extensive use of DES deep-field data, described extensively in Hartley et al. (2022). In short, we use four *deep* fields, named E2, X3, C3, and COSMOS (COS), covering areas of 3.32, 3.29, 1.94, and 1.38 square degrees, respectively (see fig. 2 in Myles et al. 2021 for a visual description). After masking regions with artefacts such as cosmic rays, artificial satellites, meteors, asteroids,

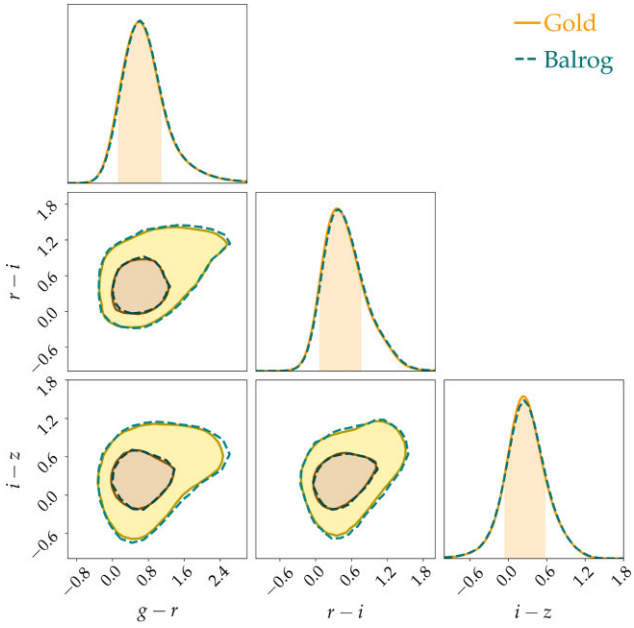


**Figure 1.** The distribution of photometric colours in the DES wide-field *griz* bands, after a pre-selection cut of $22 < i < 23.5$, using the data (Gold) and the artificial data (BALROG).

and regions of saturated pixels, 5.2 square degrees of overlap with the UltraVISTA and VIDEO near-infrared (NIR) surveys (McCracken et al. 2012; Jarvis et al. 2013) remain. We remove stars based on the kNN classification of Hartley et al. (2022), and additionally by training a SOM and removing regions of colour space with a majority of stellar contamination according to the Laigle et al. (2016) catalogue. This yields 2.8M detections with measured *ugrizJHK*$_s$ photometry with limiting magnitudes 24.64, 25.57, 25.28, 24.66, 24.06, 24.02, 23.69, and 23.58, substantially fainter than the faintest galaxies in the sample of source galaxies. In this work, we frequently refer to this sample and its photometry as *deep (field)* data.

So far we have described the wide-field DES data to be used over the full footprint and a set of deep-field photometry over a smaller area. In order to establish the relationship between these two data sets we use the BALROG (Suchyta et al. 2016) software, which injects simulated galaxies based on the DES deep fields into real images from DES wide-field observations. For this analysis, BALROG was used to inject model galaxies, with profiles fit to deep-field galaxies, into the wide-field footprint (Everett et al. 2022). After injecting galaxies into images, the output is analysed by the DES Y3 photometric pipeline (Morganson et al. 2018). Each deep-field galaxy is injected multiple times at different positions in the footprint. The resulting matched catalogue of 3194 291 injection-realization pairs, which contains both deep and wide photometric information, is a key part of our redshift calibration scheme since it quantitatively connects the two photometric spaces. This catalogue will be referred to as the *Deep/*BALROG Sample, and contains a total of 432 657 unique deep-field galaxies having at least 1 BALROG realization that passes the wide-field selection criteria.

Because we will use the BALROG sample to establish the relationship between wide and deep photometry in DES Y3, it is important that BALROG wide-field detections follow similar photometric distributions to the actual DES Y3 wide-field data in the Gold sample. Fig. 1 shows the distribution of colours in the DES Y3 photometry for the data (Gold) and for the artificial realizations of deep galaxies

(BALROG) for the pre-selected sample described in Section 2.1 ($22 < i < 23.5$). As desired, the colour distributions of data and artificial realizations of deep galaxies are in excellent agreement.

### 2.3 Redshift data

Our analysis relies on the use of galaxy samples with known redshift and deep-field photometry. To this end, we use catalogues of both high-resolution spectroscopic and multiband photometric redshifts, and we develop an experimental design that allows us to test uncertainty in our redshift calibration due to biases in these samples. The spectroscopic catalogue we use contains both public and private spectra from the following surveys: zCOSMOS (Lilly et al. 2009), C3R2 (Masters et al. 2017, 2019; Stanford et al. 2021), VVDS (Le Fèvre et al. 2013), and VIPERS (Scodeggio et al. 2018). We use two multiband photo-$z$ catalogues from the COSMOS field (Scoville et al. 2007): the `COSMOS2015` 30-band photometric redshift catalogue (Laigle et al. 2016), which includes 30 broad, intermediate, and narrow bands covering the UV, optical, and IR regions of the electromagnetic spectrum, and the `PAUS + COSMOS` 66-band photometric redshift catalogue (Alarcon et al. 2021) from the combination of PAU Survey data (Eriksen et al. 2019; Padilla et al. 2019) in 40 narrow-band filters and 26 `COSMOS2015` bands excluding the mid-infrared. We build a redshift calibration sample in the deep fields from the overlapping redshift information we find in these surveys. We prioritize information coming from spectroscopic surveys (S), then `PAUS + COSMOS` (P) and finally `COSMOS2015` (C), and we call this redshift sample SPC.[4]

## 3 REDSHIFT METHODOLOGY

This section describes our redshift inference scheme, which allows us to select and characterize samples of high-$z$ galaxies using the data described in the previous section. The next sections will describe the characterization of the uncertainties in the angular and redshift distributions of these high-$z$ samples.

We work under the framework presented in Sánchez & Bernstein (2019), in which galaxy 'types' are defined by observed properties rather than rest-frame properties, and we call them *phenotypes*. We will use the low-noise, several-band photometry available in the deep fields to define our phenotypes, and we will discretize such photometry using a SOM (Kohonen 1982; Masters et al. 2015). In this way, every cell in the Deep SOM will be a phenotype, and we will index them with $c$. This approach, proposed initially in Sánchez & Bernstein (2019), has now been successfully used in several analyses both using simulations (Buchs et al. 2019; Alarcon et al. 2020) and real data (Myles et al. 2021; Giannini et al. 2022).

We also discretize the wide-field photometry into a SOM, with wide cells indexed by $\hat{c}$. With this discretized mapping of deep and wide photometric spaces, we can estimate the redshift distribution of a given wide cell $\hat{c}$, passing a wide selection $\hat{s}$, by marginalizing over deep-field information $c$:

$$p(z|\hat{c},\hat{s}) = \sum_c p(z|c,\hat{c},\hat{s})p(c|\hat{c},\hat{s}). \quad (1)$$

The first term on the right contains information about the redshift of deep phenotypes, while the second term connects the deep and wide photometric spaces. Having the expression for the redshift distribution of a wide cell, we can construct a sample of galaxies by joining wide cells $\hat{c}$ into tomographic bins $\hat{b}$, and their redshift distribution will simply become the sum of its constituents weighted by the occupation of wide cells:

$$p(z|\hat{b},\hat{s}) = \sum_{\hat{c}\in\hat{b}} p(z|\hat{c},\hat{s})p(\hat{c}|\hat{s},\hat{b}), \quad (2)$$

$$\propto \sum_{\hat{c}\in\hat{b}}\sum_c p(z|c,\hat{c},\hat{s})p(c|\hat{c},\hat{s})p(\hat{c}|\hat{s}), \quad (3)$$

$$\approx \sum_{\hat{c}\in\hat{b}}\sum_c p(z|c,\hat{b},\hat{s})p(c|\hat{c},\hat{s})p(\hat{c}|\hat{s}). \quad (4)$$

Going from equations (2) to (3), we use the fact that $p(\hat{c}|\hat{b},\hat{s}) = p(\hat{c}|\hat{s})/\left(\sum_{\hat{c}\in\hat{b}} p(\hat{c}|\hat{s})\right)$ for $\hat{c}\in\hat{b}$, and and in the last line we approximate $p(z|c,\hat{c},\hat{s}) \approx p(z|c,\hat{b},\hat{s}.)$ The need for conditioning on bin membership rather than wide-cell measurement [going from equations (3) to (4)], and the accuracy of this approximation, will be investigated in Section 5 and Appendix A3. The final expression computes the redshift distribution of tomographic bins made of wide-field SOM cells. We use different samples to estimate the different terms in it, as we describe next:

(i) $p(\hat{c}|\hat{s})$ is computed from our wide sample, which consists of all galaxies in the DES Year 3 Gold catalogue passing the pre-selection performed in Section 2 ($22 < i < 23.5$).

(ii) $p(c|\hat{c},\hat{s})$ is computed from our Deep and BALROG Samples, which consist of all detected and selected BALROG realizations of the galaxies in the Deep Sample. We call this term the *transfer function*.

(iii) $p(z|c,\hat{b},\hat{s})$ is computed from the Redshift Sample subset of the Deep Sample, for which we have reliable redshifts, 8-band deep photometry, and wide-field BALROG realizations.[5]

The redshift scheme followed in this work is similar to that used in Myles et al. (2021) for the selection and characterization of weak lensing source galaxy samples, but there exist some important differences:

(i) We perform a pre-selection cut on our sample of $22 < i < 23.5$, to remove bright galaxies at low redshift and low S/N faint galaxies, cutting the bright end of the $18.5 < i < 23.5$ used in Myles et al. (2021).

(ii) In this work, we use DES *griz* wide photometry, while the analysis in Myles et al. (2021) uses *riz* information only.

(iii) We also use a different SOM algorithm, improved to better handle the classification of lower S/N galaxies. This will be described in detail in Section 3.1.

(iv) The tomographic bins for this work are selected using both the mean redshifts of the Wide SOM cells and also their estimated low-redshift fraction, to avoid having large low-$z$ tails in the tomographic bins. The selection in Myles et al. (2021) relies only on mean redshift information.

### 3.1 The Deep SOM

In this work, we a use SOM to characterize and discretize the deep photometric space, described in Section 2.2. The SOM algorithm uses unsupervised learning to project the eight-dimensional (8D) deep photometric data ($ugrizJHK_s$) on to a lower dimensional

[4] An identical notation was used in Myles et al. (2021).

[5] This term could, in principle, be computed from the overlapping photometry of the deep and wide fields. However, the region where these samples overlap is small and it is not representative of the observing conditions found across the whole survey footprint, which are much more well sampled by making use of BALROG.

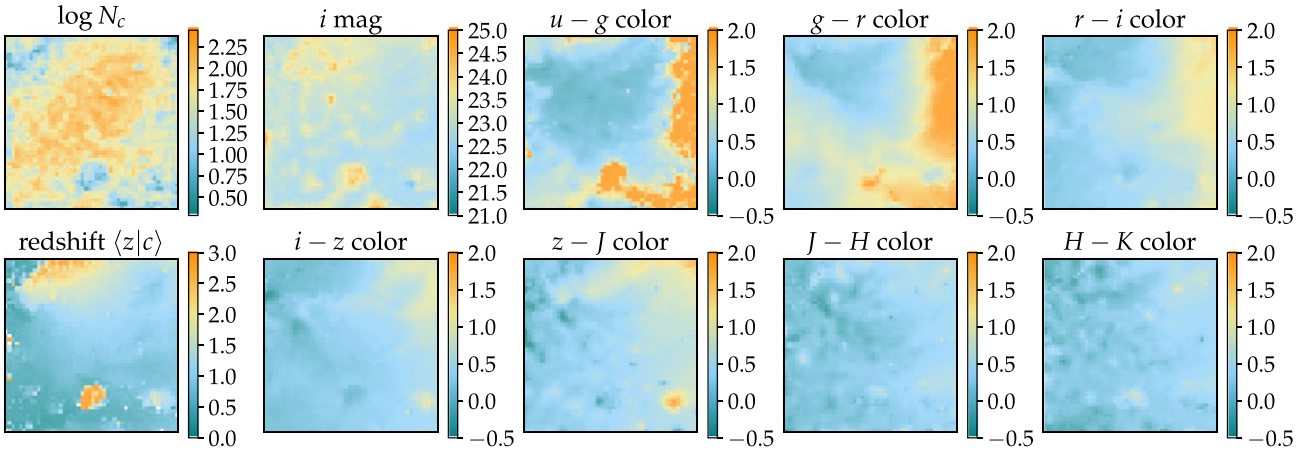


**Figure 2.** Visualization of various properties of the Deep SOM described in Section 3.1. In particular, we show the photometric properties of the map, namely the mapping of $i$-band magnitude and seven of the photometric colours, using the DES deep galaxy sample described in Section 2.2. We also show the SOM galaxy occupation, $N_c$, and the redshift mapping of the SOM using SPC redshift galaxies matched to DES deep photometry. A zoomed-in version of the Deep SOM redshift mapping is given in Fig. A3.

grid, in our case a two-dimensional (2D) grid, while attempting to preserve the topology of the 8D space. This means that similar objects in the 8D space will be grouped together in the SOM, enabling a visual understanding of features, especially in a 2D SOM. Each of the cells in the Deep-SOM 2D grid will be considered a galaxy *phenotype* in our scheme.

There is considerable flexibility in the implementation of the SOM algorithm. We alter the SOM algorithm from that used in previous DES analyses (such as Myles et al. 2021; Giannini et al. 2022) with the purpose of improving the classification of galaxies of the low- and modest-S/N photometry used in this work. This is done by altering the distance metric used by the SOM algorithm to incorporate flux uncertainties. We also allow magnitude (or flux) information, not just colours, to be used in redshift estimation, and we do not impose periodic boundary conditions on the map. This SOM algorithm was introduced and is described in detail in the appendix of Sánchez et al. (2020).

There is also flexibility in the size of the SOM. A larger number of SOM cells can improve the representative power of the map, and hence can be used to describe more complex spaces and resolve finer redshift distinctions. Using too many cells can, however, cause overfitting, with the map modelling noisy features of the data. The Deep SOM in this work uses a 48 × 48 SOM. For comparison purposes, the Deep SOM describing the DES Year 3 space in Myles et al. (2021) was 64 × 64 in size. We use a smaller SOM size since the wide-field pre-selection cut-off $22 < i < 23.5$ we apply to our sample reduces the volume of our wide-field photometric space, and our Deep SOM only uses deep galaxies whose BALROG injections have passed this criteria at least once (see Section 2.2).

Fig. 2 shows several properties of the Deep SOM used in this work. It is worth noting that the particular structure of the map depends on randomized initial conditions and training, but the overall topological structure will be similar across different runs. The figure shows different photometric properties of the SOM, mapping colours and $i$-band magnitude. The $u-g$ colour mapping shows how most of the map has a near-constant value of $u-g$, but there are well-defined areas showing strong positive (red) values of $u-g$, corresponding to breaks in the spectrum of galaxies such as the Lyman and Balmer breaks (these behaviour is also seen in other SOM analyses such as Masters et al. 2015). The $z-J$ colour shows a different structure across the map, showing variation across the regions where $u-g$ was constant and close to zero. We also show the mapping of $i$-band magnitude across the map. In this case, it is worth noting that even though our target sample has a selection of $22 < i < 23.5$, galaxies fainter than $i = 23.5$ have a non-zero probability of being selected in our sample because of noise fluctuations. Since we are including in the Deep SOM all deep galaxies whose artificial injections make the selection at least once, that means that we include some galaxies as faint as $i \simeq 25$.

Fig. 2 also shows the Deep SOM galaxy occupation, $n(c)$, the density of galaxies as a function of position in the deep photometric space probed by the SOM. Perhaps most importantly, the lower left-hand panel shows the redshift mapping of the Deep SOM. For this panel, we use the subset of deep galaxies that have a match in the SPC redshift sample (described in Section 2.3), and compute the mean redshift of the galaxies occupying each SOM cell. This plot depicts a smooth mapping of redshift in the SOM, reasonably smoother than the mapping of some colours or magnitudes, even though redshift information is never used in the SOM training.

Since we are mainly concerned about high redshift in this work, it is interesting to explore the regions of the map that correspond to that regime. There exist two main areas of high-$z$ galaxies in the SOM. There is a first high-$z$ region in the upper part of the SOM, with a smooth gradient to middling redshifts in the central part of the map. Fig. 2 shows the upper high-$z$ region to have a small $u-g$ colour (no break between the $u$ and $g$ bands), with positive and smoothly varying $z-J$ colour, and faint magnitudes in the $i$ band. There is a second 'island' in the lower centre of the SOM where very high-$z$ galaxies live, surrounded by low-redshift galaxies. This region has large (red) $u-g$ colour and also large (faint) $i$-band magnitude, i.e. is the part of photometric space where we encounter Lyman-break galaxies at high redshift. It also hosts faint Balmer-break galaxies at low redshift, and these two galaxy populations are known to present important degeneracies in the colour–redshift relation. That degeneracy is also responsible for a large redshift scatter in that part of the SOM. Finally, regarding the redshift mapping of the Deep SOM, it is important to point out that the vast majority of cells in the map contain galaxies from the SPC redshift sample, with only a four cells (out of 2304) containing no redshift information.

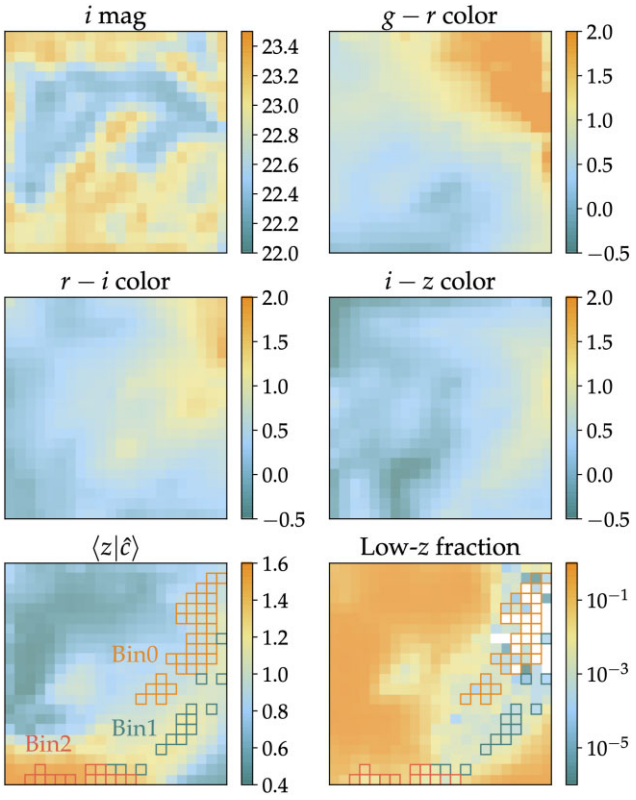


**Figure 3.** Visualization of various properties of the Wide SOM described in Section 3.2. In particular, we show the photometric properties of the map, namely the mapping of $i$-band magnitude and three of the wide photometric colours, using the DES wide galaxy sample described in Section 2. The bottom left-hand panel shows the redshift mapping of the Wide SOM, using SPC redshift galaxies matched to DES deep photometry and the BALROG transfer function between deep and wide photometry, as described in equation (1) and Section 3.2. Overlaid we can see the cells of the Wide SOM that constitute the three tomographic bins used in this work, following the procedure described in Section 3.3 and Fig. 4. The bottom right-hand panel shows the mapping of the estimated low redshift ($z < 0.5$) fraction of each Wide SOM cell, as also shown in Fig. 4.

In Section 5, when we characterize the redshift uncertainties in the defined tomographic bins, we will use the BALROG sample to estimate how the tomographic bin photometric spaces map into the Deep SOM, and quantify the (small) impact of deep galaxies in cells with no redshift information.

### 3.2 The Wide SOM

We now turn to characterizing the DES wide space, using the same SOM algorithm as for the deep space. We now use *griz* DES wide photometry as described in Section 2.1 to construct a Wide SOM having $22 \times 22$ cells. By comparison, the Wide SOM describing the DES Year 3 space in Myles et al. (2021) was $32 \times 32$ in size and was constructed using *riz* photometry (because the $g$ band was not used for galaxy selection in the weak lensing analysis). We use a smaller SOM size due to the pre-selection cut-off $22 < i < 23.5$ applied to our wide-field sample. Fig. 3 shows the photometric properties of the Wide SOM, including the mapping of $i$-band magnitude and the three observed colours.

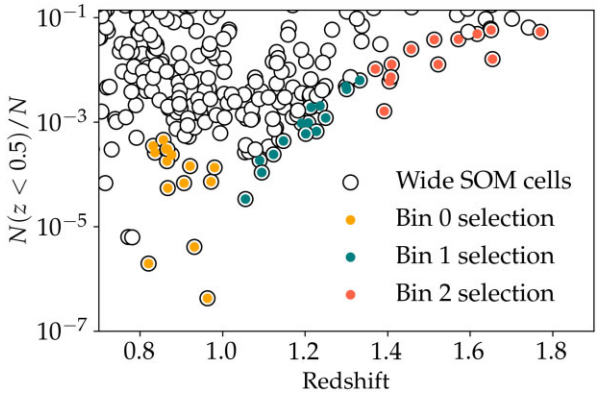


**Figure 4.** Visualization of the tomographic bin selection as groups of Wide-SOM cells, as described in Section 3.3. The plot shows the estimated low redshift ($z < 0.5$) fraction versus mean redshift for each Wide-SOM cell with mean redshift above 0.7. Cells selected for high-$z$ bin 0 are marked in orange, cells for bin 1 are marked in blue, and cells selected for bin 2 are marked in red.

Given the characterization of galaxy phenotypes in the Deep SOM and its redshift mapping using the SPC redshift sample, we can use the BALROG sample to characterize the redshift mapping of the Wide SOM using equation (1). This equation yields a probability density function for the redshift of each Wide SOM cell, using the redshift mapping of the Deep SOM with the SPC redshift sample and the transfer function between Wide and Deep spaces characterized with the BALROG sample. This is shown in the lower left-hand panel of Fig. 3, where we can see a good separation between low- and high-$z$ regions in the Wide SOM, and now we can use this redshift mapping of the Wide SOM to perform the selection of our redshift bins.

### 3.3 Selecting tomographic bins

Since each Wide galaxy can be placed in a cell of the Wide SOM, and we have an estimate of the redshift distribution $p(z|\hat{c})$ within each Wide-SOM cell, we can construct tomographic bins as groups of Wide SOM cells. With the goal of constructing tomographic bins at high redshift with the least possible low-redshift contamination, we compute the mean redshift of each Wide-SOM cell and the fraction of the redshift distribution at low redshift $z < 0.5$. We choose to define three tomographic bins at mean redshifts around 0.9, 1.2, and 1.5 that minimize the low-redshift contamination. We do this by inspecting these properties using Fig. 4, also considering the width of the $p(z|\hat{c})$. For the reproduction of the specific selection used in this paper, see Data Availability section at the end of the paper. The resulting selected cells in the Wide SOM that make up each redshift bin are depicted in the lower right-hand panel of Fig. 3. From that representation, we see how the first redshift bin comes from the upper right-hand part of the Wide SOM and hence contains galaxies with strong (red) $u - g$ and $g - r$ colours, and as the selection moves to the second and third redshift bins the corresponding galaxies will have smaller (blue) $u - g$ colours and fainter $i$-band magnitudes (the average $i$-band magnitude for bins 0, 1, 2 is 22.6, 22.9, and 23.1, respectively). To visualize these trends directly, Fig. 5 shows a small random sample of galaxy images images from each of the redshift bins, which confirm the characteristics of each bin inferred from the Wide SOM in Fig. 3.

It is notable that the Wide-SOM cells $\hat{c}$ selected for the high-$z$ samples largely *exclude* galaxies from cells $c$ in the second 'island'

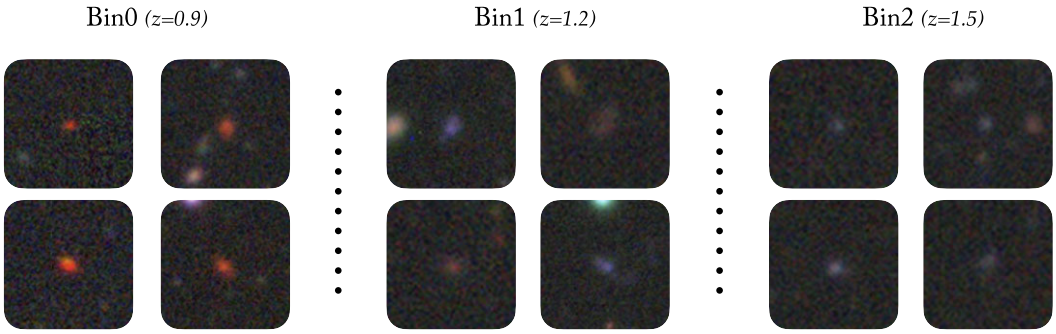


**Figure 5.** Visualization of colour images of random galaxies from each of the three redshift bins defined in Section 3.3. As apparent from Fig. 3, the first bin is made predominantly of red galaxies and then the selection moves to bluer and fainter galaxies for the second and third bin.

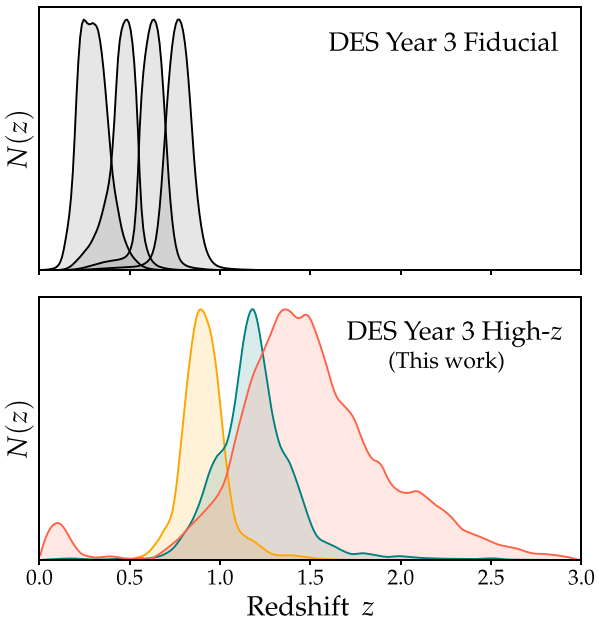


**Figure 6.** Comparison of the redshift distributions used in the fiducial DES Year 3 lens galaxy sample (MAGLIM, upper panel) with the redshift distributions of the three tomographic bins defined in this work (Section 3.3, bottom panel). The three high-$z$ redshift bins defined in this work considerably extend the lens redshift range probed by the DES Year 3 data sample. The number of galaxies, galaxy density, and mean redshift of these samples can be found in Table 1.

of high-$z$ galaxies in the Deep SOM, which contains the Lyman-break galaxies (LBGs). This is likely because the absence of $u$-band data in the wide sample makes it difficult to localize wide-field galaxies into this Deep SOM island. Hence, the DES Y3 high-$z$ sample defined in this paper is notably orthogonal to many previous high-$z$ catalogues that emphasized LBGs at $z > 2$.

Given these tomographic bin selections as lists of Wide SOM cells, we can now use equation (4) to estimate the redshift distribution of each of these bins. Fig. 6 shows the three resulting redshift distributions, and compares them with the four tomographic bins of the fiducial DES Year lens galaxy sample, the so-called MAGLIM sample (Porredon et al. 2022). As apparent from that figure, the three tomographic bins defined in this work significantly extend the redshift range probed by the DES Year 3 Fiducial lens galaxy sample. Besides extending the redshift range, the three tomographic bins from this work also provide larger number of galaxies and galaxy number densities than the MAGLIM fiducial DES lens sample, and also the REDMAGIC galaxy sample (Pandey et al. 2022; see Table 1). The characterization of the uncertainties associated with these three redshift distributions, and the way we will parametrize such uncertainties, will be described in detail in Section 5.

**Table 1.** Summary description of the lens galaxy samples defined using DES Year 3 data, as a comparison to the samples defined in this work. The fiducial lens sample in the DES Year 3 analysis consists of the first four MAGLIM bins. The other two MAGLIM bins and the REDMAGIC sample bins are marked in red as they were not part of the fiducial analysis. The table shows $N_{\rm gal}$ as the number of galaxies in each redshift bin, $n_{\rm gal}$ as the galaxy number density in units of gal arcmin$^{-2}$, and $\langle z \rangle$ as the mean redshift of each bin.

| Redshift bin | $N_{\rm gal}$ | $n_{\rm gal}$ | $\langle z \rangle$ |
|---|---|---|---|
| | DES Year 3 fiducial MAGLIM sample | | |
| 0 | 2236 473 | 0.150 | 0.30 |
| 1 | 1599 500 | 0.107 | 0.46 |
| 2 | 1627 413 | 0.109 | 0.62 |
| 3 | 2175 184 | 0.146 | 0.77 |
| 4 | 1583 686 | 0.106 | 0.89 |
| 5 | 1494 250 | 0.100 | 0.97 |
| | DES Year 3 REDMAGIC sample | | |
| 0 | 330 243 | 0.022 | 0.27 |
| 1 | 571 551 | 0.038 | 0.43 |
| 2 | 872 611 | 0.058 | 0.58 |
| 3 | 442 302 | 0.029 | 0.73 |
| 4 | 377 329 | 0.025 | 0.85 |
| | DES Year 3 high-$z$ sample (this work) | | |
| 0 | 3929 803 | 0.416 | 0.90 |
| 1 | 2551 780 | 0.270 | 1.21 |
| 2 | 2397 667 | 0.254 | 1.49 |

## 4 CHARACTERIZING THE COMPLETENESS OF THE SAMPLES IN THE FOOTPRINT

Due to the faint, low-S/N nature of the galaxies in the three tomographic bins defined in Section 3, it is expected that their selection function will fluctuate across the survey footprint because of varying observing conditions (such as exposure time, seeing, airmass) and also due to astrophysical fluctuations (such as stellar density or extinction). These variations in the selection function will induce correlations between galaxy density and survey properties for the different tomographic bins. Any such fluctuations will induce

spurious signal in the measurement of galaxy clustering, exacerbated by patterns in e.g. survey observing strategies or Galactic structure. We must correct the high-$z$ density maps for the survey selection function if we want to recover accurate measures of the high-$z$ intrinsic galaxy clustering.

These kind of corrections due to varying observing properties have been studied extensively in DES and elsewhere (Leistedt et al. 2016a; Ross et al. 2017; Elvin-Poole et al. 2018; Weaverdyck & Huterer 2021; Rodríguez-Monroy et al. 2022). In many of these cases, the relationship between survey properties and galaxy selection rates was close to linear, and therefore, the correction methodologies assumed a linear relationship. The samples in this work, however, present significant non-linearities in that relationship. Therefore, we use a non-linear, neural-network-based approach for characterizing the completeness of the sample with respect to the different survey properties, as in Rodríguez-Monroy et al. (2022; see also Rezaie et al. 2020 for a similar approach applied to the DECaLS DR7 data sample).

In this section, we describe the different survey properties we consider, the methodology used to correct for their correlations with galaxy density for the different tomographic bins, and the validation of the results. The outcome of this procedure will be a derived correction *weight* for each galaxy in the different tomographic bins, inverse to the selection rate in its vicinity. This weight will then be used throughout the analysis, for the characterization of redshift distributions and uncertainties in Section 5, for the estimation of lens magnification in Section 6, and for the calculation of correlation functions in Section 7.

### 4.1 Maps of survey properties

The DES collaboration develops spatial templates for different observing conditions and potential contaminants in the survey footprint by creating HEALPIX (Gorski et al. 2005) sky maps (at `NSIDE` = 4096, corresponding to a pixel resolution of 0.86 arcmin; see Leistedt, Mortlock & Peiris 2016b for details on the implementation). We will refer to these maps as survey property (SP) maps and we will use them to characterize and remove any possible correlations with the observed density fields of each tomographic bin. In particular, in this analysis we consider maps of the following survey observing properties, each of them having a different map for each observed photometric band *griz*:

(i) Depth: Mean survey depth, computed as the mean magnitude for which galaxies are detected at S/N = 10.
(ii) Sky variance: Estimated sky brightness, or more precisely, the standard deviation of sky pixels due to shot noise and read noise, measured in units of electrons/second/pixel.
(iii) Exposure time: Total exposure time at a given point in the survey footprint, measured in seconds.
(iv) Airmass: Mean airmass, computed as the optical path length for light from a celestial object through Earth's atmosphere (in the secant approximation), relative to that at the zenith for the altitude of the telescope site.
(v) Seeing: Mean seeing, measured in arcseconds, computed as the full-width at half maximum of the flux profile.

Those make 20 SP maps of observing properties. Additionally, we consider two maps of potential contaminants:

(i) Galactic extinction: We use the SFD dust extinction map from Schlegel, Finkbeiner & Davis (1998), which measures the $E(B-V)$ reddening, in magnitudes.
(ii) Stellar density: We use a map of stellar density, in deg$^{-2}$, using stellar sources from *Gaia* EDR3 (Gaia Collaboration 2021).

This amounts to a total of 22 survey property maps that we will use in this analysis. For a technical description of these survey observing properties, please see Leistedt et al. (2016b), Sevilla-Noarbe et al. (2021), and Rodríguez-Monroy et al. (2022). In principle, these SPs should be a complete list of all factors that could affect galaxy detectability. The images themselves should be completely specified by the passband (which is constant, with very minor airmass variation), the background noise level of the images (a.k.a. sky brightness), the point spread function (primarily seeing FWHM), and the shot noise from the sources (primarily exposure time). The Galactic dust and stellar background are the two astrophysical effects expected to alter the detectability of background galaxies. The depth map should be redundant but we include it to perhaps ease the task of training the neural network.

### 4.2 Correction method

We aim to model the relationship between the survey property maps defined above and the observed galaxy count maps for each of the tomographic bins defined in Section 3. For this, we will use a neural network (NN), with the 22 SP maps being the *features* and the observed galaxy count maps being the *label*. Naturally, the network will be able to model a non-linear relationship between the SP maps and the raw galaxy counts. It is important to note, however, that we do not include any spatial information in the process, since we do not want the network to learn about the clustering of galaxies.

The neural network is asked to predict whether or not a particular Healpixel (at the same `NSIDE` = 4096 resolution) contains *any* galaxies [that is, $p(n \geq 1)$] based on the SP values for that pixel. The loss function for the network is the binary cross-entropy between the predicted pixel occupancy and the occupancy of the training set. Note that this ignores any distinctions between Healpixels with $n = 1$ versus $n = 2$ or more galaxies. This helps prevent the network from learning any intrinsic galaxy clustering. At the resolution of `NSIDE` = 4096, most pixels contain either zero or one galaxies (the average number of galaxies per pixel for bins 0, 1, and 2 is 0.307, 0.200, and 0.187, respectively). The number of pixels with $p(n \geq 2)$ in each tomographic bin is 4.6 per cent, 2.0 per cent, and 1.7 per cent, while the number expected from a uniform density and a Poisson distribution is 3.9 per cent, 1.7 per cent, and 1.6 per cent, which represents an excess due to intrinsic clustering of 20 per cent, 16 per cent, and 10 per cent, respectively, while the usual change in density imprinted by survey property maps in our conservative mask is usually under 5 per cent (Fig. 8). On the other hand, pixels with large number of galaxies are likely to be due to intrinsic density peaks rather than survey observing property fluctuations, for example $p(n \geq 5)$ in our samples exceeds the Poisson uniform probability by 20×, 7×, and 3×, respectively.

The architecture of the network is based on our guess that the selection function scales primarily as some power-law combination of the SPs. To this end, the input SP values are all logarithmically scaled (except those, such as depth, which are already logarithmic quantities), and the output of the network is exponentiated to form the selection probability. The network output is a sum of two branches: the first branch is a simple linear combination of the 22 scaled SPs, since we expect this to capture most of the functional variation. The second branch is intended to capture departures from a simple power law: it takes the input layer of 22 dimensions through 3 hidden layers of 64, 32, and 4 fully connected neurons, respectively, and a single

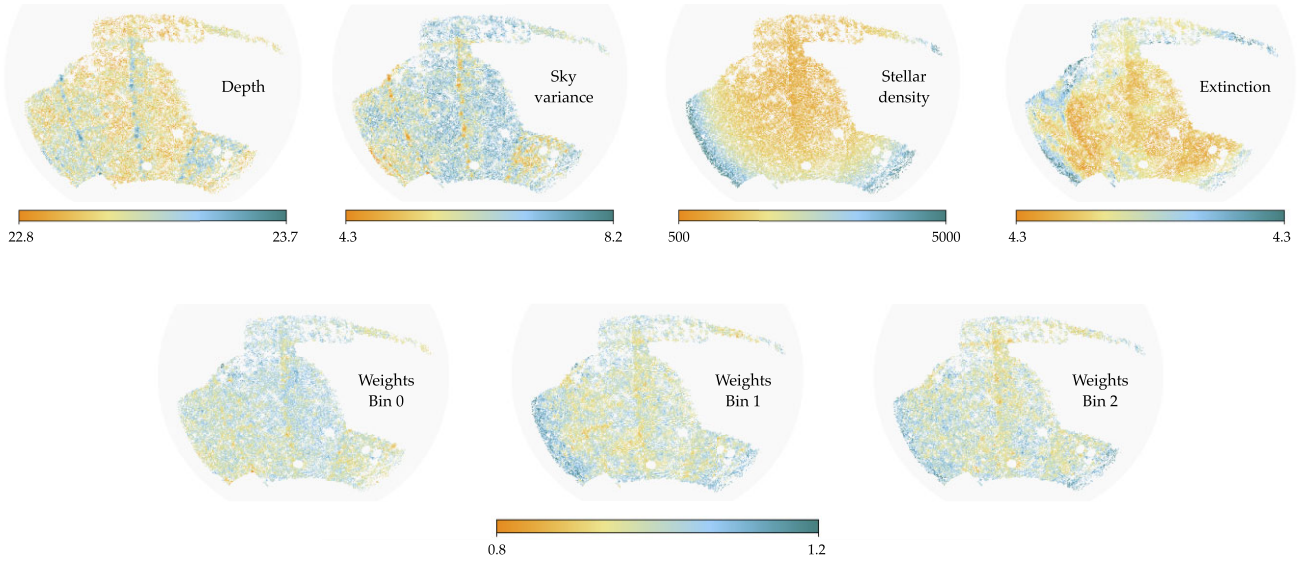


**Figure 7.** *Upper row:* Examples of four of the SP maps described in Section 4.1. In particular, we show the depth and sky variance maps in the *i* band, and the maps of stellar density and dust extinction. *Lower row:* Maps of the derived weight maps using the neural network approach described in Section 4.2, for the three tomographic bins in this work.

neuron on the output layer, each with `relu` activation. The output of the network, for each tomographic bin, consists of a single value for each Healpixel within our mask, which will be used to weight the galaxies accordingly. Fig. 7 shows the resulting weight maps for each tomographic bin, as well as four examples of survey property maps.

To prevent the network from overfitting, it is constructed with *k*-fold cross-validation, which works in the following way: The `NSIDE` = 4096 maps are re-binned into a coarser grid of `NSIDE_split` = 16 (with a resolution of about 4 degrees). We then randomly divide these cells into *k* equal-area groups. To derive the weights for a given fold `k`, we train the NN on the other folds, using fold `k` as a validation sample (the training halts when the training metric no longer improves on the validation set). This cross-validation scheme will only work to prevent overfitting on scales below the resolution defined by `NSIDE_split`, in this case around 4 degrees. In Appendix C, we test the method using unaltered simulated data and find any residual overfitting to be small compared to the statistical uncertainty on angular scales below 1 degree. Therefore, we keep the galaxy clustering analysis in this work to angular scales below one degree, and additionally test the robustness of the maximum angular scale in Section 7.3.3.

### 4.3 Validation of the derived correction weights

Different survey property maps show significant correlations with the raw galaxy density in each of the tomographic bins. Using the neural network implementation described above, Fig. 8 shows these correlations, and how the derived set of weights is able to correct for any correlations between SP maps and galaxy density. Fig. 8 shows only a limited number of examples of these correlations, for easier visualization, but we also compute the $\chi^2$ for the null hypothesis for all correlations between the 22 SP maps and the corrected galaxy density, using a jackknife approach to estimate the corresponding uncertainties. The distribution of these null $\chi^2$ values, for each of the tomographic bins, can be found in Fig. 9, and we do not find evidence of significant correlations between the SP maps and the corrected (weighted) galaxy density. The median null $\chi^2$ values for the corrected case in the three tomographic bins are 11.6, 3.4, and 7.5 for 10 degrees of freedom. On the other hand, for the raw, uncorrected case the median null $\chi^2$ values for the three bins are 92.1, 35.0, and 51.6 for 10 degrees of freedom, clearly inconsistent with the null hypothesis.

Beyond being successful at correcting for all the correlations between galaxy density and survey property maps, we need to ensure the derived neural network weights did not learn any physical galaxy clustering at the training phase. For that purpose, we compute the cross-correlation between the weight maps as shown in Fig. 7 and several tracers of the LSS of the Universe. In particular, in this work we perform the correlation of the three weight maps with the convergence field estimated from CMB lensing (using both Planck, Planck Collaboration VIII 2020; and SPT, Omori et al. 2023), the high-$z$ mass map from the DES Year 3 analysis (Jeffrey et al. 2021) and the Planck Compton $y$ map (Planck Collaboration XXII 2016). These are all tracers of the physical LSS and hence they should not present correlations with SP maps or the derived weight maps. A significant correlation would mean there has been some undesired leakage of LSS into our weights. Fig. 10 shows these correlations between weight maps and tracers of the LSS, and Table 2 shows the $\chi^2$ values for the null hypothesis, demonstrating no significant correlations between weight maps and LSS tracers.

At this point, we have now tested for the correlation of the weighted galaxy density with SP maps and the correlation of weight maps with known tracers of structure, and found a null signal in both cases. However, it is still possible that the residuals in the estimation of the weight maps could affect the clustering measurements. To account for this potential effect in the clustering analysis, we will marginalize over an additive constant in the correlation function, as done in e.g. Kwan, Sánchez et al. (2017; see also Ross et al. 2011). This procedure, which will be described in Section 7.3, will account for a potential spurious systematic effect in the clustering at first order, and it is a conservative way to marginalize over this uncertainty in the analysis. In that section, we will also explore the impact of the choice of maximum angular scale in the galaxy clustering measurements.

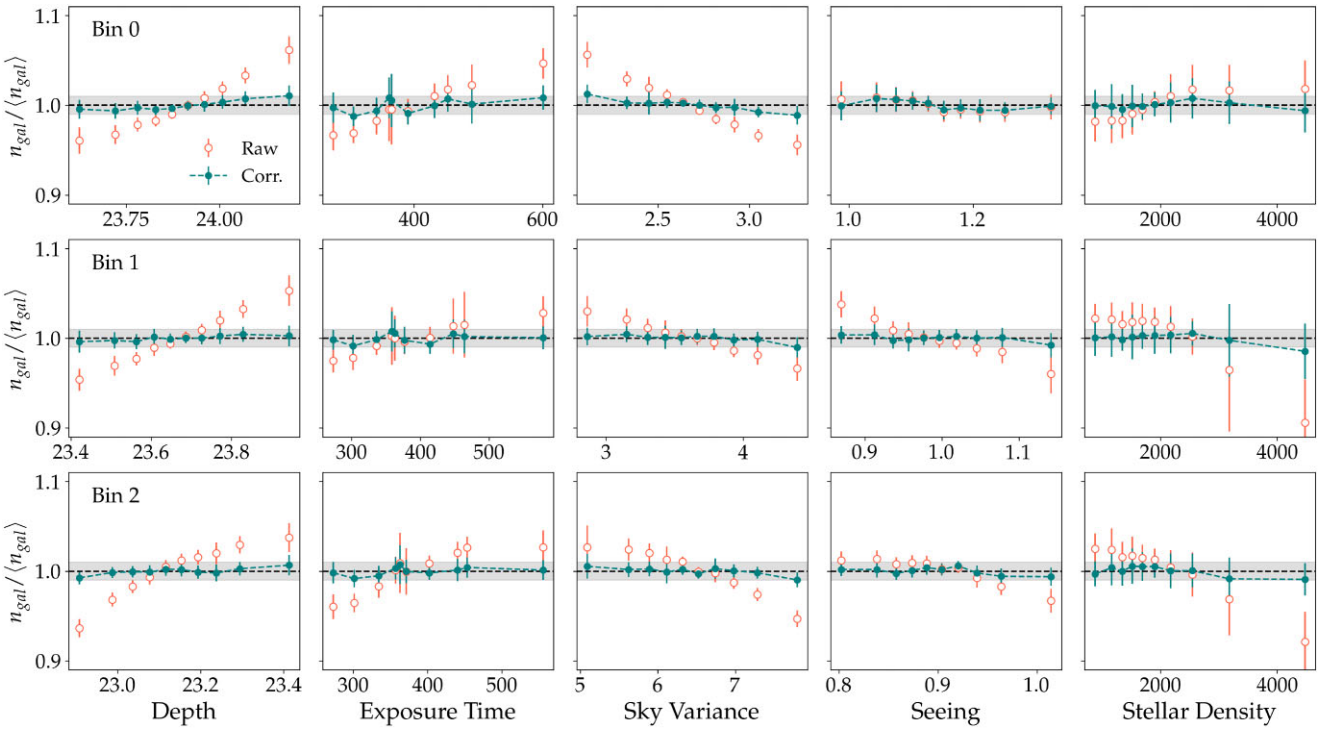


**Figure 8.** Visualization of the correlation between SPs and the observed galaxy density (relative to the mean galaxy density over the full footprint), before (red) and after (blue) the correction using the galaxy weights described in Section 4.2. We show this relationship for depth, exposure time (in seconds), sky variance (in electrons $s^{-1}$ $pixel^{-1}$) and seeing (in arcseconds), all estimated in the *i* band, and also with stellar density (in stars $deg^{-2}$), in 10 bins of equal area. The uncertainties come from jackknife resampling, and the grey-shaded region in the plot corresponds to a 1 per cent deviation. The distribution of the null $\chi^2$ values for these relationships, including all the 22 SP maps and for each of the tomographic bins, can be found in Fig. 9.

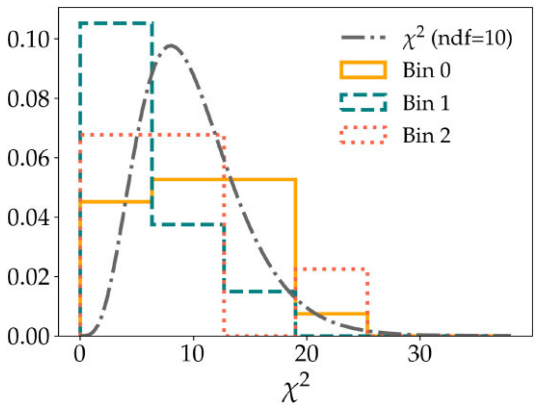


**Figure 9.** Distribution of the null hypothesis $\chi^2$ values for the relationship between survey property maps and the corrected (weighted) galaxy density, including all the 22 SP maps and for each of the tomographic bins. The median null $\chi^2$ values in the three tomographic bins are 11.6, 3.4, and 7.5 for 10 degrees of freedom. For the raw, uncorrected case the median null $\chi^2$ values for the three bins are 92.1, 35.0, and 51.6 for 10 degrees of freedom.

## 5 CHARACTERIZING REDSHIFT UNCERTAINTIES

In this section, we will describe the various sources of uncertainty in the distributions of redshift $N(z)$ within each of the three bins defined in Section 3.3, and how we will propagate them into cosmological analyses. We will follow a similar procedure to that in Myles et al. (2021), and propagate uncertainty arising from (i) sample variance (SV) and shot noise (SN) from the finite area covered by the deep fields; (ii) biases in the individual redshift estimates of deep-field galaxies having multiband photometry (COSMOS2015 and PAUS + COSMOS) but no spectroscopic redshift (PZ); (iii) uncertainty in the photometric calibration (zero-point) of deep-field galaxies (ZP); and (iv) uncertainties from the 'bin conditionalization' approximation in equation (A6) (BCE).

To model SV and SN, we use the approximate 3sDir model (a product of three Dirichlet distributions), first presented in Sánchez et al. (2020) and then further developed in Myles et al. (2021). Mathematically, the model describes $p(\{f_{zc}\}|\{N_{zc}\}) \approx$ 3sDir, where $N_{zc}$ are the number counts of galaxies that have been observed to be at redshift bin $z$ and colour phenotype $c$, and with $\{f_{zc}\}$ a finite set of coefficients indicating the probability in the redshift bin $z$ and colour phenotype $c$, where $\sum_{zc} f_{zc} = 1$ and $0 \leq f_{zc} \leq 1$. For extensive details of the model, we refer the interested reader to appendices D and E in Myles et al. (2021). The 3sDir method yields realizations of the $f_{zc}$, which then can be summed into equation (A6) to yield $N(z)$ estimates. The mean of these realizations is the fiducial $N(z)$.

We smooth the fiducial $N(z)$ distribution with a Savitzky–Golay filter: sample variance and shot noise from the small area of the calibration deep fields manifests in the $N(z)$ as rapid fluctuations in redshift and enter squared in the galaxy clustering signal, while the true redshift distribution over a larger area is smoother as these variations average out. We try different smoothing lengths and find compatible constraints on the main parameters of interest (see Appendix A6).

Deviations from the nominal $N_i(z)$ will be modelled with three parameters: a shift $\Delta z^i$, a stretch parameter $\sigma_z^i$, and an adjustment $A^i_{\mathrm{low}-z}$ of the low-redshift tail of $N_i(z)$. The main peak of the

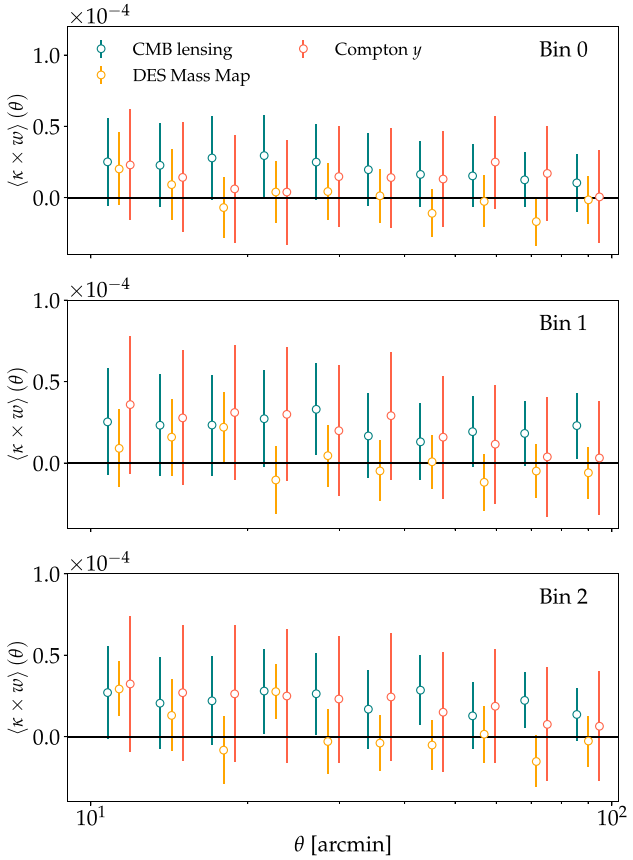


**Figure 10.** Cross-correlation of weight maps of the three tomographic bins and convergence field from three tracers of the LSS: CMB lensing (from the Planck satellite), DES Y3 mass maps, and Planck Compton $y$. Uncertainties come from jackknife resampling. The null hypothesis $\chi^2$ values can be found in Table 2, all consistent with no correlation.

**Table 2.** Values of $\chi^2_{\rm null}$/dof for different correlations between galaxy weights and tracers of the LSS of the Universe, for the three high-$z$ bins defined in this work. We find no significant correlations between weight maps and LSS tracers.

| | Bin0 | Bin1 | Bin2 |
|---|---|---|---|
| Planck CMB lensing | 9.6/9 | 6.4/10 | 6.8/10 |
| DES mass map | 8.2/9 | 9.9/10 | 16.1/10 |
| Planck Compton $y$ | 7.3/9 | 5.9/10 | 2.7/10 |

distribution is altered according to

$$N(z) \to N(\sigma_z^i(z - \Delta z^i - \bar{z}) + \bar{z}) \tag{5}$$

and the fraction of galaxies at low redshift ($z < 0.5$) is altered as

$$n(z) \to \begin{cases} n(z)\, A^i_{{\rm low}-z} & z \le 0.5 \\ n(z)\,(1 - A^i_{{\rm low}-z}) & z > 0.5 \end{cases} \tag{6}$$

Details of this transformation are in Appendix A1. Fig. 11 illustrates the effects of each of these parameters.

Priors on the $N(z)$ alteration parameters $\theta_i = \{\Delta z^i, \sigma_z^i, A^i_{{\rm low}-z}\}$ are chosen to represent the potential effects of the systematic errors by

(i) Quantifying the possible effects of the PZ, BCE, and ZP systematic errors on the input catalogues to the redshift calibration process, as detailed in Appendices A2, A3, and A4, respectively.

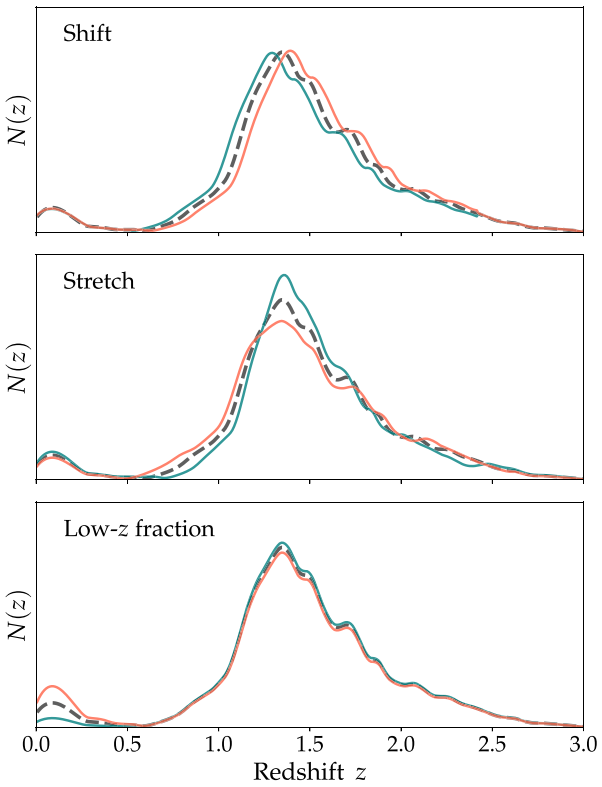


**Figure 11.** Visualization of the parametrization of redshift uncertainties, using the third tomographic bin as an example. The different rows show examples of how we account for shifts, stretches and variations in the low redshift fraction of the redshift distributions.

(ii) Creating realizations of the input catalogues drawing from these systematic errors and realizing the SV and SN variations with the 3sDir process.

(iii) Measuring the mean, width, and low-$z$ fractions of each realized $N_i(z)$.

(iv) Creating a prior for the $\theta_i$ based on the distribution of these properties of the realizations.

Fig. 12 shows the resultant distributions of the $N_i(z)$ recalibration parameters when various sources of systematic errors are included, and values of their means and standard deviations are listed in Table 3. Sample variance/shot noise, redshift biases and zero-point uncertainty all contribute significantly to the uncertainty in the mean redshift. On the other hand, the stretch uncertainty is dominated by sample variance at low redshift (Bin 0), with the zero-point uncertainty significantly increasing its importance in in the highest redshift bin. Finally, the low-redshift probability uncertainty is primarily dominated by sample variance and shot noise. Similar results for redshift uncertainties are found from the North and South subsets of the data.

## 6 CHARACTERIZING WEAK LENSING MAGNIFICATION

In this section, we study the impact of lensing magnification on the observed angular correlations of our high-$z$ galaxy samples. On top of distorting the image shapes, gravitational lensing from the foreground large-scale structure of the Universe also magnifies the images without changing the surface brightness, creating two effects:

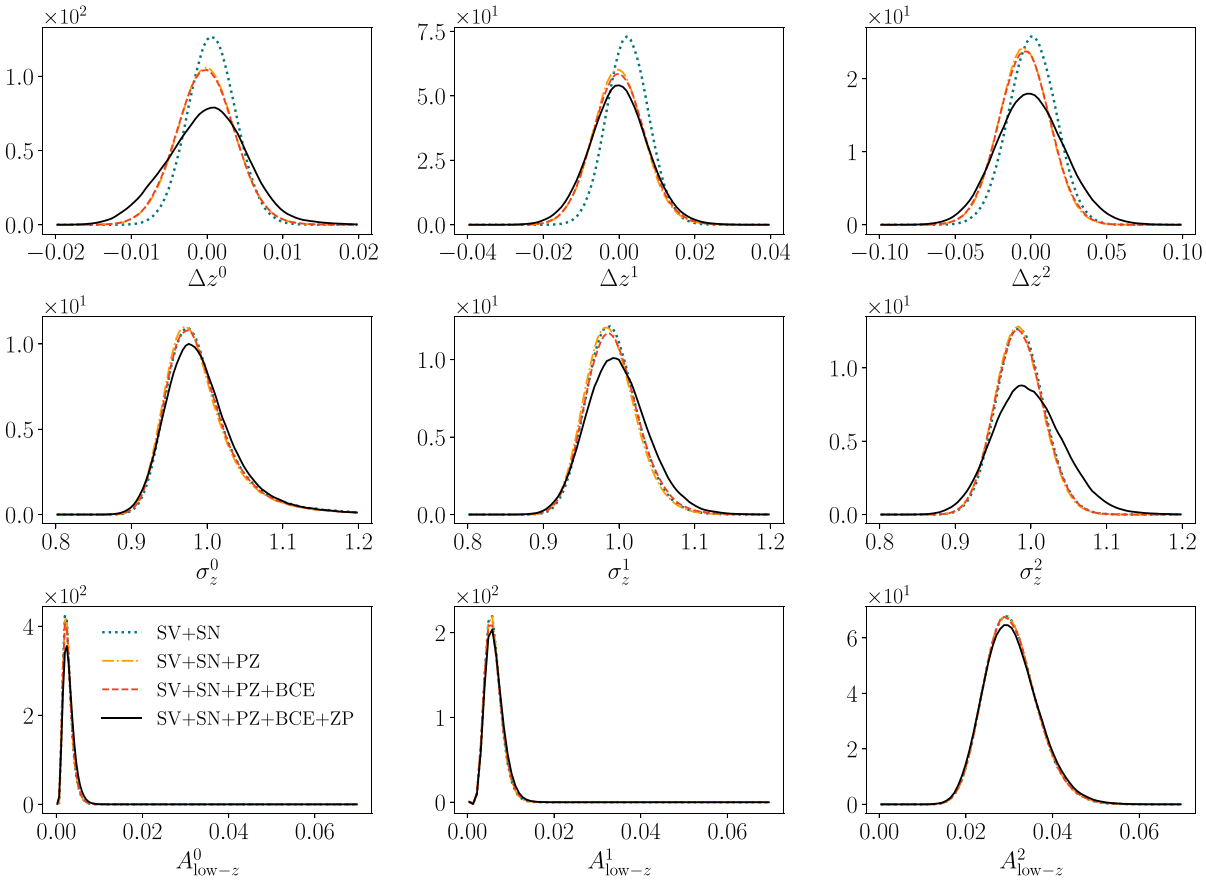


**Figure 12.** Prior distributions for each redshift uncertainty parameter. Each column shows the parameters for each tomographic bin (left: Bin 0; middle: Bin 1, right: Bin 2). Each row shows a different parameter (top: $\Delta z^i$; centre: $\sigma_z^i$, bottom: $A^i_{\mathrm{low}-z}$). The different lines show the cumulative uncertainty on each parameter from considering different effects. The dotted line shows the uncertainty from Sample Variance and Shot Noise in the calibration fields (SV + SN). The dot–dashed line adds the uncertainty from redshift biases in the redshift calibration samples (PZ). The dashed line adds uncertainty from redshift selection effects (BCE). The solid lines adds the zero-point photometric uncertainty in the deep field photometry (ZP). The distributions are measured from individual $N(z)$ samples generated to include these uncertainties. For $p(\Delta z^i)$, we measure the mean redshift of individual samples and subtract the mean redshift of the fiducial $N(z)$. For $p(\sigma_z^i)$ we measure the $N(z)$ width of individual samples and divide by the width of the fiducial $N(z)$. For $p(A^i_{\mathrm{low}-z})$, we measure the integral of each individual sample at $z < 0.5$. See Section 5 and Appendix A for details.

**Table 3.** Estimates of the parameters describing our uncertainties on the redshift distributions, as described in Section 5, for the three tomographic bins defined in this work. The parametrization is described visually in Fig. 11. We also show the estimates for the entire footprint we use, and for the independent splits of North and South regions, which will be used in Section 7 for consistency tests.

| $z$-bin | $\Delta z$ | $\sigma_z$ | $A_{\mathrm{low}-z}$ |
|---|---|---|---|
| | Entire footprint (All) | | |
| 0 | $0.0 \pm 0.0051$ | $0.997 \pm 0.068$ | $0.0044 \pm 0.0013$ |
| 1 | $0.0 \pm 0.0075$ | $0.999 \pm 0.041$ | $0.0091 \pm 0.0023$ |
| 2 | $0.0 \pm 0.0208$ | $0.998 \pm 0.044$ | $0.0383 \pm 0.0.0059$ |
| | North region (*Planck*) | | |
| 0 | $0.0 \pm 0.0054$ | $0.995 \pm 0.068$ | $0.0043 \pm 0.0015$ |
| 1 | $0.0 \pm 0.0078$ | $0.999 \pm 0.041$ | $0.008 \pm 0.0023$ |
| 2 | $0.0 \pm 0.0223$ | $0.998 \pm 0.044$ | $0.038 \pm 0.0.0065$ |
| | South region (SPT) | | |
| 0 | $0.0 \pm 0.0052$ | $0.998 \pm 0.051$ | $0.0041 \pm 0.0015$ |
| 1 | $0.0 \pm 0.0114$ | $0.996 \pm 0.081$ | $0.009 \pm 0.0027$ |
| 2 | $0.0 \pm 0.0224$ | $0.998 \pm 0.048$ | $0.0337 \pm 0.0.0065$ |

(i) a dilution of the source density due to the locally stretched image; and (ii) an increased flux of individual galaxies making them more likely to be detected (Bartelmann & Schneider 2001; Ménard et al. 2003; Hildebrandt, van Waerbeke & Erben 2009; Garcia-Fernandez et al. 2018; Gaztanaga et al. 2021; von Wietersheim-Kramsta et al. 2021; Euclid Collaboration 2022). This effect creates an additional clustering signal of the background sample that contaminates estimates of its intrinsic density fluctuations. Following the approach used in the fiducial DES Y3 analysis (DES Collaboration 2022), we model the observed projected (lens) galaxy density contrast of tomographic bin $i$, $\delta^i_{\mathrm{obs}}$, as a combination of the projected galaxy density contrast $\delta^i_{\mathrm{g}}$ and the modulation by lens magnification $\delta^i_\mu$ and redshift-space distorsions (see Section 7.1 for more details):

The change in density contrast due to magnification can be shown to be proportional to the convergence experienced by the lens galaxies $\kappa^i_l$ (Elvin-Poole et al. 2023):

$$\delta^i_\mu(\theta) = C^i \kappa^i_l(\theta). \tag{7}$$

The constant of proportionality $C^i$ is given by the response of the number of selected galaxies per unlensed area, and it can be split in two terms, one fixed term corresponding to the change of area and another term corresponding to changes in the light flux distribution

of galaxies, which will affect their selection in different samples:

$$C^i = C_{\rm area} + C^i_{\rm sample}, \tag{8}$$

where $C_{\rm area} = -2$ regardless of the sample selection. In this way, the characterization of lens magnification amounts to estimating the $C^i_{\rm sample}$ term for each tomographic bin. This term can be estimated empirically by artificially magnifying a galaxy sample and measuring the change in number density with respect to the applied magnification. In particular, if we apply some extra convergence $\delta\kappa$ to the images, the proportionality constant can be written as

$$C_{\rm sample} = \frac{\delta n}{n\,\delta\kappa}, \tag{9}$$

where $\delta n/n$ corresponds to the fractional change in number density of a given sample meeting selection criteria due to the applied magnification. In this work, we will follow the approach of Elvin-Poole et al. (2023) and estimate $C_{\rm sample}$ in two different ways, using the BALROG sample and directly perturbing the measured fluxes in the data.

### 6.1 Estimate from artificial galaxy injections

A number of BALROG catalogues were produced for the DES Year 3 analysis (Everett et al. 2022). In this analysis, we have already used BALROG to estimate the transfer function between the deep and wide photometric spaces (parametrized with SOMs), as described in Section 3. In this part, we use an additional BALROG run, in which the exact same deep field objects are injected at the same coordinates as in the fiducial run, but now with a 2 per cent magnification applied to each galaxy image, $\mu_0 = 1.02$ ($\kappa_0 \sim 0.01$). For all cases, we account for the galaxy correction weights defined in Section 4 and shown in Fig. 7.

We apply the tomographic bin selections described in Section 3.3 on both the fiducial $\kappa = 0$ BALROG run (label $i$, for *intrinsic*) and the $\kappa = \kappa_0$ run (label $o$, for *observed*). In order to estimate $C_{\rm sample}$, we need, for each tomographic bin selection:

(i) $N_i$: Selected number of galaxies in the BALROG $\kappa = 0$ run. Accounting for galaxy weights $w^j_i$, it becomes $N_i = \sum_j w^j_i$, where $j$ runs over all selected galaxies.

(ii) $N_o$: Selected number of galaxies in the magnified BALROG run, which applies a constant magnification to the galaxy images. Accounting for galaxy weights $w^j_o$, it becomes $N_o = \sum_j w^j_o$.

At this point, the estimate is simply the fractional difference between the two:

$$C_{\rm sample} = \frac{N_o - N_i}{\kappa_0 N_i}. \tag{10}$$

This estimate should capture the impact of magnification on the specific colour selection of the high-$z$ bins defined in Section 3.3, and also include possible contributions due to size selections such as the star – galaxy separation cuts. We compute the uncertainties on these estimates by following a jackknife approach, splitting the footprint over 150 regions.

### 6.2 Estimate from perturbing measured fluxes

The second method we consider uses the data itself to estimate the flux gradient of the samples. In this case, we add a constant offset $\Delta m$ to all photometric magnitudes in our sample:

$$\Delta m = -2.5\log_{10}(1 + 2\Delta\kappa), \tag{11}$$

where $\Delta\kappa = 0.01$ is the constant magnification difference we are applying to each galaxy.

Using this new *magnified* data sample, we repeat the assignment of the detected galaxies to the three high-$z$ bins, and estimate $C_{\rm sample}$ from the differential in the resultant counts in each bin, directly from equation (9), again accounting for individual galaxy weights from Section 4. This method provides an additional estimate of the magnification coefficients using only the magnification effect on the fluxes, hence ignoring other possible contributions from size selection or observational systematics.

**Table 4.** Estimates of the lens magnification coefficients $C_{\rm sample}$ using the BALROG and data-based methods described in Section 6, for the three tomographic bins defined in this work. The last column shows the final estimates of the coefficients from the combination of the two different methods. We also show the estimates for the entire footprint we use, and for the independent splits of North and South regions, which will be used in Section 7 for consistency tests.

| $z$-bin | $C^{\rm Data}_{\rm sample}$ | $C^{\rm Balrog}_{\rm sample}$ | $C^{\rm Final}_{\rm sample}$ |
|---|---|---|---|
| | Entire footprint (All) | | |
| 0 | $-0.21 \pm 0.03$ | $0.32 \pm 0.40$ | $0.05 \pm 0.48$ |
| 1 | $2.20 \pm 0.04$ | $3.02 \pm 0.63$ | $2.61 \pm 0.75$ |
| 2 | $3.88 \pm 0.04$ | $4.70 \pm 0.59$ | $4.29 \pm 0.72$ |
| | North region (*Planck*) | | |
| 0 | $-0.19 \pm 0.03$ | $0.29 \pm 0.46$ | $0.05 \pm 0.52$ |
| 1 | $2.15 \pm 0.04$ | $2.67 \pm 0.66$ | $2.41 \pm 0.71$ |
| 2 | $3.79 \pm 0.05$ | $4.85 \pm 0.65$ | $4.32 \pm 0.83$ |
| | South region (SPT) | | |
| 0 | $-0.23 \pm 0.04$ | $0.34 \pm 0.43$ | $0.05 \pm 0.52$ |
| 1 | $2.23 \pm 0.06$ | $3.33 \pm 0.97$ | $2.78 \pm 1.12$ |
| 2 | $3.95 \pm 0.04$ | $4.54 \pm 0.69$ | $4.25 \pm 0.75$ |

### 6.3 Results

Table 4 shows the estimates of $C_{\rm sample}$ using the BALROG and data-based methods described above, for the three tomographic bins and the North and South regions defined in this work. Since we have two independent methods to estimate these values, we use the average of the two methods as our final estimates $C^{\rm Final}_{\rm sample}$. For the associated uncertainties, we follow a conservative approach and add the uncertainties of the methods in quadrature, in addition to the standard deviation between the methods:

$$\sigma^{\rm Final}_C = \sqrt{\left(\sigma^{\rm Balrog}_C\right)^2 + \left(\sigma^{\rm Data}_C\right)^2 + (C^{\rm Balrog}_{\rm sample} - C^{\rm Data}_{\rm sample})^2/4}. \tag{12}$$

The derived magnification coefficients and their associated uncertainties will be used as Gaussian priors in the galaxy clustering analysis presented in the next section.

## 7 GALAXY CLUSTERING AND CONSTRAINTS ON COSMOLOGY AND GALAXY BIAS

In this section, we present the analysis of galaxy clustering in the tomographic bins defined in this work. We describe the model we use, the choice of scales, the measurements and covariance, and finally the constraints we obtain on the cosmological model and the galaxy bias of each tomographic bin, and their robustness under different analysis choices.

### 7.1 Model

Following the galaxy clustering analysis of the DES Year 3 fiducial sample (Rodríguez-Monroy et al. 2022), we model the observed projected galaxy density contrast $\delta^i_{\rm obs}(\hat{\boldsymbol{n}})$ of galaxies in tomography bin $i$ at position $\hat{\boldsymbol{n}}$ as

$$\delta^i_{g,\rm obs}(\hat{\boldsymbol{n}}) = \delta^i_{g,\rm D}(\hat{\boldsymbol{n}}) + \delta^i_{g,\rm RSD}(\hat{\boldsymbol{n}}) + \delta^i_{g,\mu}(\hat{\boldsymbol{n}})\,. \tag{13}$$

The first term is the line-of-sight projection of the three-dimensional galaxy density contrast, $\delta_g^{(\rm 3D)}$; the other terms correspond the contributions from linear redshift-space distortions (RSD) and magnification ($\mu$), which are described in detail in Krause et al. (2021). We relate the galaxy density to the matter density assuming a local, linear galaxy bias model (Fry & Gaztanaga 1993), $\delta_g(\boldsymbol{x}) = b\delta_m(\boldsymbol{x})$, with $\delta_Y \equiv (Y(\mathbf{x}) - \bar{Y})/\bar{Y}$. We assume the galaxy bias to be constant across each tomographic bin $b^i$, and we discuss more about this assumption later in this section.

Given the three terms in equation (13), the angular power spectrum $C^{ii}_{\delta_{g,\rm obs}\delta_{g,\rm obs}}(\ell)$ has six different components, corresponding to the auto- and cross-power spectra of galaxy density, RSD, and magnification. For the accuracy of the DES Year 3 analysis, it was shown by Krause et al. (2021) that the commonly used Limber approximation is insufficient to estimate these terms, and therefore we use the non-Limber algorithm of Fang et al. (2020).[6] Using the full expressions for the angular power spectrum, including RSD and magnification, from Fang et al. (2020), the angular correlation function is given by

$$w^i(\theta) = \sum_\ell \tfrac{2\ell+1}{4\pi} P_\ell(\cos\theta) C^{ii}_{\delta_{g,\rm obs}\delta_{g,\rm obs}}(\ell)\,, \tag{14}$$

where $P_\ell$ are the Legendre polynomials. For the implementation of these calculations, we use the CosmoSIS framework[7] (Zuntz et al. 2015), which in turn uses CAMB (Lewis & Bridle 2002) to obtain the evolution of linear density fluctuations and Halofit (Takahashi et al. 2012) to convert to a non-linear matter power spectrum. The modelling of redshift uncertainties has been described in detail in Section 5, and that parametrization has been implemented in CosmoSIS for this analysis.

In addition, as explained in Section 4, we marginalize over an additive constant parameter, parametrized by $R^i$, in the galaxy angular correlation function:

$$w^i(\theta) \to w^i(\theta) + 10^{R^i}\,. \tag{15}$$

This parametrization accounts for potential residuals in the calculation of galaxy weights affecting the galaxy clustering measurements (Kwan et al. 2017). Later in Section 7.3, we will explore the impact of the choice of maximum angular scale in the galaxy clustering measurements.

#### 7.1.1 *Choice of scales*

Given the fact that we assume a linear galaxy bias model for this analysis, we are required to remove small-scale information that can potentially be affected by non-linearities. We follow the approach of the DES Year 3 fiducial analysis (DES Collaboration 2022) and we remove all galaxy clustering information below 8 $h^{-1}$ Mpc (Krause et al. 2021) (corresponding to a minimum angular scale of 12.9, 10.5, and 9.0 arcmin for the three tomographic bins in this work, respectively). We also test for the robustness of the results to a minimum scale of 12 $h^{-1}$ Mpc. The maximum angular scale we use is set to 60 arcmin for all measurements. This choice is driven by the correction method of obtaining galaxy weights, described in Section 4, in particular by the cross-validation scheme to avoid overfitting, which shows no signs of overfitting at angular scales below 1 degree.

### 7.2 Measurements and covariance

Equation (14) shows the modeling of the galaxy angular 2-point correlation function, $w(\theta)$. For the measurement of this galaxy clustering observable, we use Healpix maps (`nside` = 4096) of the corrected galaxy density contrast for each tomographic bin, including the correction weights described in Section 4, and then use a pixel-based version of the Landy–Szalay estimator (Landy & Szalay 1993), following the notation of Crocce et al. (2016):

$$\hat{w}(\theta) = \sum_{i=1}^{N_{pix}} \sum_{j=1}^{N_{pix}} \frac{(N_i - \bar{N})\cdot(N_j - \bar{N})}{\bar{N}^2}\, \omega_i\, \omega_j\, \Theta_{i,j}(\theta)\,, \tag{16}$$

where $N_i$ is the galaxy number density in pixel $i$, and $\omega_i$ is the weight of each pixel $i$ (see Section 4). $\bar{N}$ is the corrected mean galaxy number density over all pixels within the footprint and $\Theta_{i,j}$ is a top-hat function which is equal to 1 when pixels $i$ and $j$ are separated by an angle $\theta$ within the bin size $\Delta\theta$. In practice, these correlation functions are computed using TreeCorr[8] (Jarvis, Bernstein & Jain 2004). Fig. 13 shows the $w(\theta)$ measurements for the galaxy autocorrelations of the three redshift bins considered in this work.

We estimate the covariance matrices using two complementary methods: using Gaussian simulations, and using Jackknife. The Gaussian simulations are generated following the procedure described in Giannantonio et al. (2008; see Appendix B for details). We generate 100 realizations of a set of four correlated maps via Healpix anafast routine. These maps, three for galaxy overdensity and one for CMB $\kappa$, are generated using the non-linear (Halofit) power spectrum with our fiducial cosmology. Each map includes its respective (uncorrelated) noise contribution. The advantage of this simulation-based approach is that it allows us to have an accurate estimation of the effects of the mask, and angular binning. The main downside is that this approach does not account for the non-Gaussian terms of the covariance. In order to cross-check the validity of this approach, we also estimate the covariance using the Jackknife technique, defining 150 subsamples for the measurements in TreeCorr. We find that both approaches are in good agreement within the range of scales used for this work, pointing to a negligible contribution of the non-Gaussian terms for this particular study. A detailed comparison can be found in Appendix B.

Defining these data measurements as $\hat{\boldsymbol{D}} \equiv \{\hat{w}^{ij}(\theta)\}$ and the covariance $\boldsymbol{C}$, we use the following expression to compute the signal to noise of the measurements:

$$\mathrm{S/N} = \sqrt{\hat{\boldsymbol{D}}\, \boldsymbol{C}^{-1}\, \hat{\boldsymbol{D}}^T - n\mathrm{d}f}, \tag{17}$$

where $ndf$ is the number of degrees of freedom, which equals the number of data points passing the scale cuts defined in Section 7.1.1. For reference, the fiducial DES Year 3 analysis had a galaxy clustering S/N = 63 (Rodríguez-Monroy et al. 2022). For the sample in this work, the total S/N, including the three autocorrelations after applying scale cuts, is S/N = 70. Breaking this into the individual measurements, the autocorrelations for bins 0, 1, and 2 get S/N = 43, 49, and 37, respectively.

[6] https://github.com/xfangcosmo/FFTLog-and-beyond
[7] https://bitbucket.org/joezuntz/cosmosis
[8] https://rmjarvis.github.io/TreeCorr

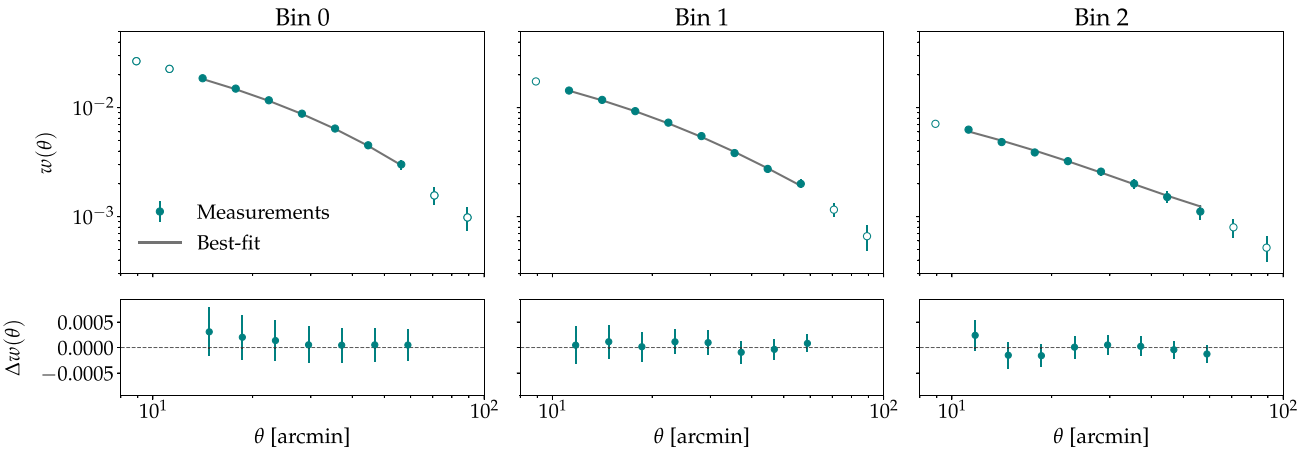


**Figure 13.** *Upper panels*: Measurements of the autocorrelations of angular galaxy clustering for the three redshift bins (0, 1, 2) defined in this work, for the entire DES Y3 footprint we use. Filled coloured points correspond to the measurements passing the scales cuts defined in Section 7.1.1. The methodology for the measurements and covariance, and the calculation of the corresponding signal to noise, can be found in Section 7.2. The solid lines show the best-fitting theory for the fiducial analysis choices, as described in Section 7.3. The goodness of fit in that case corresponds to $\chi^2/\mathrm{dof} = 8.3/8.6$. Error bars are smaller than the symbols, if not indicated. *Lower panels*: Residuals of the measurements given the best-fitting theory model shown in the upper panels.

### 7.3 Analysis and results

#### *7.3.1 Parameter inference*

In this part, we are interested in placing model constraints given the measured two-point functions of galaxy clustering shown in Fig. 13. In general, given our model $M$, we want to infer parameters $\boldsymbol{p}$ from the set of measured two-point correlation functions in our data, $\hat{\mathbf{D}}$. The theoretical model prediction for the two-point correlation functions, computed using the parameters $\boldsymbol{p}$ of the model $M$, is $\boldsymbol{T}_M(\boldsymbol{p}) \equiv \{w^{ij}(\theta, \boldsymbol{p})\}$. We compare the measurements and model predictions using a Gaussian likelihood, using the data covariance, $\boldsymbol{C}$, defined above:

$$L(\hat{\boldsymbol{D}}|\boldsymbol{p}, M) \propto \mathrm{e}^{-\frac{1}{2}\left[(\hat{\boldsymbol{D}}-\boldsymbol{T}_M(\boldsymbol{p}))^{\mathrm{T}}\boldsymbol{C}^{-1}(\hat{\boldsymbol{D}}-\boldsymbol{T}_M(\boldsymbol{p}))\right]}. \quad (18)$$

In this way, the posterior probability distribution for the parameters $\boldsymbol{p}$ of the model $M$ given the data $\hat{\boldsymbol{D}}$ is given by

$$P(\boldsymbol{p}|\hat{\boldsymbol{D}}, M) \propto L(\hat{\boldsymbol{D}}|\boldsymbol{p}, M)P(\boldsymbol{p}|M), \quad (19)$$

where $P(\boldsymbol{p}|M)$ is the prior probability distribution on the parameters.

We sample the posterior of the galaxy clustering measurements in the flat Lambda cold dark matter ($\Lambda$CDM) model, using the same parameter space as the DES Year 3 fiducial analysis (DES Collaboration 2022). The six cosmological parameters we vary are listed in Table 5, together with their respective uniform priors. These prior ranges are chosen to encompass at least five times the 68 per cent C.L. from relevant external constraints. Also, even though we sample the amplitude of primordial scalar density perturbations $A_s$, sometimes we will refer to the amplitude of density perturbations at $z = 0$ in terms of the RMS amplitude of mass on scales of $8\,h^{-1}$ Mpc in linear theory, $\sigma_8$. In addition to these cosmological parameters, our fiducial analysis includes 18 nuisance parameters to describe: galaxy bias (see Section 7.1), potential residuals in the galaxy weight calculation (see Section 4), lens magnification (see Section 6) and uncertainties in the redshift distribution of our three redshift bins (see Section 5), all of them described in Table 5.

**Table 5.** The model parameters and their priors used in the fiducial flat $\Lambda$CDM analysis, using the entire DES Y3 footprint. The parameters are defined in Section 7.3.

| Parameter | Prior | |
|---|---|---|
| Cosmology | | |
| $\Omega_\mathrm{m}$ | Flat | (0.1, 0.9) |
| $10^9 A_\mathrm{s}$ | Flat | (0.5,5.0) |
| $n_\mathrm{s}$ | Flat | (0.87, 1.07) |
| $\Omega_\mathrm{b}$ | Flat | (0.03, 0.07) |
| $h$ | Flat | (0.55, 0.91) |
| $10^3\Omega_\nu h^2$ | Flat | (0.60, 6.44) |
| Galaxy bias | | |
| $b^i (i \in [0, 2])$ | Flat | (0.8, 3.0) |
| Weight residuals | | |
| $R^0$ | Flat | (−8, −2) |
| $R^1$ | Flat | (−8, −2) |
| $R^2$ | Flat | (−8, −2) |
| Lens magnification | | |
| $C^0$ | Gaussian | (0.0275, 0.24) |
| $C^1$ | Gaussian | (1.305, 0.375) |
| $C^2$ | Gaussian | (2.145, 0.36) |
| Redshifts | | |
| $\Delta z^0$ | Gaussian | (0.0, 0.0051) |
| $\Delta z^1$ | Gaussian | (0.0, 0.0075) |
| $\Delta z^2$ | Gaussian | (0.0, 0.0208) |
| $\sigma_z^0$ | Gaussian | (0.997, 0.068) |
| $\sigma_z^1$ | Gaussian | (0.999, 0.041) |
| $\sigma_z^2$ | Gaussian | (0.998, 0.044) |
| $A^0_{\mathrm{low}-z}$ | Gaussian | (0.0044, 0.0013) |
| $A^1_{\mathrm{low}-z}$ | Gaussian | (0.0091, 0.0023) |
| $A^2_{\mathrm{low}-z}$ | Gaussian | (0.0383, 0.0059) |

#### *7.3.2 Blinding procedure*

In order to minimize a potential impact of experimenter bias, we have adopted a blinding procedure throughout this work. In that way, we have kept the results on the main parameters constrained in this analysis (those depicted in Figs 14 and 15) blinded to the analysis until the robustness tests performed in Section 7.3.3 satisfied the tension metrics reported there. An internal review committee set-up

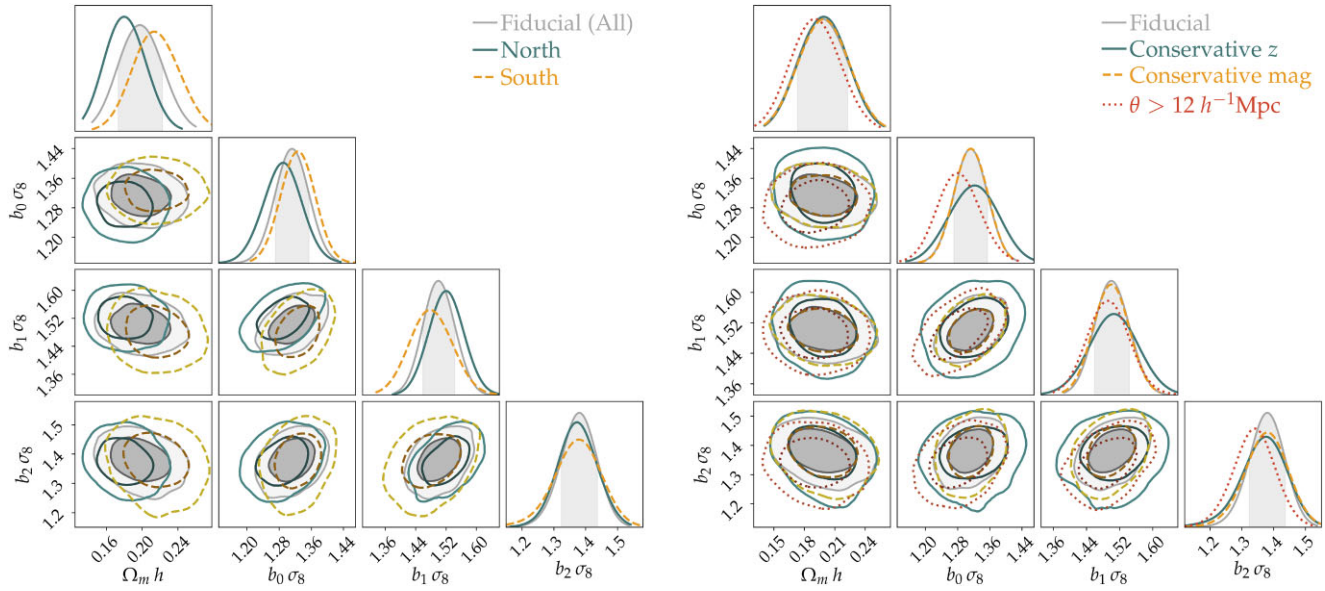


**Figure 14.** Constraints on the combination of cosmological parameters and galaxy bias derived from out measurements of galaxy clustering for various analysis configurations. The left-hand panel shows the fiducial constraints using the entire footprint (All), compared to the constraints using the independent splits in North and South regions. The right-hand panel shows the comparison between the fiducial constraints and three analysis variations, one with conservative redshift priors (×2 width in all redshift parameter priors), one with conservative magnification priors (×2 width in all magnification parameter priors), and larger minimum angular scales.

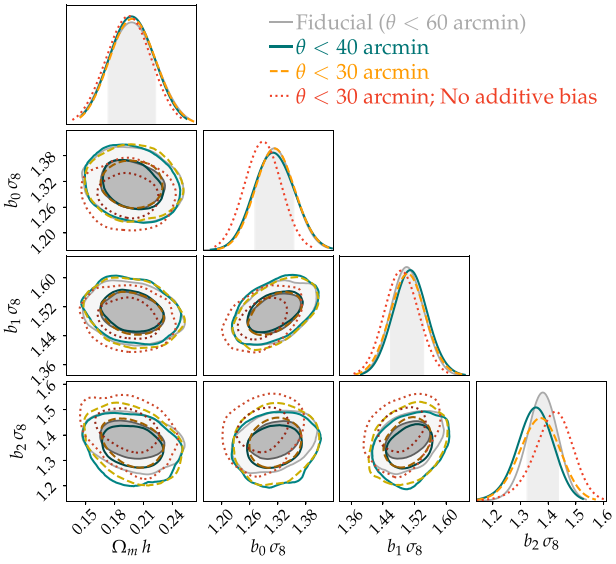


**Figure 15.** Comparison of the parameter constraints from galaxy clustering using different choices for the maximum angular scale, as well as not marginalizing over an additive constant in the galaxy clustering measurements.

by the DES collaboration was in charge of overviewing this procedure and allowing for the unblinding of the constraints.

### *7.3.3 DES Y3 high-z results and robustness tests*

Next, we analyse the model constraints from the measurements of galaxy clustering. In this case, there exists a strong degeneracy between galaxy bias and the amplitude of matter fluctuations, $\sigma_8$, and therefore, the analysis presented here is not sensitive to $\sigma_8$. The combination of clustering and weak gravitational lensing can be used to break these degeneracies, and that will be presented in a companion paper (*in preparation*), using CMB lensing from the SPT and *Planck*. However, for the clustering-only case analysed here, the shape of the galaxy clustering measurements is sensitive to the scale of matter–radiation equality in the matter power spectrum, which in turn depends on a combination of the matter density $\Omega_{\rm m}$ and the Hubble constant $h$, close to the direction $\Omega_{\rm m}h$ (see e.g. Philcox et al. 2021).

Fig. 14 shows the constraints we obtain for the parameters we are sensitive to, namely $\Omega_{\rm m}h$ and the product of $\sigma_8 b^i$ for the three redshift bins we use. The fiducial constraints use the entire survey footprint, the autocorrelations shown in Fig. 13, the scale cuts described in Section 7.1.1 and the priors shown in Table 5, and they result in constraints on a combination of the fraction of matter in the Universe $\Omega_{\rm m}$ and the Hubble parameter $h$, $\Omega_m h = 0.195^{+0.023}_{-0.018}$, and 2–3 per cent measurements of the amplitude of the galaxy clustering signals for the three redshift bins, probing galaxy bias and the amplitude of matter fluctuations, $b\sigma_8$. The best-fitting theory model for this fiducial case is shown together with the measurements in Fig. 13, and the corresponding $\chi^2/ndf$ is 8.3/8.1, where $ndf$ is the estimated effective number of degrees of freedom. Using the Update Difference in Mean tension metric from Lemos et al. (2021), we find the posterior constraints to be compatible with the redshift prior, with a tension of $0.34\sigma$, and also compatible with the magnification prior, with a tension at $0.03\sigma$.

In addition, to assess the robustness of the results, in Fig. 14 we show constraints for various alternative cases. First, we analyse the constraints we obtain from the independent North and South regions, where we split the data into two independent patches: 'North' (Dec. $> -39°$) and 'South' (Dec. $< -40°$). This is motivated by the fact that we will combine the clustering measurements shown here with CMB lensing measurements from *Planck* and SPT in a companion paper *(in preparation)*. Since SPT only covers the South region in this split, we do this test to check for the consistency of the

clustering measurements. In this test, the redshift and magnification priors are computed specifically for each region, although they are largely consistent (see Tables 3 and 4), and the galaxy clustering measurements are also performed separately for the two regions. The analysis of the North and South regions yields best-fitting theory models with $\chi^2/ndf$ is 10.1/8.6 and $\chi^2/ndf$ is 15.7/8.6, respectively. When using the entire parameter space, the constraints from the two independent regions are in agreement, with an estimated tension of $0.65\sigma$, using the non-Gaussian parameter difference tension metric from Lemos et al. (2021) and Raveri & Doux (2021). When restricting the set of parameters to $\Omega_m$, $\Omega_m\, h$, $b^0\,\sigma_8$, $b^1\,\sigma_8$, $b^2\,\sigma_8$, the constraints from the independent North and South regions are also in agreement, with an estimated tension of $0.41\sigma$.

Fig. 14 also shows the galaxy clustering constraints under some different analysis choices. In particular, we study the impact of redshift and magnification priors, both described in Table 5, by studying the conservative case of doubling the width these priors. When broadening the width of redshift priors by a factor of 2, the constraints on $b^0\sigma_8$, $b^1\sigma_8$, and $b^2\sigma_8$ widen by a factor of 1.47, 1.41, and 1.27, respectively. When broadening the width of magnification priors by a factor of 2, the constraints on $b^2\sigma_8$ broaden by a factor of 1.20. Therefore, redshift priors are relevant for all bins, especially for bins 0 and 1, while lens magnification is only relevant in bin 2, at higher redshift. None of these changes has an important effect on $\Omega_m h$, which shows very robust constraints under all different analysis choices. Using larger minimum angular scales, corresponding to 12 $h^{-1}$ Mpc, as opposed to the fiducial 8 $h^{-1}$ Mpc, broadens the constraints on $b^i\sigma_8$ by a factor of 1.28, 1.21, and 1.17 for bins $i = 0$, 1, and 2, while having no significant effect on $\Omega_m h$.

We also explore the impact of the choice of maximum angular scale on the clustering analysis. The fiducial value for the maximum angular scale is 60 arcmin, driven by the method used to correct for correlations between galaxy density and survey properties. In order to account for any residuals coming from that method, we also marginalize over an additive constant parameter for each tomographic bin $R^i$ (see equation 15). Fig. 15 shows the galaxy clustering constraints when limiting the maximum angular scale to 40 and 30 arcmin, and also, for the latter case, when not marginalizing over additive constants. The figure shows how the galaxy clustering constraints are robust to these choices. The constraints on $\Omega_m h$ are not sensitive to the variations, and the main impact of limiting the maximum angular scale is a ∼20 per cent decrease in constraining power for $b^2\sigma_8$. Regarding the posterior values of $R^i$, we find $R^0 = -5.13^{+0.84}_{-1.93}$, $R^1 = -3.42^{+0.31}_{-0.65}$, $R^2 = -3.21^{+0.06}_{-0.09}$. We can see how this parameter is constrained to be very small for the first bin, and its importance grows with redshift (and $i$-band magnitude) of the tomographic bin.

### 7.3.4 *Comparison with other DES Y3 clustering analyses*

Given the parameter constraints obtained in the analysis of galaxy clustering with the DES Y3 High-$z$ sample presented in this work, we can now compare how these constraints compare with the corresponding clustering analyses of the other DES Y3 lens samples already defined and used in other works. The fiducial DES Y3 lens sample is the so-called MAGLIM sample (Porredon et al. 2021), while the alternative lens sample is REDMAGIC (Pandey et al. 2022; see Table 1 for a comparison of the number densities of the three samples). Fig. 16 shows the constraints on the cosmological parameter combination of $\Omega_m h$ provided by each of the three DES Y3 lens samples, together with the Planck 2018 constraint. The

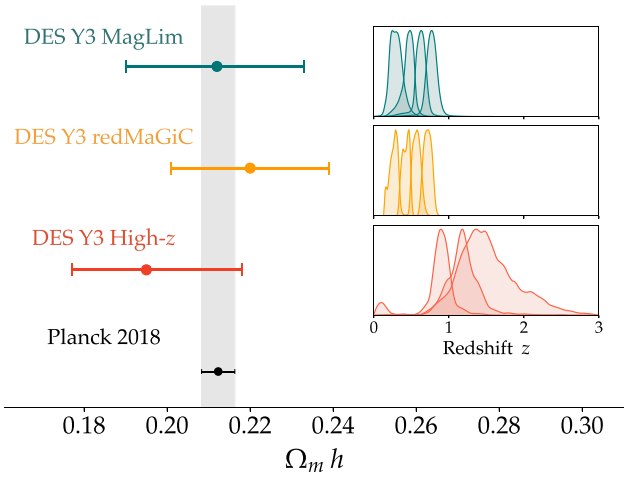


**Figure 16.** Comparison of the constraints on the parameter combination $\Omega_m h$ from galaxy clustering analyses using three different lens samples in DES Y3. The constraints from the DES Y3 Fiducial sample, also called MagLim sample [ref], are shown in blue; the constraints from the redMaGiC sample are showing in yellow and the constraints from the High-$z$ sample described in this work are shown in red. The Planck 2018 constraint is shown in black. The inset panel on the right of the plot depicts the different redshift range probed by the three DES Y3 lens samples.

figure shows the DES Y3 constraints to be in agreement between the three samples, and with the Planck result, and also having similar constraining power. However, while the constraints from MAGLIM and REDMAGIC probe similar redshift ranges, the High-$z$ constraints come from significantly higher redshifts, extending the redshift range probed by the DES Y3 data. This results demonstrate the robustness of the clustering measurements in this work and our ability to produce a well-characterized high-redshift sample, which is complimentary to the DES fiducial analysis in terms of the redshift range it probes. Note that the upcoming analyses combining the High-$z$ galaxy clustering presented in this work with cross-correlation with weak gravitational lensing will be able to break the degeneracy between galaxy bias and the amplitude of matter fluctuations, $\sigma_8$, allowing us to place constraints on the latter at higher redshifts than probed in the fiducial DES analysis.

## 8 SUMMARY AND OUTLOOK

The cosmological analysis of imaging galaxy surveys provides powerful measurements of the amplitude of matter fluctuations in the late time Universe. In recent years, the analyses of different surveys like DES, KiDS and HSC, probing the regime at $z < 1$, have reported persistent tensions with the predicted value from the CMB, a problem known as the $S_8$ tension. Measurements at a higher redshift regime ($1 < z < 3$) would be crucial for understanding the origin of this tension. In addition, such measurements would probe the matter-dominated epoch and would shed light on dynamical dark energy models that can mimic a cosmological constant at late times but differ substantially during the matter-dominated era.

In this work, we describe the selection and characterization of three galaxy samples covering the approximate redshift range $0.8 < z < 2.5$ (see Fig. 6) using data from the third year of the Dark Energy Survey Year 3 (DES Y3). To enable the selection and characterization of these high-$z$ samples, which push the limits of DES Y3 data, we introduce several changes with respect to the fiducial DES Y3 lens galaxy sample:

(i) We start from a fainter galaxy selection, excluding all lens galaxies used in the DES Y3 fiducial analysis. The average $i$-band magnitude of the three High-$z$ redshift bins is 22.6, 22.9, and 23.1, respectively, while all four redshift bins used in the fiducial analysis had average $i$-band magnitudes brighter than $i = 22$.

(ii) Both the selection and redshift characterization of the samples are based on a principled, Bayesian scheme using a novel SOM algorithm better suited for the characterization of lower S/N galaxies (Sánchez et al. 2020).

(iii) We use a redshift marginalization scheme that explicitly accounts for uncertainties in the tails of redshift distributions.

(iv) We use a non-linear, machine-learning-based approach to correct for correlations between galaxy number density and survey observing properties like depth, stellar density, and sky noise.

Out of this list of changes with respect to the fiducial analysis, steps (i) and (ii) are responsible for the selection of high redshift galaxies, and steps (iii) and (iv) are required due to the faint, high-redshift selection. The procedure results in the definition of three redshift bins with mean redshifts around $z = 0.9$, 1.2, and 1.5, which significantly extend the redshift coverage of the fiducial DES Year 3 analysis. In addition, these samples contain a total of about 9 million galaxies, resulting in a galaxy density that is more than 2 times higher than those in the DES Year 3 fiducial case (Porredon et al. 2022).

After the selection and characterization of the high-$z$ galaxy samples, we perform an analysis of their galaxy clustering autocorrelation measurements. The analysis provides robust constraints on the product of the fraction of matter in the Universe $\Omega_{\rm m}$ and the Hubble parameter $h$, $\Omega_m h = 0.195^{+0.023}_{-0.018}$, and 2–3 per cent measurements of the amplitude of the galaxy clustering measurements for the three redshift bins, probing galaxy bias times the amplitude of matter fluctuations, $b\sigma_8$. The constraints on $\Omega_{\rm m}h$ are compatible and show comparable uncertainties to the clustering analyses on the fiducial and alternative lens galaxy samples using DES Y3 data (Porredon et al. 2022; Pandey et al. 2022), but probing a complementary, much higher redshift range. This part also showcases the robustness of the galaxy clustering analysis, which is highly non-trivial when using galaxy samples going as faint as $i \sim 23$ in DES Y3 data.

The definition and characterization of high-redshift galaxy samples in this work represents the first step to analyse the $0.8 < z < 2.5$ redshift range made by DES and other Stage III surveys. It therefore develops the tools that will enable similar analyses with other data sets, including Rubin LSST and *Euclid*, and it opens the door to a range of scientific analyses exploiting the unique nature of the selections. In subsequent publications, we will explore this set of applications using the samples defined in this work. We will present the cross-correlation of High-$z$ galaxies with CMB lensing maps from SPT and *Planck*, providing crucial constraints on $S_8$ at high redshift (Planck Collaboration VIII 2020; Omori et al. 2023). We will also study their cross-correlations with galaxy lensing, probing $S_8$, lensing magnification and intrinsic alignments at high redshifts, and the clustering cross-correlations with lower redshift galaxies, probing lensing magnification and the redshift evolution of galaxy bias. The redshift regime of these samples is also well suited to study the star formation history using cross-correlations with the cosmic infrared background (Jego et al. 2023a, b). The outcome of these analyses will provide important information about this particularly unexplored period in the Universe, and will set the tools and expectations for future analyses with more powerful data sets.

## ACKNOWLEDGEMENTS

CS acknowledges financial support from Ramón y Cajal project RYC2021-031194-I, funded by MCIN/AEI/10.13039/501100011033 and by the 'European Union NextGenerationEU/PRTR', and by a Junior Leader fellowship from the 'la Caixa' Foundation (ID 100010434), with code LCF/BQ/PI22/11910018. CS and GMB are supported by grants AST-2009210 from the U.S. National Science Foundation, and DE-SC0007901 from the U.S. Department of Energy (DOE). Argonne National Laboratory's work was supported by the U.S. Department of Energy, Office of High Energy Physics. Argonne, a U.S. Department of Energy Office of Science Laboratory, is operated by UChicago Argonne LLC under contract no. DE-AC02-06CH11357.

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

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, NSF's NOIRLab, 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 MICINN under grants ESP2017-89838, PGC2018-094773, PGC2018-102021, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA programme of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework 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) do e-Universo (CNPq grant 465376/2014-2).

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

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

We thank the developers of NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), JUPYTER (Kluyver et al. 2016), IPYTHON (Pérez & Granger 2007), CONDA-FORGE (conda-forge community 2015), MATPLOTLIB (Hunter 2007), and KERAS (Chollet et al. 2015) for their extremely useful free software.

## DATA AVAILABILITY

The authors will share the data needed to reproduce the galaxy selections developed in this paper, i.e. the galaxies making up the three tomographic bins, the Wide SOM, and the cells making up each tomographic bin, upon reasonable request.

A general description of DES data releases is available on the survey website at https://www.darkenergysurvey.org/the-des-project/data-access/. DES Y3 cosmological data has been partially released on the DES Data Management website hosted by the National Center for Supercomputing Applications at https://des.ncsa.illinois.edu/releases/y3a2.

## REFERENCES

Abbott T. M. C. et al., 2019, Phys. Rev. D, 100, 023541
Alarcon A., Sánchez C., Bernstein G. M., Gaztañaga E., 2020, MNRAS, 498, 2614
Alarcon A. et al., 2021, MNRAS, 501, 6103
Alonso D., Bellini E., Hale C., Jarvis M. J., Schwarz D. J., 2021, MNRAS, 502, 876
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Baxter E. J. et al., 2019, Phys. Rev. D, 99, 023508
Buchs R. et al., 2019, MNRAS, 489, 820
Bull P., White M., Slosar A., 2021, MNRAS, 505, 2285
Carlstrom J. E. et al., 2011, PASP, 123, 568
Chang C. et al., 2023, Phys. Rev. D, 107, 023530
Chollet F. et al., 2015, Keras. Available at: https://keras.io
conda-forge community, 2015, The Conda-forge Project: Community-based Software Distribution Built on the conda Package Format and Ecosystem. Zenodo, available at: https://doi.org/10.5281/zenodo.4774216
Crocce M. et al., 2016, MNRAS, 455, 4301
DES Collaboration, 2022, Phys. Rev. D, 105, 023520
De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, MNRAS, 459, 3078
Elvin-Poole J. et al., 2018, Phys. Rev. D, 98, 042006
Elvin-Poole J. et al., 2023, MNRAS, 523, 3649
Eriksen M. et al., 2019, MNRAS, 484, 4200
Euclid Collaboration, 2022, A&A, 662, A93
Everett S. et al., 2022, ApJS, 258, 15
Fang X., Krause E., Eifler T., MacCrann N., 2020, J. Cosmol. Astropart. Phys., 2020, 010
Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015, MNRAS, 448, 2987
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Gaia Collaboration, 2021, A&A, 649, A1
Garcia-Fernandez M. et al., 2018, MNRAS, 476, 1071
Gaztanaga E., Schmidt S. J., Schneider M. D., Tyson J. A., 2021, MNRAS, 503, 4964
Giannantonio T., Scranton R., Crittenden R. G., Nichol R. C., Boughn S. P., Myers A. D., Richards G. T., 2008, Phys. Rev. D, 77, 123520
Giannini G. et al., 2022, preprint (arXiv:2209.05853)
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Hang Q., Alam S., Peacock J. A., Cai Y.-C., 2021, MNRAS, 501, 1481
Harikane Y. et al., 2018, PASJ, 70, S11
Harris C. R. et al., 2020, Nature, 585, 357
Hartley W. G. et al., 2022, MNRAS, 509, 3547
Heymans C. et al., 2021, A&A, 646, A140
Hikage C. et al., 2019, PASJ, 71, 43
Hildebrandt H., van Waerbeke L., Erben T., 2009, A&A, 507, 683
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338
Jarvis M. J. et al., 2013, MNRAS, 428, 1281
Jeffrey N. et al., 2021, MNRAS, 505, 4626
Jego B., Ruiz-Zapatero J., García-García C., Koukoufilippas N., Alonso D., 2023a, MNRAS, 520, 1895
Jego B., Alonso D., García-García C., Ruiz-Zapatero J., 2023b, MNRAS, 520, 583
Joudaki S. et al., 2020, A&A, 638, L1
Kluyver T. et al., 2016, in Positioning and Power in Academic Publishing: Players, Agents and Agendas. IOS Press, p. 87
Kohonen T., 1982, Biol. Cybern., 43, 59
Krause E. et al., 2021, preprint (arXiv:2105.13548)
Krolewski A., Ferraro S., Schlafly E. F., White M., 2020, J. Cosmol. Astropart. Phys., 2020, 047
Krolewski A., Ferraro S., White M., 2021, J. Cosmol. Astropart. Phys., 2021, 028
Kwan J., Sánchez C. et al., 2017, MNRAS, 464, 4045
Laigle C. et al., 2016, ApJS, 224, 24
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Le Fèvre O. et al., 2013, A&A, 559, A14
Leistedt B. et al., 2016a, ApJS, 226, 24
Leistedt B., Mortlock D. J., Peiris H. V., 2016b, MNRAS, 460, 4258
Lemos P. et al., 2021, MNRAS, 505, 6179
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Lilly S. J. et al., 2009, ApJS, 184, 218
Malz A. I., Hogg D. W., 2022, ApJ, 928, 127
Marques G. A., Bernui A., 2020, J. Cosmol. Astropart. Phys., 2020, 052
Masters D. et al., 2015, ApJ, 813, 53
Masters D. C., Stern D. K., Cohen J. G., Capak P. L., Rhodes J. D., Castander F. J., Paltani S., 2017, ApJ, 841, 111
Masters D. C. et al., 2019, ApJ, 877, 81
McCracken H. J. et al., 2012, A&A, 544, A156
Ménard B., Hamana T., Bartelmann M., Yoshida N., 2003, A&A, 403, 817
Miyatake H. et al., 2022, Phys. Rev. Lett., 129, 061301
Morganson E. et al., 2018, PASP, 130, 074501
Myles J. et al., 2021, MNRAS, 505, 4249
Omori Y. et al., 2019a, Phys. Rev. D, 100, 043501
Omori Y. et al., 2019b, Phys. Rev. D, 100, 043517
Omori Y. et al., 2023, Phys. Rev. D, 107, 023529
Ono Y. et al., 2018, PASJ, 70, S10
Padilla C. et al., 2019, AJ, 157, 246
Pandey S. et al., 2022, Phys. Rev. D, 106, 043520
Pérez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Philcox O. H. E., Sherwin B. D., Farren G. S., Baxter E. J., 2021, Phys. Rev. D, 103, 023538
Planck Collaboration XXII, 2016, A&A, 594, A22
Planck Collaboration VIII, 2020, A&A, 641, A8
Porredon A. et al., 2021, Phys. Rev. D, 103, 043503
Porredon A. et al., 2022, Phys. Rev. D, 106, 103530
Rau M. M., Morrison C. B., Schmidt S. J., Wilson S., Mandelbaum R., Mao Y. Y., Mao Y. Y., LSST Dark Energy Science Collaboration 2022, MNRAS, 509, 4886
Raveri M., Doux C., 2021, Phys. Rev. D, 104, 043504

Rezaie M., Seo H.-J., Ross A. J., Bunescu R. C., 2020, MNRAS, 495, 1613
Rodríguez-Monroy M. et al., 2022, MNRAS, 511, 2665
Ross A. J. et al., 2011, MNRAS, 417, 1350
Ross A. J. et al., 2017, MNRAS, 464, 1168
Rozo E. et al., 2016, MNRAS, 461, 1431
Sánchez C., Bernstein G. M., 2019, MNRAS, 483, 2801
Sánchez C., Raveri M., Alarcon A., Bernstein G. M., 2020, MNRAS, 498, 2984
Schlafly E. F., Meisner A. M., Green G. M., 2019, ApJS, 240, 30
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Schmittfull M., Seljak U., 2018, Phys. Rev. D, 97, 123540
Scodeggio M. et al., 2018, A&A, 609, A84
Scoville N. et al., 2007, ApJS, 172, 1
Sevilla-Noarbe I. et al., 2021, ApJS, 254, 24
Stanford S. A. et al., 2021, ApJS, 256, 9
Suchyta E. et al., 2016, MNRAS, 457, 786
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
van den Busch J. L. et al., 2022, A&A, 664, A170
Virtanen P. et al., 2020, Nat. Methods, 17, 261
von Wietersheim-Kramsta M. et al., 2021, MNRAS, 504, 1452
Weaverdyck N., Huterer D., 2021, MNRAS, 503, 5061
Zuntz J. et al., 2015, Astron. Comput., 12, 45

## APPENDIX A: REDSHIFT DISTRIBUTION UNCERTAINTIES

In this section, we go over the redshift calibration presented in Section 5 in detail.

### A1 Redshift uncertainty parametrization

We can express the parametric $N(z)$ error model as

$$
\begin{aligned}
N_i(z, \theta^i, A^i_{\text{low}-z}) &= C_{N_i} \times \begin{cases} G_i\, A^i_{\text{low}-z} & z \leq 0.5 \\ G_i\,(1 - A^i_{\text{low}-z}) & z > 0.5 \end{cases} \\
G_i(z, \theta^i) &= C_{G_i} \times \begin{cases} F_i(y) & |z - \bar{z}_i| \leq 2\Sigma_{z_i} \\ F_i(z) & |z - \bar{z}_i| > 2\Sigma_{z_i} \end{cases} \\
y &= \sigma^i_z (z - \Delta z^i - \bar{z}) + \bar{z} \\
\bar{z}_i &= \int z\, F_i(z)\, dz \\
\Sigma_{z_i} &= \sqrt{\int (z - \bar{z}_i)^2\, F_i(z)\, \mathrm{d}z} \\
F_i(z) &= i\text{-th Fiducial redshift distribution} \\
\theta^i &= \{\Delta z^i,\ \sigma^i_z\} \\
(C_{N_i})^{-1} &= \int N_i(z, \theta^i, A^i_{\text{low}-z})\, \mathrm{d}z \\
(C_{G_i})^{-1} &= \int G_i(z, \theta^i)\, \mathrm{d}z
\end{aligned}
\tag{A1}
$$

with $\Delta z^i$ the shift, $\sigma^i_z$ the stretch and $A^i_{\text{low}-z}$ the low redshift fraction free parameters of the model.

A visualization of the shift, stretch and low-$z$ fraction parameters can be seen in Fig. 11. On the one hand, the galaxy clustering signal cares both about the mean redshift of the distribution but also of its spread in redshift, as the more spread out galaxies are the less physically correlated they become, reducing the clustering signal. On the other hand, the majority of the selected galaxies live primarily at high redshift, but with *griz* colours a population of low-$z$ galaxies leaks into the selection, especially in our highest redshift bin, producing a distinct clustering signal than that of the high redshift galaxies. Furthermore, we smooth the fiducial redshift distribution with a Savitzky–Golay filter: sample variance and shot noise from the small area of the calibration deep fields manifests in the $N(z)$ as rapid fluctuations in redshift and enter squared in the galaxy clustering signal, while the true redshift distribution over a larger area is way more smooth as these variations average out. We try different smoothing lengths and find compatible constraints on the main parameters of interest (see Appendix A6).

### A2 Redshift biases

To measure the colour–redshift relation in the deep fields, we build our redshift sample from a combination of the redshift information that we have available from spectroscopic and multiband photometric redshifts, SPC (see Section 2.3). Whenever a galaxy has spectroscopic measurements, we use them. Alternatively, we use photometric redshifts from the PAUS + COSMOS, and when that is not available we use redshifts from COSMOS2015. After removing colour regions with significant stellar contamination and retraining the Deep SOM (see Section 3), we find that only 9 out of 2304 cells (0.4 per cent) do not have any overlapping redshifts, but relative to the probability of finding galaxies in these cells $p(c)$, they amount to only 0.1 per cent of the probability. Each tomographic bin relates with different probability to each deep cell, and when we take that into account the relative probability without redshift information in each tomographic bin is 0.1 per cent, 0 per cent, and 0 per cent.

We only use high-quality spectroscopic redshifts; therefore, we assume the spectroscopic redshifts are accurate and precise. However, the photo-$z$ from COSMOS2015 and PAUS + COSMOS are estimated from multiband photometric band data, with band filters spanning a wide range in wavelength and with multiple intermediate and narrow bands. The individual $p(z)$ from these catalogs are broader, but their width is still negligible compared to the redshift resolution from noisier wide field observations with *griz* broad bands, and so we simply stack the individual $p(z)$. Stacking the $p(z)$ is statistically incorrect, and for galaxies where the $p(z)$ is degenerate between two different redshift values, or if the $p(z)$ were wider, then a more correct technique should be used (e.g. Leistedt et al. 2016b; Sánchez & Bernstein 2019; Alarcon et al. 2020; Malz & Hogg 2022; Rau et al. 2022). We defer the application of such techniques for future work.

An additional concern is whether the photo-$z$ estimates from these catalogs are systematically biased from an incorrect modeling of the galaxy SEDs (e.g. Joudaki et al. 2020; Myles et al. 2021; van den Busch et al. 2022). Here, we measure the bias by comparing the photo-$z$ estimates of individual objects in both catalogues to overlapping spectroscopic measurements (described in Section 2.3). For each of these objects, we calculate $(z_{\text{phot}} - z_{\text{spec}})/(1 + z_{\text{spec}})$, with $z_{\text{phot}}$ the mode of the $p(z)$, and we plot the distributions. By visual inspection, we find that the distributions of COSMOS2015 and PAUS + COSMOS are generally unimodal, but sometimes slightly biased. We define the median bias as a function of the DES deep field $i$-band magnitude as

$$
b(i) = \text{Median}\left(\frac{z_{\text{phot}} - z_{\text{spec}}}{1 + z_{\text{spec}}} \mid i\right). \tag{A2}
$$

Fig. A1 shows $b(i)$ from both catalogues: we find a slight positive bias $b(i) \sim 0.002$ at faint magnitudes in the PAUS + COSMOS catalogue, while the COSMOS2015 catalogue presents a negative bias reaching a minimum value of $b(i = 22.5) \sim -0.005$. We model the redshift bias uncertainty in these samples with a parameter $\alpha$ that shifts the individual $p(z)$ of COSMOS2015 or PAUS + COSMOS galaxies (one $\alpha$ parameter for each catalogue). This $\alpha$ parameter shifts $p(z) \rightarrow p(z - \delta(\alpha, i) \cdot (1 + z))$ by an amount $\delta$ that is proportional to the

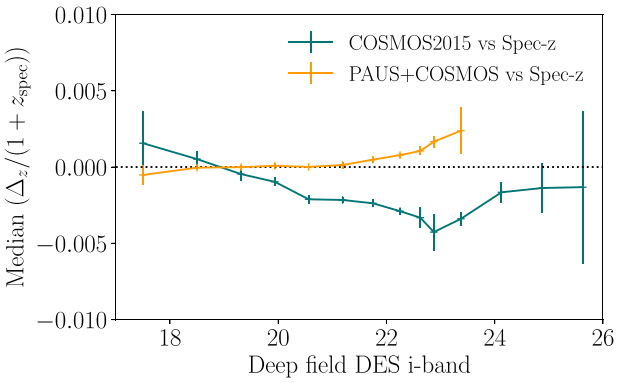


**Figure A1.** Median photo-$z$ bias as a function of the deep field DES $i$-band. The bias is calculated for `COSMOS2015` and `PAUS + COSMOS` for galaxies where a spectroscopic measurement also exists, with $\Delta z = z_{\rm phot} - z_{\rm spec}$. This measured bias is used to estimate the redshift bias of this catalogues and is marginalized over in our analysis. See Section A2 for more details.

median bias of a galaxy of magnitude $i$:

$$\delta(\alpha, i) = \alpha\, b(i) \tag{A3}$$

We place a Gaussian prior on this parameter and marginalize over it, $p(\alpha) = N(\mu = 1, \sigma = 1)$. Therefore, our most likely guess for the systematic bias is centred at the measured median bias $b(i)$, but we assign an uncertainty equal to the magnitude of $b(i)$. Note that the value $\alpha$ is the same for all galaxies in the same catalog, but the magnitude of the shift to the $p(z)$ ultimately depends on both the redshift and magnitude of each galaxy: $\delta(\alpha, i) \cdot (1 + z)$.

### A3 Selection biases

We empirically measure the prior on the colour–redshift relation from the galaxies in the deep field that have overlapping redshifts. Since we do not parametrize this prior and let the parameters update hierarchically with wide field galaxies, it is crucial to include all selection effects for the final estimate to be unbiased. BALROG injects versions of these galaxies into the wide field and allows us to measure the probability they will be selected into each of our tomographic bins, and therefore to correct for these selection effects. However, due to the limited number of BALROG injections, we cannot always measure these effects accurately, leading to several approximations to the SOMPZ methodology described in Section 3. In this section we explain these approximations and their validity, and provide a way to marginalize over the potential systematic biases that they might introduce.

The first row of panels (from the top) of Fig. A2 show the distribution of deep field galaxies in the Deep SOM weighted by their probability of being selected in each tomographic bin as measured by BALROG. This distribution is different than the one presented in Fig. 2, where we show the distribution of deep field galaxies weighted by their probability of being selected at $22 \leq i \leq 23.5$ according to BALROG. Note how in each panel the distribution peaks around Deep SOM cells with high redshift and has little to no overlap with cells at lower redshift, as expected (compare to Fig. A3 for the distribution of mean redshift in the Deep SOM).

The redshift distribution of each Deep SOM cell formally depends on the pre-selections $\hat{s}$ and on the Wide SOM cell where galaxies are selected, $p(z|c, \hat{c}, \hat{s})$, see equation (A5):

$$p(z|\hat{b}, \hat{s}) = \sum_{\hat{c}\in\hat{b}} p(z|\hat{c}, \hat{s}, \hat{b})\, p(\hat{c}|\hat{s}, \hat{b}) \tag{A4}$$

$$= \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c, \hat{c}, \hat{s})\, p(c|\hat{c}, \hat{s})\, p(\hat{c}|\hat{s}, \hat{b}) \tag{A5}$$

$$\approx \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c, \hat{b}, \hat{s})\, p(c|\hat{c}, \hat{s})\, p(\hat{c}|\hat{s}, \hat{b}) \tag{A6}$$

$$\approx \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c, \hat{B}, \hat{s})\, p(c|\hat{c}, \hat{s})\, p(\hat{c}|\hat{s}, \hat{b}) \tag{A7}$$

$$\approx \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c, \hat{s})\, p(c|\hat{c}, \hat{s})\, p(\hat{c}|\hat{s}, \hat{b}). \tag{A8}$$

Using BALROG we can empirically measure how often deep field galaxies $c$ will get through our pre-selections $\hat{s}$ and also how often they get selected in the different wide field cells $\hat{c}$. However, due to the limited number of BALROG injections it is not possible to accurately measure the relation between all $(z, c, \hat{c})$. Following Myles et al. (2021), we use the approximation shown in equation (A6) for our fiducial estimation of the redshift distribution of deep cells using $p(z|c, \hat{c}) \approx p(z|c, \hat{b})$, with $\hat{b}$ representing the set of $\hat{c}$ of a tomographic bin. When no redshift galaxy satisfies both $c$ and $\hat{b}$ then we use $p(z|c, \hat{B}, \hat{s})$ (equation A7) using redshift information from galaxies that are selected into any of the tomographic bins $\hat{B} \equiv \{\hat{b}_0, \hat{b}_1, \hat{b}_2\}$, or else $p(z|c, \hat{s})$ (equation A8), using redshift information from any galaxies satisfying our pre-selection $\hat{s}$.

The second row of panels of Fig. A2 shows the difference in the mean redshift of each cell from including the tomographic bin selection, showing:

$$\Delta\langle z\rangle_i \equiv \int z\, p(z|c, \hat{b}_i, \hat{s})\, \mathrm{d}z - \int z\, p(z|c, \hat{s})\, \mathrm{d}z$$
$$\equiv \langle z \mid \hat{b}_i\rangle - \langle z\rangle. \tag{A9}$$

Note how the $\Delta\langle z\rangle_i$ values tend to be close to 0 where the distribution of $p(c|b_i)$ peaks (top panels), as most galaxies from these cells get selected very often into that tomographic bin. However, note that $\Delta\langle z\rangle_i$ shows larger differences at the tails of the $p(c|b_i)$ distribution. In such cells, generally speaking, galaxies with a redshift that is closer to the average redshift of the tomographic bin get preferentially selected, and consequently cells with a $\langle z\rangle$ smaller than the average redshift of the bin tend to have a positive $\Delta\langle z\rangle_i$, and vice versa. This effect is very clear in bin 0, where cells at the lower part of the SOM have a $\langle z\rangle$ that is smaller than the typical redshift of galaxies in bin 0, and they show a positive $\Delta\langle z\rangle_0$, implying that additionally conditioning on the tomographic bin tends to increase the mean redshift of these cells. We find the contrary for cells at the top of the SOM, which have a $\langle z\rangle$ that is larger than the typical redshift of galaxies in this bin and they present a negative $\Delta\langle z\rangle_0$ that lowers the average redshift of the cell when we condition their selection to the bin.

This highlights how important it is to at least include the so-called *bin conditionalization*,[9] i.e. using $p(z|c, \hat{b}, \hat{s})$ instead of just $p(z|c, \hat{s})$. Otherwise one will introduce important selection effect biases, as those found by Buchs et al. (2019), where they found a positive bias for low redshift bins relative to the average redshift and a negative bias for high redshift bins, as a result of just using $p(z|c, \hat{s})$. More quantitatively, the average difference in mean redshift per tomographic bin, measured as $\sum_c p(c|\hat{b}_i)(\langle z \mid \hat{b}_i\rangle - \langle z\rangle)$, is $(-3, 11, 12) \times 10^{-3}$, which is non-negligible.

[9]We follow the notation introduced in Myles et al. (2021).

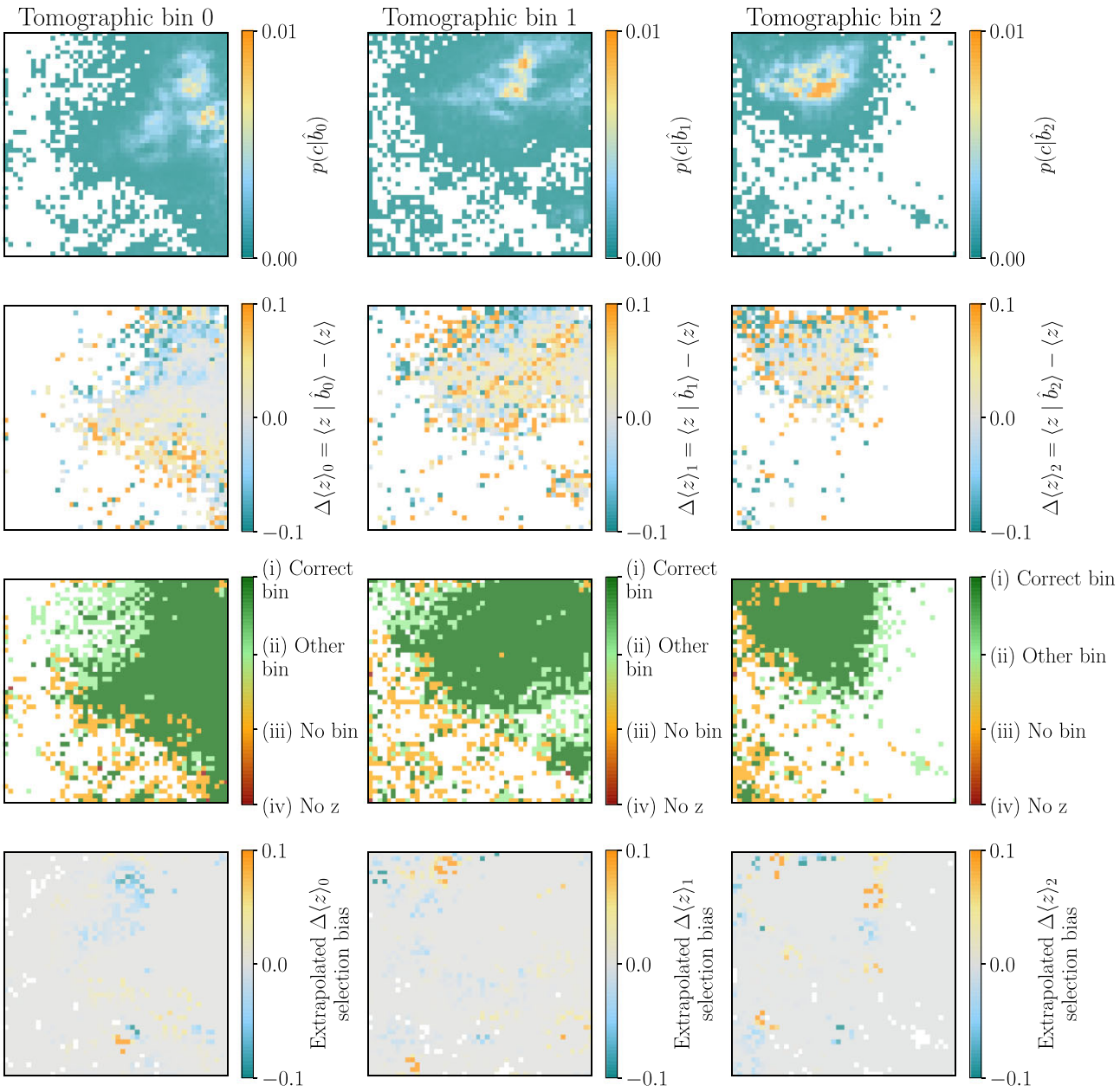


**Figure A2.** The redshift selection effects and the extrapolated $\Delta z$ selection effect bias. Each column shows a different tomographic bin. The first row of panels shows the pdf of deep field cells conditioned on each tomographic bin, $p(c|\hat{b}_i)$. The second row of panels shows the mean redshift difference of deep field cells when galaxies are additionally conditioned to be observed by BALROG into our each tomographic bin. The third row of panels shows which cells have some galaxy with redshift information selected into the bin by BALROG (i), which do not (ii)–(iii), and also which do not have any $z$ information (iv) (only five cells for bin 0, four for bin 1 and one for bin 2). The fourth row of panels show an extrapolated redshift bias. The redshift bias due to the additional selection of galaxies into the bin is extrapolated from (i) cells that have galaxies selected into the bin to cells (ii)–(iii)–(iv) that do not. See Section A3 for more details.

The third row of panels in Fig. A2 shows with a colour code which cells have redshift estimates that include accurate tomographic bin selection effects. The colour code goes as follows:

(i) *Dark green*: Cells that have at least one redshift galaxy that has been selected by BALROG into the corresponding tomographic bin, we use equation (A6), $p(z|c, \hat{b}, \hat{s})$.

(ii) *Light green*: Cells that have do not have any galaxy selected into the corresponding tomographic but at least one redshift galaxy that has been selected by BALROG into one of the other two tomographic bins, we use equation (A7), $p(z|c, \hat{B}, \hat{s})$.

(iii) *Light red*: Cells that have do not have any galaxy selected into any tomographic bin, but at least some galaxy satisfying our pre-selection $\hat{s}$. We use equation (A8), $p(z|c, \hat{s})$.

(iv) *Dark red*: Cells that have do not have any redshift galaxy satisfying our pre-selection $\hat{s}$. We do not have direct redshift information for these cells.

Note how the $\Delta\langle z\rangle_i$ from the second row of panels can only be calculated for (i)/Dark Green cells in the third row of panels. The remaining cells do not have any galaxy selected by BALROG into the corresponding tomographic bin, and bin conditionalization

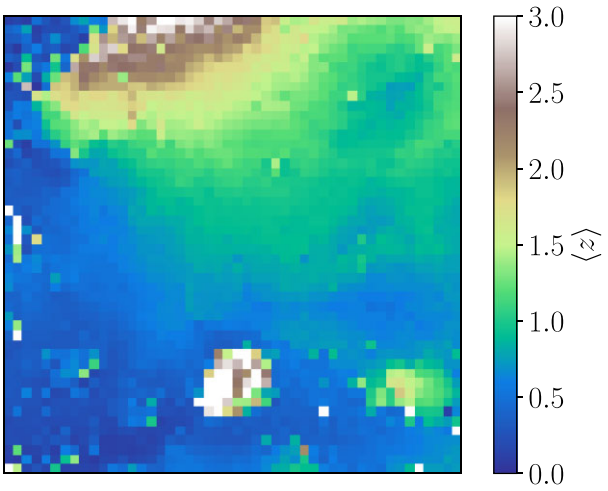


**Figure A3.** Deep SOM mean redshift. The modified terrain colour map highlights the different mean redshift levels, with the flooded area roughly showing redshifts below our samples and our high-$z$ galaxies lifting out of the oceans of low-$z$ galaxies. The grassy area roughly shows the redshifts of our first two tomographic bins, while the north-northwest hill shows the area of our highest redshift bin. Going south we find the snowed peaky island showing the area of very high redshift Lyman-break galaxies, with very low redshift Balmer-break galaxies lurking below the icy glaciers of Lyman-break galaxies.

cannot be estimated directly, which is a source of potential systematic uncertainty. We test this effect by calculating the mean redshift bias in Dark Green cells from neglecting the bin conditionalization, and extrapolating it to other nearby cells using a Gaussian smoothing. The last row of panels in Fig. A2 shows the bias values from extrapolation for every deep cell, showing that certain groups of cells have under-/overestimated mean redshifts. We parametrize this possible systematic bias with the same parameter $\epsilon$ that shifts the $p(z|c) \rightarrow p(z - \epsilon(\beta, c)|c)$ of each deep cell; with $\epsilon(\beta, c) = \beta\, b(c)$; and $b(c)$ the estimated systematic bias from the last row of panels in Fig. A2. We place a Gaussian prior on this parameter and marginalize over it, $p(\beta) = N(\mu = 1, \sigma = 1)$. Fig. 12 shows that this missing selection effect (labelled as BCE in the figure) has a very negligible effect to all the $N(z)$ parameters relative to the other sources of uncertainty.

### *A3.1 Cell conditionalization*

An additional source of systematic error comes from the approximation of using bin conditionalization (or bincond, equation A6) instead of the exact *cell conditionalization* (or cellcond, equation A5). Fig. A4 explores the difference in mean redshift for (i)/Dark Green cells between using cellcond and bincond. We find a clear (but somewhat noisy) trend, where cells within a tomographic bin with a lower than average mean redshift have an overestimated mean redshift, and vice versa, as expected. The overall trend within the same tomographic bin is centred around 0, as bincond already corrects for most of the overall redshift selection effect bias.

We have calculated the $N(z)$ using cellcond, and despite the large biased trend seen in Fig. A4, we have found that the resulting $n(z)$ from using cellcond presents very similar mean redshift, width and low redshift fraction values to those obtained from just using bincond. Upon closer inspection, the $p(z|c, \hat{c})$ and $p(z|c, \hat{b})$ distributions differ at their tails, which produces significant changes to their mean redshifts $\langle z|c, \hat{c}\rangle$ and $\langle z|c, \hat{b}\rangle$, but this effect ends up cancelling out after adding up the contributions from each deep field cell to

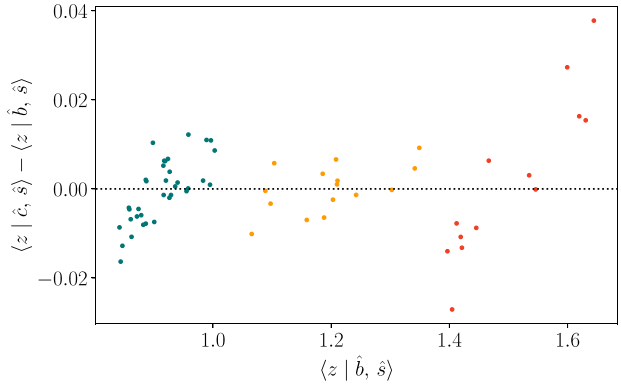


**Figure A4.** Mean redshift difference for each wide field cell between using cell conditionalization (cellcond) or bin conditionalization (bincond). Cellcond refers to conditioning the redshift distribution of Deep SOM cells to galaxies that are selected into each Wide SOM cell (i.e. using equation A5). In contrast, Bincond only requires galaxies to have been selected into any Wide SOM cell belonging to the tomographic bin (i.e. using equation A6). As expected, Wide SOM cells with lower redshift within the bin have a lower estimated mean redshift when we additionally require deep field galaxies to be selected into that particular Wide SOM cell. See Section A3.1 for more details.

calculate the final $N(z)$ for each bin. Although this effect cancels for these samples and for the summary statistics relevant to this work, we suggest it should be verified for other galaxy samples.

## A4 Zero-point uncertainty

As measured in Hartley et al. (2022), the deep field photometry has some residual photometric zero point error. This error is largest in the $u$-band (0.055), and much smaller in the other bands: 0.005 in *griz* and 0.008 in *JHK* (table 5 in Hartley et al. (2022)). This in principle impacts our analysis in two ways. First, most of the redshift information is in the COSMOS fields, while X3, C3, E2 have little or no redshift information. Therefore, we are extrapolating the redshift information measured in one field to the colours of all fields, and measuring the colour abundance from all fields. The zero-point uncertainty affects the accuracy of this extrapolation, as well as the measured deep colour abundance. On the other hand, a zero-point error on the deep field fluxes introduces an error in the input injected model fluxes used by BALROG, which in turn will induce a slight error on the distribution of recovered wide field BALROG fluxes. Since the error in the $u$ band is the largest, there is no $u$ band in the wide field, and the zero-point errors in *griz* are small, we assume the former is the only form of zero-point error we need to worry about.

Since the zero-point photometric uncertainty is mainly measured from the variance of the stellar and red galaxy loci between each band and field (for full details see Hartley et al. 2022), we perturb the zero-point magnitude of each deep field (X3, C3, E2, COSMOS) and band by an amount drawn from a Gaussian distribution with zero mean and variance equal to the measured variance from Hartley et al. (2022). Since only the relative zero-point matters, we fix the zero-point of one of the fields (COSMOS) and perturb the zero-point of the remaining fields (X3, C3, E2). We marginalize over this uncertainty by (i) drawing 3 zero-point shifts for each X3, C3, E2 field, (ii) we modify the fluxes and flux errors by the corresponding amount, (iii) we reassign each galaxy to the Deep SOM based on the perturbed fluxes, and (iv) we re-calculate the $n(z)$ based on this new assignment.

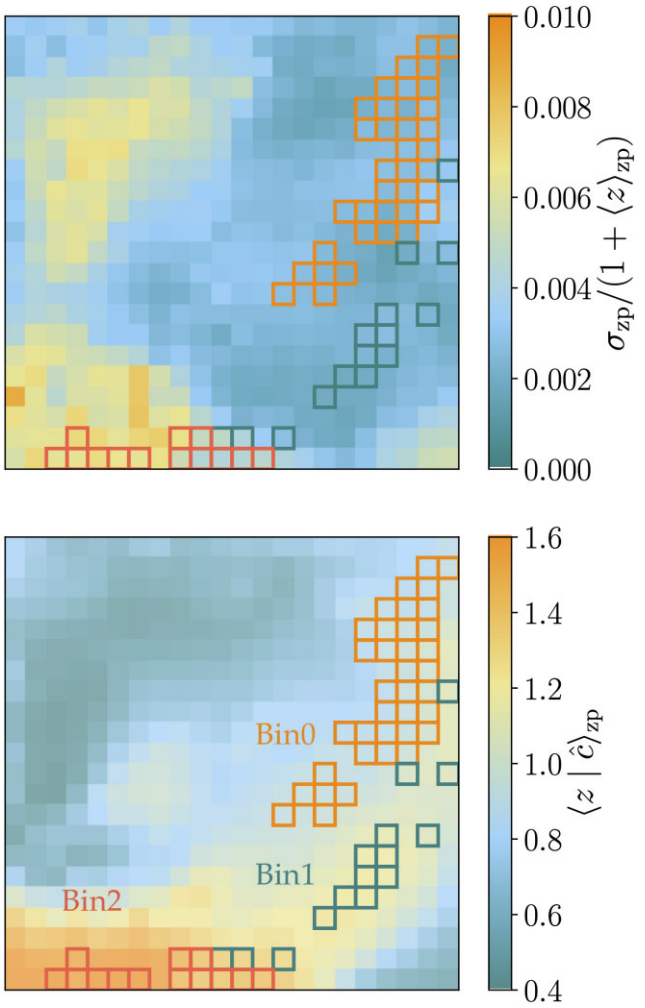


**Figure A5.** Wide SOM mean redshift variance from zero-point photometric uncertainty. The top panel shows the scatter in the mean redshift of each Wide SOM cell $\sigma_{zp}$ from perturbing the fluxes of Deep SOM galaxies with the zero-point photometry uncertainty, weighted by $1 + \langle z\rangle_{zp}$, with the latter being the average mean redshift from the same variations. The bottom panel shows $\langle z\rangle_{zp}$ for each Wide SOM cell for reference. Overlaid we can see the cells of the Wide SOM that constitute the three tomographic bins used in this work, following the procedure described in Section 3.3 and Fig. 4. See Section A4 for more details.

Fig. A5 shows the resulting variance in mean redshift for each Wide SOM cell in the top panel, as a result of perturbing the fluxes of the deep field galaxies. The average mean redshift shown in the bottom panel for reference, with the cells pertaining to each tomographic bin indicated with different colours. As expected, we find a large effect in cells with a low redshift, as the *u*-band uncertainty is the largest, which affects the classification of low redshift galaxies. We also find a large effect in some of the wide cells that have a high mean redshift but that are next to wide cells with low redshift, i.e. cells that are near colour-redshift degeneracies.

### A5 Redshift uncertainty parameter priors

To estimate the priors $p(\Delta z^i)$, $p(\sigma_z^i)$ and $p(A^i_{\mathrm{low}-z})$ on these parameters we draw $\{n^k(z)\}$ samples from the sources of uncertainty described in this Appendix, and for each individual realization $k$ we calculate its summary statistics $\{\Delta_z^{i,k}, \sigma_z^{i,k}, \mathrm{A}^{i,k}_{\mathrm{low-z}}\}$. We find

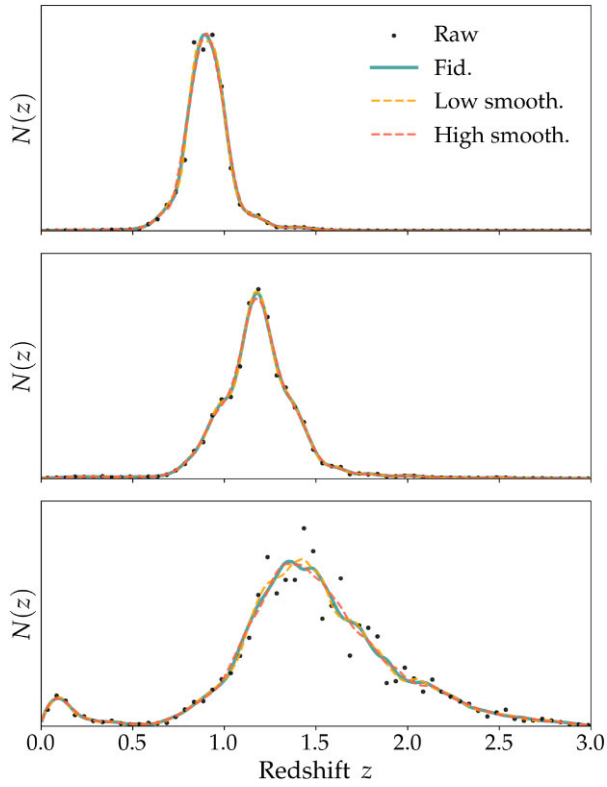


**Figure A6.** Visualization of the smoothing procedure applied to the raw redshift distributions using a Savitzsky Golay (SG) filter. A fiducial set of distributions is presented, along with two alternative sets using lower (higher) amounts of smoothing, as described in Appendix A6.

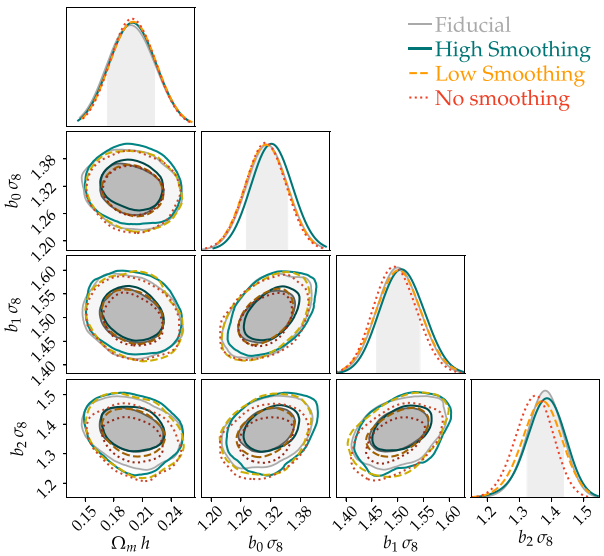


**Figure A7.** Comparison of the parameter constraints from galaxy clustering using higher and lower amounts of $N(z)$ smoothing, and no smoothing, as described in Appendix A6, demonstrating the small impact of the smoothing step in the analysis.

that the sampled distributions look nearly Gaussian (Fig. 12), and therefore we assume a Gaussian distribution for each $p(\Delta z^i)$, $p(\sigma_z^i)$ and $p(A^i_{\mathrm{low}-z})$, and calculate their mean and width from the average and standard deviation of the sampled values.

In total, we have 24 zero point systematic shifts (8 bands and 3 out of 4 fields), two redshift systematic shifts (one for `COSMOS2015` and one for `PAUS + COSMOS`) and one selection effect bias parameter. We draw 100 samples in quantile space using Latin hypercube sampling, a stratified random sampling technique for generating near-random samples of parameter values that is more efficient than a pure random sampling. For each of these 100 samples we shift the $p(z)$ of individual galaxies, we shift the deep fluxes of galaxies and reassign them to deep cells. Then for each of these 100 samples we generate 5,000 N(z) samples using 3sDir. We properly weight deep field galaxies injected by Balrog by the clustering weight (Section 4) of the spot where they were injected. We produce samples for all the area, and the North (*Planck*) and South (SPT) regions.

The fiducial redshift distribution $F(z)$ of equation (A1) is the average $N(z)$ of the distribution samples with an additional smoothing. We apply a Savitzky-Golay filter on the average $N(z)$, using a 0.21 smoothing length in redshift for Bins 0 and 1, while for Bin 2 we use a combination of two smoothing lengths: we use a length of 0.21 at $z < 0.5$ and a length of 0.45 for $z > 0.5$.

### A6 Smoothing of the redshift distributions

The redshift inference methodology described in Section 3 is subject to effects of shot noise and especially sample variance in the redshift samples (Sánchez et al. 2020), which result in noisy estimates of the redshift distributions of our tomographic bins. The uncertainties coming from these effects are properly taken into account in Section 5. In addition, we also apply a smoothing procedure to the redshift distributions used in this work, since noise in the redshift distributions might cause instabilities in the analysis of galaxy clustering. For that purpose, we apply a Savitzsky Golay (SG) filter with a third-order polynomial to the raw redshifts distributions, as depicted in Fig. A6. In our fiducial case, the length of the filter window is set to 0.21 in redshift for the low redshift part of the distributions ($z < 0.5$), and 0.45 in redshift for the higher redshift part of the distributions ($z > 0.5$). In order to test the stability of our results to the particular smoothing filter choices, we define two alternative sets of smoothed redshift distributions, corresponding to lower (higher) smoothings, using SG filters with window lengths of 0.15 (0.27) in redshift for the low redshift part of the distributions ($z < 0.5$), and 0.27 (0.55) in redshift for the higher redshift part of the distributions ($z > 0.5$). The comparison between the raw estimates and the smoothed versions of the redshift distributions for the three tomographic bins is shown in Fig. A6. Then, in Fig. A7, we test the impact of the smoothing step in the parameter constraints from the galaxy clustering measurements used in this paper, and find negligible impact, even for the case of no smoothing of the redshift distributions.

## APPENDIX B: COMPARISON BETWEEN JK AND THEORY COVARIANCE

In this section, we compare the two covariance estimates (based on Gaussian simulations, and based on Jackknife estimates) presented in Section 7. In order to generate each realization of the Gaussian simulations, we generate a set of four maps following the procedure detailed in Giannantonio et al. (2008). In order to obtain correlated maps with the correct power spectrum, we have to generate a set of correlated (in-phase) screens with an amplitude $T_{i,k}$, where the

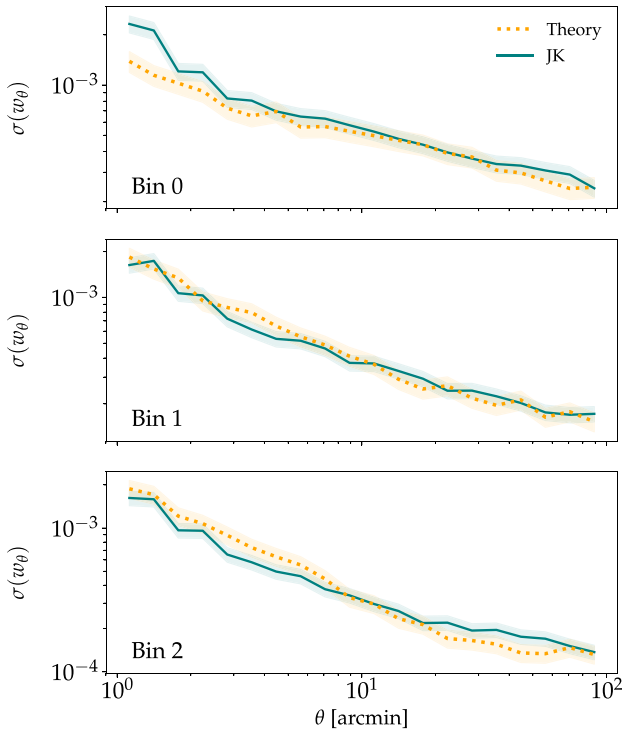


**Figure B1.** Comparison of the diagonal elements of the theory and jackknife covariance matrices for the autocorrelations of angular galaxy clustering for the three redshift bins (0,1,2) defined in this work. The methodology for the measurements and covariance can be found in Section 7.2 and Appendix B.

subindex $i$ refers to the final map, and $k$ to the phase. So we add all contributions with the same index $i$ to get the $i$th map, and all screens that have the same index $k$ are generated using the same random seed (are in-phase). Each screen is generated using `hp.anafast(T**2_ij, nside)`. The amplitudes $T_{ik}$ are calculated as follows:

$$T_{1a} = \sqrt{C_\ell^{00}}, \tag{B1}$$

$$T_{2a} = \frac{C_\ell^{01}}{T_{1a}}, \tag{B2}$$

$$T_{2b} = \sqrt{C_\ell^{11} - T_{2a}^2}, \tag{B3}$$

$$T_{3a} = \frac{C_\ell^{02}}{T_{1a}}, \tag{B4}$$

$$T_{3b} = \frac{C_\ell^{12} - T_{2a}}{T_{3a}}, \tag{B5}$$

$$T_{3c} = \sqrt{C_\ell^{22} - T_{3a}^2 - T_{3b}^2}, \tag{B6}$$

$$T_{4a} = \frac{C_\ell^{0\kappa}}{T_{1a}}, \tag{B7}$$

$$T_{4b} = \frac{C_\ell^{1\kappa} - T_{2a}T_{4a}}{T_{2b}}, \tag{B8}$$

$$T_{4c} = \frac{C_\ell^{2\kappa} - T_{3a}T_{4a} - T_{3b}T_{4b}}{T_{3c}}, \tag{B9}$$

$$T_{4d} = \sqrt{C_\ell^{\kappa\kappa} - T_{4a}^2 - T_{4b}^2 - T_{4c}^2}. \tag{B10}$$

We generate 100 realizations of these maps, and get their covariance. We compare the resulting covariance with the Jackknife estimate in

Fig. B1. In this figure, we can see that the diagonal terms from both covariance estimates are in excellent agreement in the range of scales that we are considering.

## APPENDIX C: FALSE CORRECTION TEST OF NN-WEIGHTS

In this section, we verify that the neural network architecture and the k-fold cross-validation described in Section 4.2 do not introduce significant artificial correlations due to overfitting or from treating pixels with $n > 2$ galaxies as $n = 1$. In general, correction methods work by removing the spurious clustering introduced by varying observing conditions, therefore reducing the amplitude of $w(\theta)$ after correction has been applied. Here, we use unaltered simulations to test if the NN introduces any overcorrection.

We use the public MICEv2 simulations (Fosalba et al. 2015) and perform the following cuts: $0.8 <$ `z_cgal` $< 1.0$ and $18 <$ `des_asahi_full_i_true` $< 23.5$. Since MICE only spans one octant, which we replicate and mirror it 8 times so it spans the full sky, after which we apply the High-z redshift samples mask and subsample to the galaxy density of Bin 0. We use this unaltered mock catalog, together with the same DES survey property maps from Section 4.1, to train the neural network weights. We do this multiple times to account for possible run-to-run variations of the NN best-fitting weights.

Fig. C1 shows the ratio between the simulated $w(\theta)$ and the corrected one after training the NN and applying the resulting weights. The black bands show the distribution for different multiple training runs. Values greater than unity indicate an overcorrection of the neural network training. We find the overcorrection is well within the statistical uncertainty in the scales of interest, always lower than 40 per cent of the statistical uncertainty at any given used scale, and much smaller than that in the combined set of used angular scales.

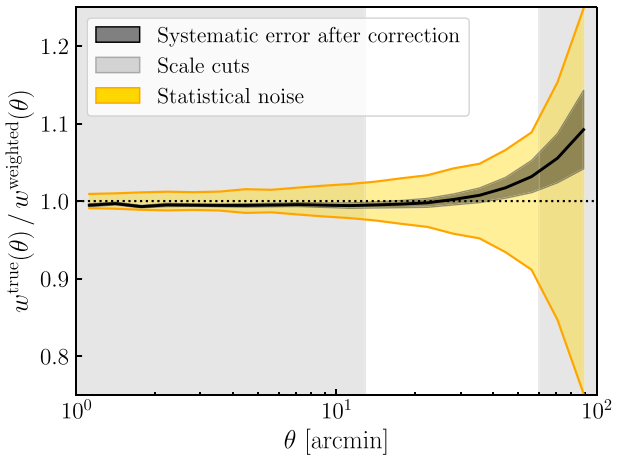


**Figure C1.** False detection bias in unaltered mock catalogues. It shows the simulated angular clustering from the unaltered mock catalog $w^{\rm true}(\theta)$ divided by the same angular clustering corrected by the NN-weights $w^{\rm weighted}(\theta)$. The black bands show the $\pm 1\sigma$ distribution for multiple runs of the NN. Values greater than unity indicate an overcorrection of the neural network training. The yellow bands indicate the statistical uncertainty using the diagonal of the covariance, relative to the measured angular clustering of Bin 0. The bias is well within the statistical uncertainty, and always less than 40 per cent of the statistical uncertainty in our fiducial scales.

[1]*Departament de Física, Universitat Autònoma de Barcelona (UAB), E-08193 Bellaterra (Barcelona), Spain*
[2]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*
[3]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain*
[4]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[5]*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA*
[6]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA*
[7]*Department of Physics, University or Genova and INFN, Genova Division, Via Dodecaneso 33, I-16146, Genova, Italy*
[8]*Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*
[9]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[10]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[11]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[12]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid 28040, Spain*
[13]*Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*
[14]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[15]*Kavli Institute for Particle Astrophysics & Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA*
[16]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[17]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[18]*Physics Department, 2320 Chamberlin Hall, University of Wisconsin-Madison, 1150 University Avenue Madison, WI 53706-1390, USA*
[19]*Department of Physics, Northeastern University, Boston, MA 02115, USA*
[20]*NASA Goddard Space Flight Centre, 8800 Greenbelt Rd, Greenbelt, MD 20771, USA*
[21]*Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15312, USA*
[22]*Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*
[23]*Laboratório Interinstitucional de e-Astronomia – LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ-20921-400, Brazil*
[24]*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[25]*Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark Street, Urbana, IL 61801, USA*
[26]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 West Green Street, Urbana, IL 61801, USA*
[27]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[28]*Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[29]*NSF AI Planning Institute for Physics of the Future, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[30]*Department of Astronomy/Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA*
[31]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*
[32]*Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[33]*Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*
[34]*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany*
[35]*School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK*

[36]*Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland*
[37]*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*
[38]*Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK*
[39]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[40]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[41]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[42]*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*
[43]*Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA*
[44]*Department of Physics, Duke University Durham, NC 27708, USA*
[45]*Cerro Tololo Inter-American Observatory, NSF's National Optical-Infrared Astronomy Research Laboratory, Casilla 603, La Serena, Chile*
[46]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[47]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[48]*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[49]*Physics Department, William Jewell College, Liberty, MO 64068, USA*
[50]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[51]*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*
[52]*Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy*
[53]*INAF − Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[54]*Institute for Fundamental Physics of the Universe, Via Beirut 2, I-34014 Trieste, Italy*
[55]*Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany*
[56]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
[57]*Université Grenoble Alpes, CNRS, LPSC-IN2P3, F-38000 Grenoble, France*
[58]*Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029 Blindern, NO-0315 Oslo, Norway*
[59]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain*
[60]*School of Mathematics and Physics, University of Queensland, Brisbane QLD 4072, Australia*
[61]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[62]*Centre for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*
[63]*Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia*
[64]*Lowell Observatory, 1400 Mars Hill Rd, Flagstaff, AZ 86001, USA*
[65]*George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*
[66]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[67]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[68]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[69]*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[70]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*
[71]*Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK*

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