Mon. Not. R. Astron. Soc. **000**, 1–23 (0000) Printed 16 December 2015 (MN LaTeX style file v2.2) FERMILAB-PUB-15-305 (accepted)

Galaxy clustering, photometric redshifts and diagnosis of systematics in the DES Science Verification data

M. Crocce[1⋆], J. Carretero[1,2], A. H. Bauer[1], A. J. Ross[3], I. Sevilla-Noarbe[4,5], T. Giannantonio[6,42], F. Sobreira[7,8], J. Sanchez[4], E. Gaztanaga[1], M. Carrasco Kind[5,10], C. Sánchez[2], C. Bonnett[2], A. Benoit-Lévy[9], R. J. Brunner[5,10], A. Carnero Rosell[8,11], R. Cawthon[12], P. Fosalba[1], W. Hartley[13], E. J. Kim[5], B. Leistedt[9], R. Miquel[2], H. V. Peiris[9], W.J. Percival[15], R. Rosenfeld[8,16], E. S. Rykoff[17,18], E. Sánchez[4], T. Abbott[19], F. B. Abdalla[9,43], S. Allam[7], M. Banerji[6,20], G. M. Bernstein[21], E. Bertin[22,23], D. Brooks[9], E. Buckley-Geer[7], D. L. Burke[17,18], D. Capozzi[15], F. J. Castander[1], C. E. Cunha[17], C. B. D'Andrea[15], L. N. da Costa[8,11], S. Desai[24,25], H. T. Diehl[7], T. F. Eifler[21,26], A. E. Evrard[27,28], A. Fausti Neto[8], E. Fernandez[2], D. A. Finley[7], B. Flaugher[7], J. Frieman[7,12], D. W. Gerdes[28], D. Gruen[29,30], R. A. Gruendl[5,10], G. Gutierrez[7], K. Honscheid[3,31], D. J. James[19], K. Kuehn[32], N. Kuropatkin[7], O. Lahav[9], T. S. Li[33], M. Lima[8,34], M. A. G. Maia[8,11], M. March[21], J. L. Marshall[33], P. Martini[3,35], P. Melchior[3,31], C. J. Miller[27,28], E. Neilsen[7], R. C. Nichol[15], B. Nord[7], R. Ogando[8,11], A. A. Plazas[26], A. K. Romer[36], M. Sako[21], B. Santiago[8,37], M. Schubnell[28], R. C. Smith[19], M. Soares-Santos[7], E. Suchyta[3,31], M. E. C. Swanson[10], G. Tarle[28], J. Thaler[38], D. Thomas[15], V. Vikram[39], A. R. Walker[19], R. H. Wechsler[17,18,40], J. Weller[24,29,30], J. Zuntz[41]

(The DES Collaboration)

Affiliations are listed at the end of the paper

arXiv:1507.05360v2 [astro-ph.CO] 15 Dec 2015

16 December 2015

ABSTRACT

We study the clustering of galaxies detected at $i < 22.5$ in the Science Verification observations of the Dark Energy Survey (DES). Two-point correlation functions are measured using 2.3×10^6 galaxies over a contiguous 116 deg^2 region in five bins of photometric redshift width $\Delta z = 0.2$ in the range $0.2 < z < 1.2$. The impact of photometric redshift errors are assessed by comparing results using a template-based photo-z algorithm (BPZ) to a machine-learning algorithm (TPZ). A companion paper (Leistedt et al 2015) presents maps of several observational variables (e.g. seeing, sky brightness) which could modulate the galaxy density. Here we characterize and mitigate systematic errors on the measured clustering which arise from these observational variables, in addition to others such as Galactic dust and stellar contamination. After correcting for systematic effects we measure galaxy bias over a broad range of linear scales relative to mass clustering predicted from the Planck ΛCDM model, finding agreement with CFHTLS measurements with χ^2 of 4.0 (8.7) with 5 degrees of freedom for the TPZ (BPZ) redshifts. We test a "linear bias" model, in which the galaxy clustering is a fixed multiple of the predicted non-linear dark-matter clustering. The precision of the data allow us to determine that the linear bias model describes the observed galaxy clustering to 2.5% accuracy down to scales at least 4 to 10 times smaller than those on which linear theory is expected to be sufficient.

Key words: photometric surveys – galaxies clustering – systematic effects – large-scale structure of Universe.

1 INTRODUCTION

Vast galaxy surveys trace the large-scale structure (LSS) of the Universe at late times and therefore complement and improve the wealth of information already provided by cosmic microwave background (CMB) and supernovae experiments. In particular our understanding and characterization of the cosmic accelerated expansion. Imaging from multi-band photometry, e.g., SDSS (York et al. 2000), PanSTARRS (Kaiser, Tonry & Luppino 2000), KiDS (de Jong et al. 2013), HSC (Miyazaki et al. 2012) and the planned LSST (Tyson et al. 2003), provides the angular positions and detailed color information of the galaxies. From this color information, photometric redshifts (photo-z) can be measured for each galaxy providing distance estimates that have typically low resolution. While obtaining

⋆ crocce@ice.cat

© 0000 RAS

detailed radial information requires a spectroscopic redshift survey, e.g., 2dF (Colless et al. 2001), VVDS (Le Fèvre et al. 2005), WiggleZ (Drinkwater et al. 2010) and BOSS (Dawson et al. 2013), obtaining spectra is more time consuming. Hence the data volume available to a photometric redshift sample will naturally far exceed that of a spectroscopic one. Many LSS studies have already obtained constraints on cosmological parameters from photo-z samples, such as the sum of neutrino masses, the matter/energy content of the Universe, and the nature of dark energy (Padmanabhan et al. 2007; Thomas, Abdalla & Lahav 2010; Crocce et al. 2011; Ho et al. 2012; Carnero et al. 2012; de Simoni et al. 2013).

The Dark Energy Survey (DES, Flaugher (2005)) will image 5000 deg^2 of the South Galactic Cap to a depth of $i < 24$, recording 300 million galaxies in 5 broadband filters ($grizY$), thereby providing high quality photometric redshifts up to redshift $z = 1.4$ (Sánchez et al. 2014). The DES camera (DECam, Flaugher et al. 2015; Diehl & and the Dark Energy Survey Collaboration 2012) was installed and commissioned in the second semester of 2012. A Science Verification (SV) period of observations followed, lasting from November 2012 to February 2013. The DES observations officially started in late August 2013.

In this paper, we analyze the clustering of galaxies observed during the SV period. The SV region has been observed to match the nominal 24th mag depth in the i- band expected for the full DES survey, and we can therefore use the data to characterize the properties of the galaxy samples that DES can reliably observe. We perform detailed systematic tests on the galaxy samples and develop methods to mitigate systematic effects on the measured clustering. We asses the impact on clustering measurements from uncertainties in photometric redshift estimation by obtaining our samples using two different photo-z algorithms, one based on a template fitting method and another on a machine learning one. Each of these methods has advantages and disadvantages, which we discuss. Our resulting clustering measurements allow us to characterize the evolution of bias as a function of redshift, to assess the validity of linear bias models and to provide baseline clustering measurements that other DES studies using SV data can use to compare against (e.g., Giannantonio et al. 2015; Suchyta et al. 2015). We were able to complete this work without the benefit of a large number of simulated DES SV galaxy samples, but future DES LSS studies (covering larger volumes) will rely on such simulations.

The outline of the paper is as follows: we introduce the clustering estimators and the theory used throughout the paper in Section 2, describe our data set in Sections 3 and 4, and present the maps of potential systematics in Section 5. In Section 6 we discuss the extent and mitigation of possible systematics, summarize and discuss our results in Section 7, before concluding in Section 8.

2 ANGULAR CLUSTERING: THEORY AND ESTIMATORS

In this section we review the modeling of the angular correlation function and its covariance matrix, in photo-z bins. We also describe the algorithms used to estimate these quantities from the data.

2.1 Two-point Angular Correlations: Modeling

Galaxy clustering can be modeled starting from the dark matter overdensity field at the angular position $\hat{\mathbf{n}}$ and at redshift z: $\delta(\hat{\mathbf{n}}, z)$. If we assume a linear bias model, the projected overdensity of our galaxy sample is given by,

$$\delta_g(\hat{\mathbf{n}}) = \int_0^\infty b(z)\, \frac{dn}{dz}(z)\, \delta(\hat{\mathbf{n}}, z)\, dz\,. \tag{1}$$

where dn/dz is the probability of detecting a galaxy at a given redshift (i.e. the normalised redshift distribution) and $b(z)$ is the bias. The angular two-point correlation function between redshift bins A and B is then defined as,

$$\begin{aligned} w_{AB}(\theta) &\equiv \langle \delta_g(\hat{\mathbf{n}})\delta_g(\hat{\mathbf{n}}+\hat{\theta})\rangle = \\ &= \int dz_1\, \phi_A(z_1) \int dz_2\, \phi_B(z_2)\, \xi(r_{12}(\theta), \bar{z}) \end{aligned} \tag{2}$$

where $r_{12}^2 = r(z_1)^2 + r(z_2)^2 - 2r(z_1)r(z_2)\cos(\theta)$ with $r(z)$ being the comoving distance to redshift z (assuming a flat Universe), and

$$\phi(z) = \frac{D(z)}{D(\bar{z})} b(z) \frac{dn}{dz}(z),$$

where $D(z)$ is the linear growth factor. Equation (2) should in principle be written in terms of a spatial correlation function ξ that evolves with redshift. Instead we evaluate the spatial correlation at some mean redshift $\bar{z}$ and encode the growth evolution relative to this redshift into the selection function ϕ (this is exact at the linear level) which we make proportional to $D(z)/D(\bar{z})$. In turn, rather than assuming a parametric function for $b(z)$, we will measure a single bias value at each tomographic photo-z bin. Lastly, the way we estimate the redshift distribution of the sample, dn/dz, is described in Sec. 4 and shown in Fig. 3.

We calculate the linear and nonlinear power spectra using CAMB (Lewis et al. 2000). For nonlinear dark matter clustering we use the recently re-calibrated `Halofit` prescription (Takahashi et al. 2012) built into CAMB. We then Fourier transform these into configuration space to obtain ξ and evaluate Eq. (2). Even though the effect of redshift space distortions is negligible for the angular scales considered in this paper, we do include them in our predictions (e.g. see formulae in Crocce, Cabré & Gaztañaga 2011).

Throughout the paper we assume a fiducial flat ΛCDM+ν (one massive neutrino) cosmological model based on *Planck* 2013 + *WMAP* polarisation + ACT/SPT + BAO, with parameters (Ade et al. 2014) : $\omega_b = 0.0222$, $\omega_c = 0.119$, $\omega_\nu = 0.00064$, $h = 0.678$, $\tau = 0.0952$, $n_s = 0.961$ and $A_s = 2.21 \cdot 10^{-9}$ at a pivot scale $\bar{k} = 0.05$ Mpc^{-1} (yielding $\sigma_8 = 0.829$ at $z = 0$) where $h \equiv H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\omega_i \equiv \Omega_i h^2$ for each species i.

2.2 Two-point Angular Correlations: Estimators

In this paper we use two different estimators for angular correlation functions. One is a direct pair-count algorithm particularly suited to investigating small-scale clustering, and another based on pixelized maps using `HEALPix` which is particularly useful to investigate the cross-correlation of our galaxy sample with maps of potential systematics (such as observing conditions) in an efficient way at the expense of decreased angular resolution.

The pair-counting algorithm uses the publicly available tree code ATHENA [1] to measure the angular correlation function with the standard Landy & Szalay (1993) estimator,

$$w(\theta) = \frac{DD - 2DR + RR}{RR} \tag{3}$$

where DD, DR and RR represent the number of pairs of objects, taken from the galaxy catalogue (D) or from a randomly generated sample covering the angular footprint (R), which are separated by a scale θ within a bin size $\Delta\theta$. The random points are distributed uniformly over the footprint. As we discuss in Sec. 4.3 our angular footprint considers only regions of the survey that are deeper than the

[1] `www.cosmostat.org/software/athena`

flux limit in the sample definition, therefore the distribution of random points is uniform in regions where the sample is complete.

The pixel-based estimator for the correlation between maps 1 and 2 is given by,

$$w_{1,2}(\theta) = \sum_{i=1}^{i=N_{\rm pix}} \sum_{j=i}^{j=N_{\rm pix}} \frac{(N_{i,1}-\bar{N}_1)(N_{j,2}-\bar{N}_2)}{\bar{N}_1\bar{N}_2}\Theta_{i,j}, \quad (4)$$

where $N_{\rm pix}$ is the number of pixels used, N_i is the value in pixel i of the quantity of interest (either the number of galaxies or, e.g., the seeing), $\bar{N}$ represents the mean over all pixels, and $\Theta_{i,j}$ is 1 if pixels i, j are separated by a scale θ within a bin size $\Delta\theta$ and is zero otherwise. The sum only runs through those pixels that fall within the footprint.

2.3 Covariance Matrix

In order to estimate the covariance matrix for our $w(\theta)$ measurements we combine two different approaches: a modeling of the off-diagonal elements and the jackknife (JK) method for the variance.

The jackknife technique, which has been widely applied to angular galaxy clustering measurements at small scales (e.g., Scranton et al. 2002; Myers et al. 2006; Ross & Brunner 2009; Coupon et al. 2012) works by splitting the footprint into $N_{\rm JK}$ small equal-area sections and recomputing $w(\theta)$ with each section removed. The covariance matrix between $w(\theta)$ measurements at two angular separations between two redshift bins, A and B, is estimated by,

$$\mathrm{Cov}^{JK,AB}_{\theta_i,\theta_j} = \frac{N_{\rm JK}-1}{N_{\rm JK}} \sum_{k=1}^{N_{\rm JK}} \Delta w^A[\theta_i]\,\Delta w^B[\theta_j], \quad (5)$$

$$\Delta w^{A(B)}[\theta] = \left(w_k^{A\,(B)}[\theta] - \bar{w}^{A\,(B)}[\theta]\right), \quad (6)$$

where w_k is the correlation function measured with the k-th JK region removed and $\bar{w}$ is the mean of the $N_{\rm JK}$ jackknife $w(\theta)$'s. In our case we use 40 JK regions, in such a way that each patch is $\sim 2\,\mathrm{deg}$ across, matching our largest scales of interest.

Theoretical estimates of the covariance for angular clustering have also been explored and tested (e.g. Dodelson (2003); Cabré et al. (2007); Crocce, Cabré & Gaztañaga (2011); Ross et al. (2011b)). They rely on the assumption that errors scale as the square root of the survey area $\mathcal{C}_{\theta_i,\theta_j} \sim f_{sky}^{-1}\mathcal{C}^{\rm full\,sky}_{\theta_i,\theta_j}$. The covariance for a full-sky survey is then easily derived considering that angular power spectrum measurements C_ℓ are uncorrelated, $\mathrm{Cov}_{\ell\ell'} = \mathrm{Var}(C_\ell)\delta_{\ell\ell'}$ with $\mathrm{Var}(C_\ell) = 2\,C_\ell^2/(2\ell+1)$. One can then use a Legendre transform to express this error in real space for $\mathcal{C}^{\rm full\,sky}_{\theta_i,\theta_j}$ and arrive to,

$$\mathrm{Cov}^{TH}_{\theta_i,\theta_j} = \frac{2}{f_{sky}} \sum_{\ell\geqslant 0} \frac{2\ell+1}{(4\pi)^2} P_\ell(\cos\theta_i)P_\ell(\cos\theta_j)(C_\ell + 1/\bar{n})^2 \quad (7)$$

for auto-correlations in one bin, and

$$\mathrm{Cov}^{TH,AB}_{\theta_i,\theta_j} = \frac{1}{f_{sky}} \sum_{\ell\geqslant 0} \frac{2\ell+1}{(4\pi)^2} P_\ell(\cos\theta_i)P_\ell(\cos\theta_j) \left[\left(C_\ell^{AB}\right)^2 + \left(C_\ell^{AA} + 1/\bar{n}_a\right)\left(C_\ell^{BB} + 1/\bar{n}_b\right)\right] \quad (8)$$

for cross-correlations between bins A and B. In Eqs. (7,8) we have included the standard shot-noise contribution arising in the variance of the C_ℓ estimates ($\bar{n}$ is the number of objects per steradian). The angular power spectrum between bins A and B is given by

$$C_\ell^{AB} = \frac{2}{\pi}\int dk k^2 P(k,\bar{z}) W_\ell^A(k) W_\ell^B(k) \quad (9)$$

with the kernel defined for each bin by

$$W_\ell(k) = \int dz\, b(z)\, \frac{dn}{dz}(z)\, \frac{D(z)}{D(\bar{z})}\, j_\ell(kr(z)) \quad (10)$$

As before, we evaluate the power spectrum in Eq. (9) at some mean redshift and then compute the growth within W_ℓ with respect to this reference point. In Eq. (10) each bin is characterized by its dn/dz and b, while j_ℓ are the spherical Bessel functions.

Measurements of the correlation function at different angular scales are considerably correlated. The JK estimation is limited by the number of independent samples one can build out of the given footprint and hence yields noisy off-diagonal correlations. Obtaining good estimates of the off-diagonal elements is important, as the correlation $w(\theta)$ between adjacent θ bins are typically greater than 0.9. For the diagonal elements of the covariance matrix, we expect the JK estimation to be adequate and favor it over the theoretical estimation as it is extracted from the data itself and hence can trace effects such as non-linearities, the effect of the mask, and residual systematic variation. The theory estimate on the other hand is adequate for off-diagonal elements because it avoids the intrinsic noise from the limited number of realizations. Therefore we combine both approaches by defining our covariance matrix as,

$$\mathrm{Cov}^{MIX}_{\theta_i,\theta_j} = \mathrm{Corr}^{TH}_{\theta_i,\theta_j}\sigma_{JK}(\theta_i)\sigma_{JK}(\theta_j) \quad (11)$$

where $\sigma^2_{JK}(\theta_i) \equiv \mathcal{C}^{JK}_{\theta_i,\theta_i}$ from Eq. (5) and $\mathrm{Corr}^{TH}_{\theta_i,\theta_j}$ is the correlation matrix (or reduced covariance matrix) defined by

$$\mathrm{Corr}^{TH}_{\theta_i,\theta_j} = \frac{\mathrm{Cov}^{TH}_{\theta_i,\theta_j}}{\sigma_{TH}(\theta_i)\sigma_{TH}(\theta_j)} \quad (12)$$

from Eqs. (7,8). We refer to this as the "mixed" approach.

The methodology we use to model our covariance matrix is further justified by the results presented in a companion paper (Giannantonio et al. 2015), where we present a discussion of various estimators for the covariance matrix of auto-correlations. As naively expected, it is found that the JK method performs better than theory errors for diagonal correlations (i.e. variance) when both are compared to a covariance matrix derived from an N-body simulation. But under-performs compared to theory ones for the off-diagonal elements. In Appendix B we show that on large scales both the JK and theory approaches yield similar results for the variance, and for derived best-fit biases.

3 THE DES SCIENCE VERIFICATION (SVA-1) PHOTOMETRIC SAMPLE

3.1 The Science Verification dataset

The Science Verification (SV) observations, taken over 78 nights in 2012 and 2013, provided science-quality data for more than 250 deg^2 at close to the nominal depth of the survey. The SV data footprint was chosen to contain areas already covered by several deep spectroscopic galaxy surveys, including VVDS (Le Fèvre et al. 2005), zCOSMOS (Lilly et al. 2007), and ACES (Cooper et al. 2012), which together provide a calibration sample for the DES photometric redshifts (Sánchez et al. 2014). In addition to these, two large contiguous regions (termed SPT-E and SPT-W, due to their partial coverage of the South Pole Telescope (Carlstrom et al. 2011) fields) were observed.

3.2 Image processing: the DESDM pipeline

The raw images taken each night by DECam are sent for processing to the National Center for Supercomputer Applications (NCSA[2], Urbana-Champaign, United States) using the DES Data Management (DESDM) pipeline described in Sevilla et al. (2011) and Mohr et al. (2012). We briefly describe the process below.

[2] http://www.ncsa.illinois.edu/

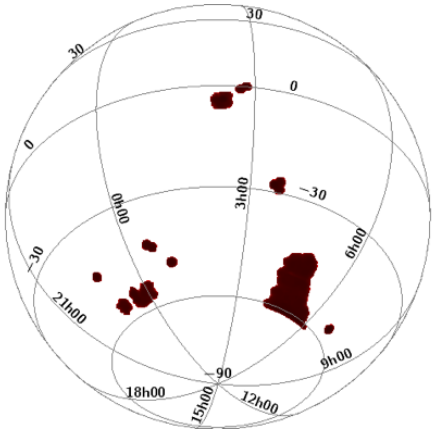

Figure 1. Footprint of the SVA1 Gold catalog. In this analysis, we only use data from the largest compact footprint occupying the lower-right, known as 'SPT-E'.

These single-epoch images are 'detrended' by applying bias and flat-fields corrections. Then, cross-talk, pupil, illumination, fringing and non-linearity corrections are carried out. Bad pixels are flagged and traced in a map while satellite trails, cosmic rays and saturation effects from bright stars are detected and flagged. The astrometric calibration is performed using `SCAMP` (Bertin 2006), fitting the best local-to-celestial reference system transformation using the single-epoch image positions of bright sources. One then relates them to reference celestial positions, correcting for distortions from the large field-of-view. Nightly absolute calibration is performed using reference stars which are observed every night, to obtain a photometric solution for that period including the zero point, a color, and airmass terms for each CCD.

The resulting reduced images are then re-mapped into a uniform pixel grid in which the coaddition of said images is performed. The purpose of this process is to obtain increased depth.

Lastly, the catalog creation stage uses the `SExtractor` (Bertin & Arnouts 1996) software , which detects objects on the coadded images for each band and stores them as unique single object entries at a database at NCSA. During this procedure, the coadded astrometric and photometric information is measured and stored as well. The astrometry is accurate to $0.2''$. A global photometric relative calibration step is performed for this dataset using the exposure overlap across different fields, which provides repeated observations of the same stars so a minimization procedure can be performed (Glazebrook et al. 1994). Observations from PreCAM (Kuehn et al. 2013) provide a grid of standards to which this relative calibration can be anchored to in order to minimize large scale variations[3]. Finally, observations of spectrophotometric standards ties this photometry to the absolute AB system (Tucker et al. 2007). This final data set represents the "coadd" catalog.

3.3 The Gold catalog

The coadd catalog produced by DESDM of the SV dataset was analyzed and tested resulting in the generation of the *SVA1 Gold* catalog (Rykoff et al., in prep.) The footprint of the catalog in equatorial coordinates is shown in Fig.1.

The basic additions of this value added catalog include:

- Incorporating satellite trail and other artifact information to mask out specific areas;
- Removing areas where the colors are severely affected by stray light in the images and areas with a small exposure count (at the borders of the footprint);
- Removing the area below declination of -61° to avoid the increased stellar contamination in our galaxy catalog due to the Large Magellanic Cloud (LMC), as well as the area dominated by the nearby star R Doradus.
- Applying an additional Stellar Locus correction (High et al. (2009)) to tighten the calibration even further, ensuring the agreement of stellar colors with respect to those of reference stars. After this correction, magnitudes are calibrated to 2 per cent accuracy, with a 1 per cent accuracy in colors (Rykoff et al. in prep.).

After this selection, about 244 deg^2 remain in the catalog.

3.3.1 *Survey mask*

Concurrently with the generation of the catalog, another pipeline builds a nominal `MANGLE`[4] mask (Swanson et al. 2008) that takes into account the DECam CCD pointings and properties of the sky during each night. Artifacts such as airplane or satellite trails, cosmic rays, etc., and areas near bright stars are masked out. An estimation of the depth for the remaining regions is calculated as the magnitude at which an object is measured with signal-to-noise of 10 in a 2 arcsecs aperture. This aperture-based depth must be converted into a total-magnitude depth corresponding to `SExtractor`'s `MAG_AUTO` measurements used in the galaxy selection. This conversion, described in Rykoff et al. (in prep.), takes into account properties of the data such as the seeing, which affects the relationship between a fixed aperture and a total magnitude measurement. The resulting 10σ depth map is then translated to an averaged weighted pixelized map of resolution given by $N_{\rm side} = 4096$ (an angular scale of 0.74 arcmin^2) using the `HEALPix` software (Górski et al. 2005). These maps will be a key component of our final masks (see below). See also the appendix of Suchyta et al. (2015) for a description of how the mask is constructed.

4 THE LSS BENCH-MARK SAMPLE

From the *SVA1 Gold* catalog described above we select a sample of galaxies that enables robust clustering measurements by applying further magnitude and color cuts and restricting our analysis to the largest contiguous area overlapping the SPT-E field[5],

$$\begin{aligned} 60 < ra\,[\mathrm{deg}] < 95 \\ -60 < dec\,[\mathrm{deg}] < -40 \end{aligned} \tag{13}$$

We then focused on an flux limited sample defined by

$$18 < i < 22.5, \tag{14}$$

where i refers to the `SExtractor`'s `MAG_AUTO` quantity. In Sec. 4.3 we discuss that this sample is complete in regions of the survey deeper than $i = 22.5$ magnitudes. Therefore we will only consider such regions as our baseline footprint, covering an area of $116.2\,\mathrm{deg}^2$. Within this footprint we have $2,333,294$ objects in the bench-mark LSS sample with number density $n_g = 5.6\,\mathrm{arcmin}^{-2}$.

[3] This calibration is improved using the Stellar Locus correction described below.

[4] http://space.mit.edu/~molly/mangle/

[5] Note that we are being very conservative with potential contamination from the LMC by removing one extra stripe at $-61 < dec\,[\mathrm{deg}] < -60$.

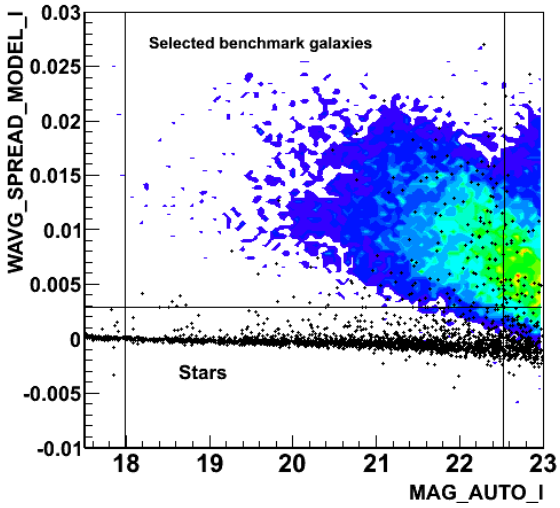

Figure 2. Star-galaxy separator distribution of true stars and galaxies vs. magnitude from a combined spectroscopic (Sánchez et al. 2014) and space imaging sample (Leauthaud et al. 2007). Color contour corresponds to the true galaxy locus whereas true stars are depicted as dots. Magnitude cuts for the sample used in this work are shown as vertical lines, together with the selection of stars and galaxies using the `WAVG_SPREAD_MODEL` classifier in horizontal lines.

We also perform the following additional color cuts

$$\begin{aligned} -1 < g - r < 3 \\ -1 < r - i < 2 \\ -1 < i - z < 2, \end{aligned} \tag{15}$$

in order to remove outliers in color space. While `MAG_AUTO` measures galaxy properties independently for each band, `MAG_DETMODEL` applies a consistent morphological model across all bands, yielding more accurate galaxy colors. We therefore use `MAG_DETMODEL` when making the above color cuts on the sample.

4.1 Star-galaxy separation

The selection of galaxies (as opposed to stars) from the Gold catalog is achieved using a star-galaxy classifier that relies on the `SExtractor SPREAD_MODEL` parameter (Desai et al. 2012; Soumagnac et al. 2013) which measures the light concentration of each object using a linear discriminant analysis at the pixel level. This value is calculated for all single epoch images contributing to the final coadded object and the classifier is built using a weighted average expression of the individual `SPREAD_MODEL` values. The final value is then termed `WAVG_SPREAD_MODEL`. The star-galaxy separation is done selecting objects with,

$$\texttt{WAVG_SPREAD_MODEL} > 0.003 \tag{16}$$

In order to assess the efficacy of this cut, we use DES observations over the COSMOS field, in which we can make use of the measurements from the ACS instrument of the Hubble Space Telescope in that area as a source of 'truth' information for morphological separation of stars and galaxies (Leauthaud et al. 2007), and also the DES observations that overlap with the spectroscopic samples that were used in Sánchez et al. (2014) (which provide truth information based on the spectra). The distribution of `WAVG_SPREAD_MODEL` as a function of i-band magnitude for these objects is shown in Fig. 2. This data allow us to determine a contamination level of 1.5% from stars, with a cut efficiency in the galaxy sample of 96% (for DES objects with $i < 22.5$). The impact of stellar contamination on clustering measurements is discussed in Section 6.3.

Photo-z bin	$N_{\rm BPZ}$	$N_{\rm TPZ}$
$0.2 < z < 0.4$	$684,416$	$441,791$
$0.4 < z < 0.6$	$759,015$	$721,696$
$0.6 < z < 0.8$	$494,469$	$586,510$
$0.8 < z < 1.0$	$270,077$	$361,937$
$1.0 < z < 1.2$	$55,954$	$93,958$

Table 1. The number of galaxies in each of our photo-z bins, when using the template based method BPZ ($N_{\rm BPZ}$) and when using the machine learning method TPZ ($N_{\rm TPZ}$).

4.2 Photometric redshift estimation and redshift binning

A priori, different photo-z codes will lead to different estimates of galaxy redshifts, which then will propagate into theory predictions and eventually into the fitted parameters. There exist two main approaches for photometric redshift estimation: template fitting methods and machine learning ones. Each of these approaches have their advantages and disadvantages, in particular they depend differently on how the spectroscopic set is used, which can be used for training a machine learning approach or to derive fitting priors on the template-based one. Their differences also depend on the distribution of the magnitude errors, the survey depth, the observing conditions and the galaxy population among others.

A direct comparison of clustering observables and derived quantities such as bias between these two approaches is a good way of assessing the impact of photometric redshift estimation on clustering measurements. To our knowledge this test has not been done in clustering analyses in the literature. In this paper, we employ two different algorithms to estimate photometric redshift: BPZ (Bayesian Photometric Redshifts), a well known template-fitting based method (Benítez 2000; Coe et al. 2006) and TPZ (Carrasco Kind & Brunner 2013, 2014), a high-performing machine learning algorithm for DES data.

4.2.1 Template Fitting method

BPZ[6] compares the broad-band galaxy spectral energy distribution to a set of galaxy templates until a best fit is obtained, which determines both the galaxy spectral type and its photometric redshift. The details and capabilities of BPZ on early DES data are presented in Sánchez et al. (2014), where it shows the best performance among template-based codes. The primary set of templates used contains the Coleman, Wu & Weedman (1980) templates, two starburst templates from Kinney et al. (1996) and two younger starburst simple stellar population templates from Bruzual & Charlot (2003), added to BPZ in Coe et al. (2006). We calibrate the Bayesian prior by fitting the empirical function $\Pi(z, t|m_0)$ proposed in Benítez (2000), using a spectroscopic sample matched to DES galaxies and weighted to mimic the photometric properties of the DES-SV sample used in this work.

4.2.2 Machine Learning method

TPZ[7] (Carrasco Kind & Brunner 2013, 2014) is a machine learning algorithm that uses prediction trees and random forest techniques

[6] http://www.stsci.edu/∼dcoe/BPZ/
[7] http://lcdm.astro.illinois.edu/research/TPZ.html

to produce robust photometric redshift probability density functions (PDFs) together with ancillary information for a given galaxy sample. The prediction tree is built by asking a sequence of questions that recursively split the input data taken from the spectroscopic sample, frequently into two branches, until a terminal leaf is created that meets a stopping criterion (e.g., a minimum leaf size or a variance threshold). By perturbing the data using their magnitude errors and by taking bootstrapping samples, many (600 in this case) uncorrelated trees can be created whose results are aggregated to construct each individual photo-z PDF.

Sánchez et al. (2014) contains detailed information about the application of TPZ and BPZ to DES Science Verification data including the derivation of photometric redshifts used in this paper. It also presents a comparison of these two algorithms to numerous other existing photo-z methods. They both proved to be among the best performing codes in the tests presented in that work, motivating their use in further science analyses using DES-SV data. The 68th percentile widths (corresponding to the scatter in the photo-z solution) were found to be 0.078 for TPZ and 0.097 for BPZ, with 3σ outlier fractions being 2 per cent for both algorithms, for detailed metrics see Sánchez et al. (2014).

4.2.3 Redshift binning

Because we wish to study the evolution of galaxy clustering with redshift, we split our sample in five photo-z bins of width $\Delta z = 0.2$, from $z_{phot} = 0.2$ up to $z_{phot} = 1.2$. For this analysis we use two photometric redshift estimation algorithms which provide photo-z probability density functions (PDF). One is a standard template-fitting based code, BPZ, which we take as our default photo-z algorithm. For robustness and cross-validation we also use a machine learning based algorithm, TPZ. Both are discussed in more detail in Sec. 4.2. The number of galaxies in each photo-z bin is given in Table 1.

In Fig. 3 we show the estimated "true" redshift distribution corresponding to each of these five bins for BPZ and TPZ; we will use these in Eq. 2 to make model predictions for the clustering analysis. These were computed by stacking the individual photo-z PDFs of each galaxy in the binned sample. The accuracy of this approach in DES-SV data is discussed in Sánchez et al. (2014) and its performance in simulated data in a companion paper, Leistedt et al. (2015).

As can be seen by inspecting Table 1, there are some large differences in the numbers of galaxies within the same z_{phot} bin but using different photo-z algorithms. These differences can be understood when also considering the reconstructed $n(z)$ of each sample. For instance, in the $0.2 < z_{phot} < 0.4$ bin, the distribution is considerably wider for BPZ. Thus, the bin has more objects in it, as objects that are truly at higher redshift (and occupy a larger volume) are mis-estimated to occupy the low redshift bin. Therefore, despite their differences, we expect each set of galaxies to effectively probe the mean clustering properties of the galaxies at the effective redshift of the sample, when properly analyzed within the context of the dn/dz.

4.3 Survey depth and baseline angular mask

Within any given redshift range, minimization of the expected uncertainty on clustering measurements is a trade-off between area and number of objects (see, e.g., Crocce, Cabré & Gaztañaga 2011). A fainter flux-limited sample maximizes the number density of observed galaxies, but one should discard areas where this limit is not met in order to secure the completeness of the sample. The available SV area decreases by approximately 20 deg^2 for limiting magnitudes in the range $20 <$ `MAG_AUTO` < 22.5 but exhibits a sharp decrease of deeper

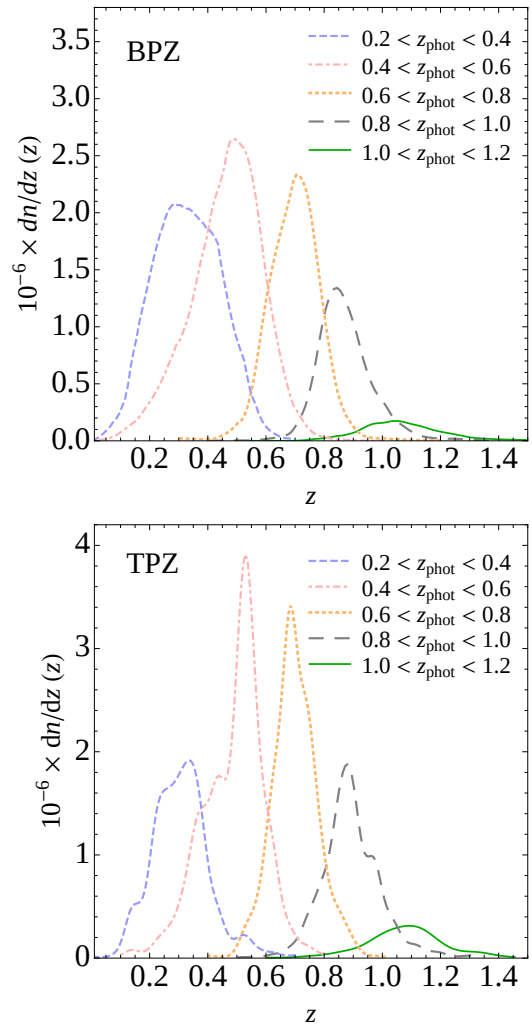

Figure 3. Redshift distributions obtained from stacking the photo-z PDF of objects selected in five top-hat photo-z bins using either the template based estimation (BPZ, top panel) or the machine learning one (TPZ, bottom panel).

regions. For example, from the approximately 150 deg^2 SV area overlapping SPT-E region, only about 80 deg^2 have a limiting magnitude of $i = 23$.

Thus, we set the flux-limit of $i = 22.5$ for the definition of the sample and consider only the regions with i-band limiting magnitude `MAG_AUTO` > 22.5 in our clustering analysis [8]. The resulting angular footprint, after combining with the survey mask of Sec. 3.3.1, is shown in Fig. 4. It occupies a contiguous area of 116.2 deg^2. We consider this footprint as our baseline mask which will be used for all subsequent clustering measurements in this work. But we note that further masking may be due in some photo-z bins to mitigate systematic effects, this will be the subject of Sec. 6.

We also note that our area and sample are very close to the ones used in the CHFTLS analysis of Coupon et al. (2012); with an area of 133 deg^2 to a limiting depth of $i = 22.5$. However our footprint is contiguous, which makes the impact of the integral constraint on large scales negligible as opposed to measurements in the four separated fields of CFHTLS. The sample we analyze is the largest contiguous area demonstrated to have a reliable i-band depth of 22.5 for extra-

[8] We have checked that the mean number density of objects in the sample does not change any longer if we restrict to deeper areas.

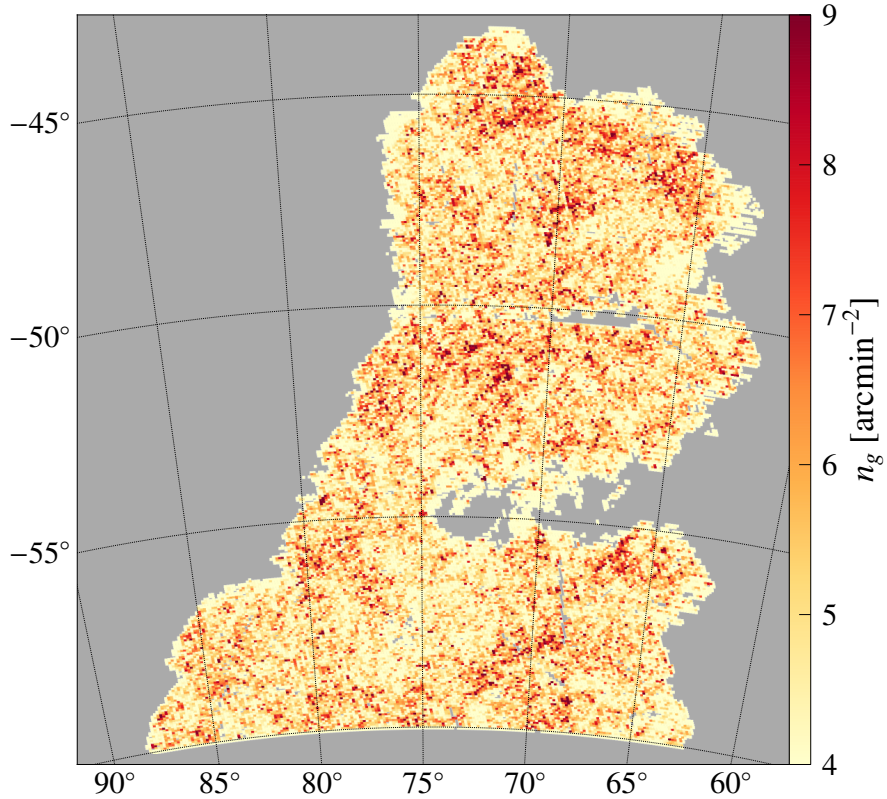

Figure 4. The distribution of LSS bench-mark sample galaxies over the angular footprint defined by regions with survey limiting magnitude in the i-band > 22.5. The sample is selected to be flux limited to $i \leqslant 22.5$ and has a mean density of 5.6 arcmin^{-2}. All the regions considered provide at least S/N 10 measurements for objects at i-band $= 22.5$. This choice balances concerns between using the maximum depth and area possible, as described in the text. The x-axis (y-axis) corresponds to right ascension (declination) measured in degrees.

galactic sources, despite being less than 3% of the size of the final DES footprint.

5 POTENTIAL SOURCES OF SYSTEMATICS AND MAP PROJECTIONS

Quantities that may modulate the detection efficiency of galaxies and produce spurious galaxy correlations have been recorded and mapped so that any systematic effects can be empirically studied. The maps used in this paper are presented below.

Galaxy catalogs can be affected by the time-dependent fluctuations in the observing conditions of the survey. These fluctuations affect the depth in non-trivial ways because they occur in the single-epoch images, which are coadded and post-processed before extracting the galaxy properties. In other words, the transfer function between the raw images and final catalogs is a complicated function of the input single-epoch images and the coadding and source-extraction pipelines, also coupled to the galaxy density field and astrophysical foregrounds.

A significant effort is dedicated in DES to understanding the transfer function from intrinsic to observed quantities (see, e.g., the BCC-UFig framework, Chang et al. 2014 and also the `Balrog` framework, Suchyta et al. 2015), which will be critical to precisely identify sources of systematics and eliminate them in the cosmological analyses. Such techniques can in principle be used to account for the observational effects that cause fluctuations in the observed galaxy density. These methods are still maturing (though see Suchyta et al. 2015 for a working example of applying such techniques to LSS measurements), however, and thus for the DES-SV data we instead test the observed galaxy density against a large number of maps of potential sources of systematics (e.g., the mean seeing). These maps are then used to run null-tests and correct for systematic shifts in the clustering measurements, as detailed in subsequent sections. This requires one to project the properties of the single-epoch images onto the sky, accounting for the geometrical overlap and also weighting due to the coadding procedure. This projection is fully detailed in Leistedt et al. (2015), where the full set of maps, their potential applications, and a pedagogical example using catalogs extracted from the BCC-UFig simulated images are presented.

For this paper, we consider single-epoch properties that are known to affect the depth and therefore might affect the clustering measurements. Each of these observational properties is mapped on the sky into high-resolution (Nside = 4096) `HEALPix` format maps, and we reduce the full set of values in each pixel into a weighted mean using inverse-variance weights (as in the coadd and `mangle` pipelines). The four quantities used in this work, computed in the *grizY* bands, are (see Leistedt et al. 2015 for a full description):

- `FWHM`: the mean seeing (in pixel units), measured as the full width at half maximum of the flux profile.
- `airmass`: the mean airmass.
- `skybrite`: the mean sky brightness.
- `skysigma`: the mean sky background noise derived as flux variance per amplifier in each CCD chip[9].

[9] In principle, the sky brightness and the mean sky background should be

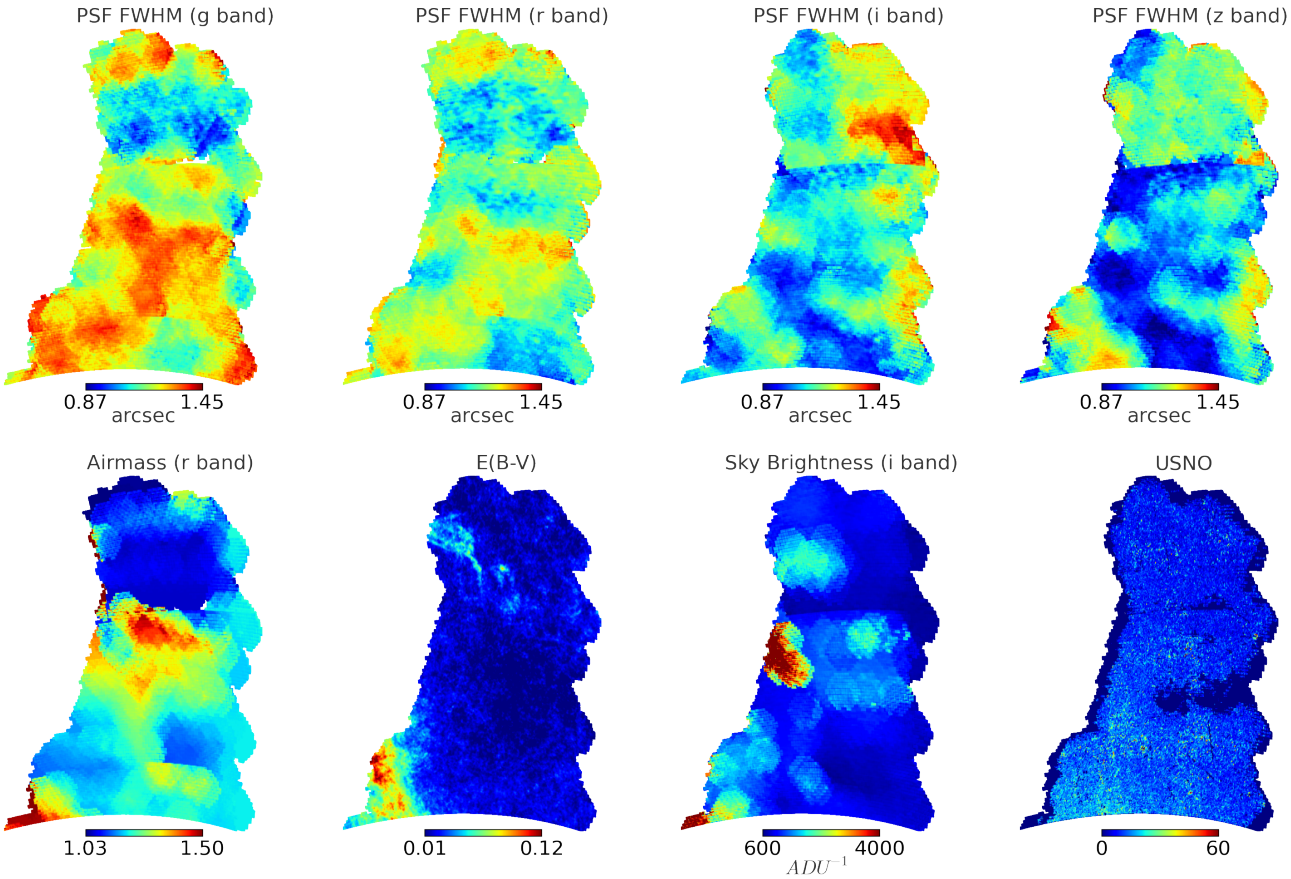

Figure 5. Maps of potential sources of systematic errors that can lead to spatial variations in the number of observed galaxies (through depth fluctuations or sample contamination) or degrade the data quality itself (impacting the determination of magnitudes, colors or photo-z). The values at each pixel are computed as single statistical estimators of the single-epoch values for the images contributing to that map pixel.

We will consider only the mean quantities; in the future it might be important to test against maps of the variance of these properties as well. In addition to the above maps, we consider the following:

- `Galactic dust extinction`: As described above, the data are calibrated using the technique of stellar locus regression, which in addition to correcting for instrumental and atmospheric effects also removes reddening due to Galactic extinction. To test for residual effects of Galactic dust on the photometry, we include the Schlegel, Finkbeiner & Davis (1998) dust map as a potential systematic, pixelized with `HEALPix` resolution Nside=4096.

- `USNO`: Contamination of our galaxy sample by stars will affect the measured clustering; stellar density increases towards the galactic plane, following a gradient that causes significant clustering signal on large scales (Crocce et al. 2011; Ross et al. 2011a; Ho et al. 2012). Furthermore, because the footprint used in this work neighbors the Large Magellanic Cloud, stellar contamination may introduce more complicated spurious clustering. To investigate this effect, we include a map of stellar density across the field, as measured by the USNO-B1 catalog (Monet et al. 2003). We take USNO stars with B magnitude brighter than 20; deeper than this limit the depth of the USNO catalog varies across the field. Although this catalog is brighter than our sample, we expect the stars in our sample to trace the same galactic density distribution as those in the USNO catalog. Due to low statistics we use a coarser map `HEALPix` resolution of Nside=256.[10]

- `depth_mask`: As described in section 4.3, we cut our sample at a magnitude of `MAG_AUTO` < 22.5 and use a footprint where the data provides $> 10\sigma$ measurements at that depth. We include the depth map as a potential systematic in order to probe possible inaccuracy of the map and incompleteness at the faint end of the sample.

- `chi2_psf_fit`: The coadded image on which we base our photometry can have complicated structure, as the number of input images varies across the field and each input image has a unique PSF. It is therefore difficult for the photometric reduction to characterize the image and its PSF across an entire coadd unit. Poor characterization of the coadd image quality at the location of a star will cause the star to be poorly fit by the PSF model. In locations of a poor PSF fit to stars, galaxy photometry will also be adversely affected by the inaccuracy of the estimate of the object's size. We therefore use a map of the average χ^2 of objects' fits to the PSF model to test against this potential source of systematic uncertainty. We only consider very bright objects, with $16 < i < 18$, to ensure that the large majority of them are stars.

strongly correlated. We test against both as a validation that this process has worked properly.

[10] We also tested results against a sample of fainter DES detected stars and found no significant differences.

- `chi2_detmodel_fit`: `MAG_DETMODEL` magnitudes are measured by fitting a Sérsic profile, convolved with the PSF model, to the object image. Poor `MAG_DETMODEL` fits over a region of data would indicate photometric systematic errors in the same manner as poor PSF fits, as described above. We therefore also test our data against a map of the average χ^2 of the `detmodel` magnitude fits; we expect this map to be somewhat degenerate with the `chi2_psf_fit` map.

We show some of the most relevant maps in Fig. 5.

6 MITIGATING SYSTEMATIC EFFECTS IN ANGULAR CLUSTERING

In this section we give a detailed explanation of our general procedure to address the impact of potential systematic effects due to varying observing conditions, stellar contamination, or any other source of spurious density fluctuations across our footprint. We assume that all potential sources of systematics are encoded in the maps described in Section 5.

Let us first describe our step-by-step procedure in broad terms with more details and results given in subsequent sections,

[1] *Galaxy Density vs. Potential Systematics* : Our first step is to study the galaxy number density in each tomographic bin as a function of the value of each potential systematic variable. If the galaxy density is independent of the value of the given potential systematic we do not consider it as impacting our data. Otherwise we try to either mask it or correct for its impact in the clustering measurement, as shown below.

[2] *Bad Regions Masking* : The relation between galaxy density vs. potential systematics can take several forms. If this relation is such that the galaxy density is constant as function of the potential systematic and changes sharply after some threshold value, we mask the regions of such anomalous dependence (such as bad seeing, or high airmass) to minimize its effect. This defines a veto mask that we then use in all subsequent clustering analysis. The quality of the data is such that we typically mask only small fractions of the footprint. We will only mask data worse than a given quantity, as when the data quality is worse, we naturally expect photo-zs, star/galaxy separation, object detection, etc., to all perform worse and thus be more likely to cause spurious fluctuations in the observed galaxy density.

[3] *Clustering corrections using Cross-correlations* : If the relation between galaxy density vs. potential systematic is smooth and close to linear (after imposing the veto mask discussed above) we use cross-correlation between galaxies maps and the observational maps to correct our measurements.

[4] *Stellar Contamination* : As described in Section 4.1, we estimate 1.5% of the SV galaxy sample is comprised of mis-classified stars. The clustering of stars over the SV footprint is close to zero, implying that the main effect of the contamination is to proportionally remove power from the $w(\theta)$ measurements. We estimate the stellar contamination in each tomographic bin and derive how to correct our clustering measurements given such contamination in Section 6.3.

6.1 Galaxy Density vs. Potential Systematics

Our first step is to study how the galaxy density depends on each potential systematic variable, if at all. The starting point for these measurements is pixelized maps of the distribution of galaxies and of the potential systematics. We divide the range of the values of each potential systematic variable into roughly ten bins. For the sky area corresponding to each bin, we calculate the mean galaxy number density normalized by the mean across the whole footprint. In Fig. 6 we show the resulting normalized galaxy density as a function of the different potential systematic variables. After examining all z bins, we only show the relevant cases. Error bars are computed using the jackknife method by splitting the footprint into 100 regions. Hence errors in Fig. 6 might be under-estimated due to a limited sampling of large-scale fluctuations. Furthermore, note that extreme values of the potential systematics are typically not sampled in most JK regions. In such cases (i.e. the first and/or last points in each panel of Fig. 6) we choose to estimate error bars as Poisson sampling noise.

We perform this procedure in every redshift bin, and define an angular mask due to systematics per bin. These are:

- $0.2 < z < 0.4$: The top left panel of Fig. 6 shows the relation of galaxies selected in this bin with BPZ as a function of FWHM in g-band and z-band. For high values of seeing we see a clear drop of galaxy density and therefore choose to remove these areas. We cut regions with g-band FWHM > 1.34 arcsecs (1.09 of the mean g-band FWHM) which removes 13% of the area. Similarly, we remove area with z-band FWHM > 1.24 arcsecs (1.12 of the mean z-band FWHM) which removes a further 6%. In all 19% of our nominal footprint is masked for the BPZ sample at this photo-z bin. In the top left and right panels of Fig. 7 we show histograms of FWHM shading the portion removed by this selection (at the same time, Fig. 5 shows the maps for FWHM in all bands, providing an idea of which regions are removed). The same procedure but using galaxies selected with TPZ yields the top right panel of Fig. 6. In this case we remove i-band FWHM > 1.28 arcsecs (1.11 of the mean i-band FWHM). Note that there is also a dependence with z-band FWHM. However after the regions with worst i-band FWHM are removed, this dependence disappears. This is shown by the triangular symbols with dashed error-bars . Hence for TPZ we only mask bad i-band FWHM regions removing $\sim 11\%$ of the original footprint. The distribution of i-band FWHM is shown in the bottom right panel of Fig. 7, where the masked out part has been shaded.

- $0.4 < z < 0.6$: For this bin only the g-band FWHM affects the galaxy density in a way suitable for masking. It does so in a very similar manner for both photo-z samples, as shown in the bottom left panel of Fig. 6. Hence we define a common veto mask for BPZ and TPZ removing regions with g-band FWHM > 1.34 arcsecs (1.09 of the mean g-band FWHM). This removes 13% of the area. According to Fig. 7 the threshold value could be smaller because the relation turns around at ~ 1.28 arcsecs. We have checked that this extra masking does not change the final clustering results, but it implies removing about 30% of the area. Hence we cut g-band FWHM > 1.34 arcsecs what should account for the bulk of the effect. Note that after removing this region there is a clear dependence of galaxy density increasing with g-band FWHM. This residual is exactly the type that can be addressed by means of cross-correlations between galaxies and potential systematics, as detailed in Sec. 6.2.

- $1.0 < z < 1.2$: For the highest redshift bin we again define one additional mask for BPZ and one for TPZ based on Fig. 6 bottom right panel. For BPZ we remove regions with r-band FWHM > 1.28 arcsecs (1.08 of the mean $r-$band FWHM) which amounts to 6% of the original footprint (leaving 109.2 deg^2). For TPZ we cut regions with i-band FWHM > 1.28 arcsec (1.1 of the mean i-band FWHM and 10% of the area). Above these threshold values the density of galaxies drops sharply compared to the mean.

For the bins at intermediate redshifts ($0.6 < z < 0.8$ and

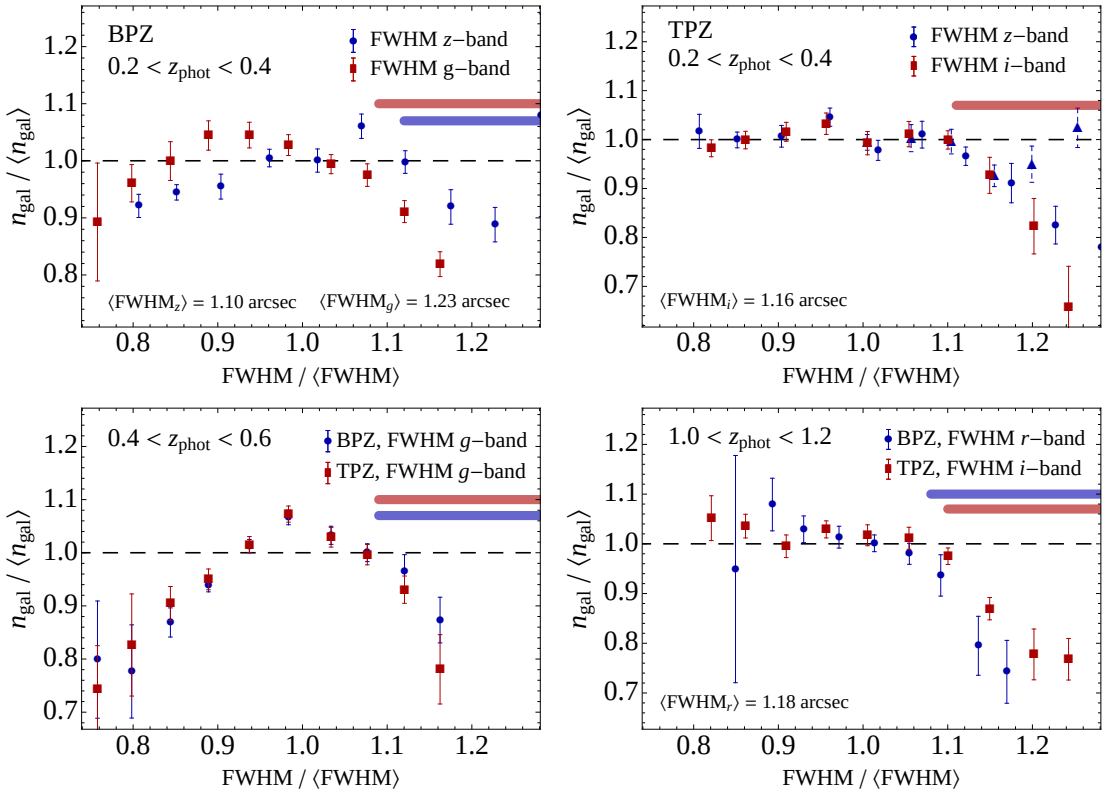

Figure 6. Galaxy density (normalized to the mean over the footprint) as a function of the value of different potential sources of systematic errors, in regions where those galaxies reside. This is displayed for our two photo-z estimators and different tomographic bins. We only show those cases where the density of galaxies drops steadily after some threshold value indicating regions that may lead to systematic effects. To minimize them we mask these regions on top of our nominal mask (as a function of photo-z bin). Values masked out in each case are indicated by the inset boxes in blue/red. Cases where this relation is roughly linear (e.g. $0.4 < z < 0.6$ after high g-band FWHM values are masked) induce a change to the clustering measurements that can be corrected for using the cross-correlations between the maps for galaxy distribution and the ones for systematic effects (see Sec. 6.2)

.

$0.8 < z < 1.0$) we find no dependence of the data with the potential systematics. Thus, we do not correct the clustering or introduce additional cuts to the nominal footprint when analyzing these bins.

6.2 Galaxy-Systematics cross correlations

Once regions with bad observing conditions that translate into an anomalous drop of the density of galaxies have been identified and masked out we devote our attention to correcting more subtle dependencies where the density varies roughly proportionally with the strength of the systematic. For this we follow and extend the work of Ross et al. (2011a); Ho et al. (2012). We start by assuming that fluctuations in the potential sources of systematic effects $\delta^i_{\rm sys}$ will induce spurious fluctuations into the observed galaxy number density $\delta_{\rm obs} = n_g/\bar{n}_g$, such that

$$\delta_{\rm obs} = \delta_{\rm true} + \sum_i \alpha_i \, \delta^i_{\rm sys} \qquad (17)$$

where $\delta_{\rm true}$ are the true underlying fluctuations we want to recover and i refers to any of the maps described in Section 5. By definition we assume that true galaxy fluctuations do not correlate with the ones

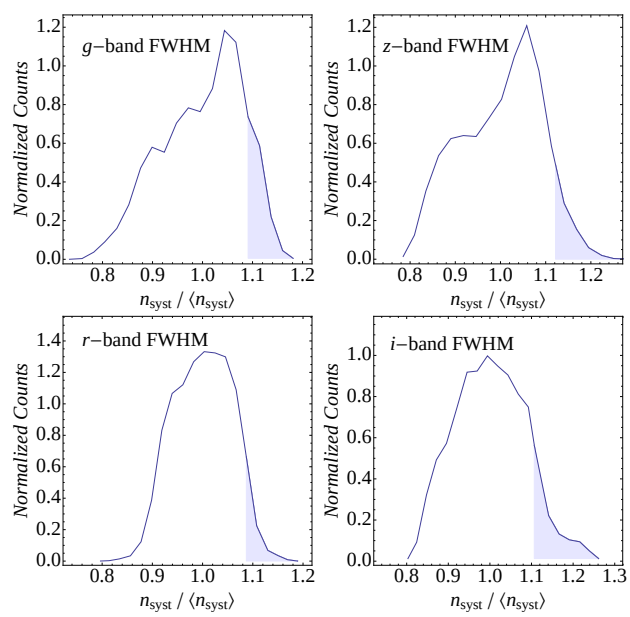

Figure 7. Distribution of FWHM weighted co-added values in different bands across our footprint. The shaded parts of the histograms correspond to regions where large decrement of galaxy density in different photo-z bins is found. To avoid a systematic effect they were removed from the corresponding clustering analysis. See Sec. 6.1 for further details.

induced by the systematics[11],

$$\langle \delta_{\rm obs}\, \delta^j_{\rm sys} \rangle = \sum_i \alpha_i \, \langle \delta^i_{\rm sys} \delta^j_{\rm sys} \rangle. \tag{18}$$

If we define,

$$w_{{\rm cross},j} \equiv \langle \delta_{\rm obs}\, \delta^j_{\rm sys} \rangle,$$
$$w_{{\rm auto},ij} \equiv \langle \delta^i_{\rm sys}\, \delta^j_{\rm sys} \rangle, \tag{19}$$

then Eq. (18) can be written as,

$$w_{{\rm cross},j} = \sum_i \alpha_i \, w_{{\rm auto},ij}, \tag{20}$$

and the solution is a matrix inversion,

$$\vec{\alpha} = \vec{w}_{\rm cross} \cdot (w_{\rm auto})^{-1}. \tag{21}$$

Hence if we have 6 systematic maps we have to invert a 6x6 matrix for each θ-bin (notice that we are a priori assuming that α is spatially independent to factor it out of the correlations). From Eq. (17) one can work out the auto correlation of true fluctuations

$$w(\theta)_{\rm true} = w(\theta)_{\rm obs} - \sum_i \sum_j \alpha_i \alpha_j \langle \delta^i_{\rm sys} \delta^j_{\rm sys} \rangle \tag{22}$$
$$= w(\theta)_{\rm obs} - \vec{\alpha} \cdot \vec{w}_{\rm cross}, \tag{23}$$

which reduces to the standard case,

$$w(\theta)_{\rm true} = w(\theta)_{\rm obs} - w^2_{\rm cross}/w_{\rm auto} \tag{24}$$

for just one systematic.

Note that our starting assumption in Eq. (17) of a roughly linear relation between $\delta_{\rm obs}$ and $\delta_{\rm sys}$ can be tested and confirmed with the one-dimensional relation of galaxy density as a function of potential systematics as done in Sec. 6.1. If strong non-linear dependencies are found then Eq. (17) will not apply and the corrections induced will bias the measurements. For this reason, it is important to mask the regions (e.g., high g-band seeing) where the relationships are most clearly non-linear, as is done in the previous section.

In addition, note that our approach does account for the correlation among the potential systematics themselves. Nevertheless, solving for an arbitrarily large number of maps might still induce over-corrections and biases due to the inversion in Eq. (23) and the noise of the measurements particularly on large-scales.

Further below we explain our quantitative criteria based on cross-correlations to choose which maps impact the galaxy sample and need to be corrected for. Then we follow a combined approach in which we also cross check against the density relations as in Sec. 6.1 for those systematics that pass the criteria that a linear relation exists.

To use a quantitative criterion to select the most relevant systematics maps, we cross-correlate them with the galaxy distribution maps at each redshift bin after applying the masking described in Section 6.1. For each potential systematic we calculate the correction to the galaxy correlation according to Eq. (24), $w^2_{\rm cross}/w_{\rm auto}$, ignoring cross-correlations among systematics to begin with. This correction is determined for each angular bin θ. Using JK resampling over the footprint, we calculate an error on the correction. Then, if the correction is inconsistent with zero at a 1σ level[12], the systematic is deemed significant and taken into account in further analysis; otherwise, it is neglected.

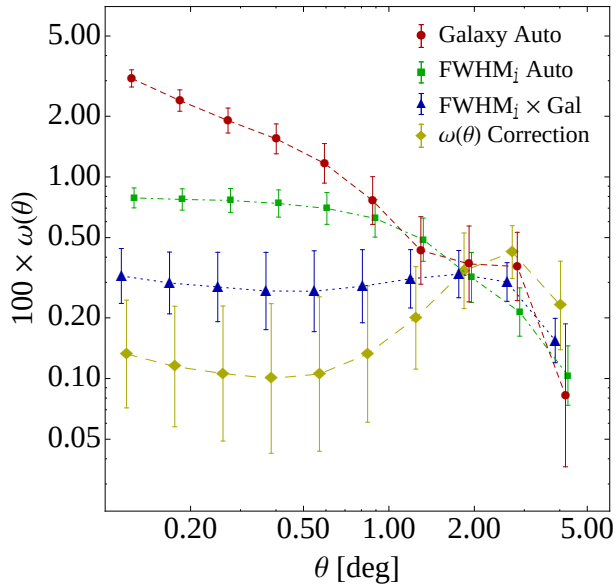

Figure 8. Different two-point correlations entering the clustering correction in the bin $0.2 < z < 0.4$ for the BPZ sample due to spatial variations in i-band FWHM over the DES footprint. See Sec. 6.2 for details.

Figure 8 shows as an example the correlation between the central redshift bin, $0.2 < z < 0.4$ for the BPZ sample and the FWHM map in the i band. The figure shows the auto-correlation of the galaxy sample, the auto-correlation of the FWHM map, the cross-correlation between the data and the systematic, and the corresponding correction that should be applied to the data (ignoring covariance with any other systematic effects), see Eq. (24). The correction is significantly non-zero, and therefore we consider the i-band FWHM map to be a relevant systematic in our analysis.

For the BPZ photo-z catalog the systematics that correlate significantly with the data set are:

- $0.2 < z < 0.4$: i-band r-band and z-band FWHM, r-band Skybrite and dust extinction (5 maps)
- $0.4 < z < 0.6$: g-band and r-band FWHM (2 maps)
- $0.6 < z < 0.8$: None
- $0.8 < z < 1.0$: None
- $1.0 < z < 1.2$: None after r-band FWHM masking

For the TPZ photo-z catalog the systematics that correlate significantly with the data set are:

- $0.2 < z < 0.4$: r-band FWHM, r-band and z-band Skybrite (3 maps)
- $0.4 < z < 0.6$: r-band FWHM (1 map)
- $0.6 < z < 0.8$: None
- $0.8 < z < 1.0$: None
- $1.0 < z < 1.2$: None after i-band FWHM masking

Despite the fact that the limiting depth is the same in all of the redshift bins we use, we find slight variations in the type and degree of systematic contamination as a function of redshift. The corrections arise mainly at low redshifts, hence they might be due to a correlation between the observing conditions and the determination of photometric redshifts (the DES filter systems does not contain u-band, which degrades the low redshift photo-z). Nonetheless, the significance of the corrections is never beyond 2σ at any given angular scale which signals that the data is not very impacted by systematics and our re-

[11] This means we must be careful in choosing systematics: e.g. photometric redshift error is typically smallest for luminous red galaxies, which cluster strongly; we therefore must not use a map of redshift error (or, similarly, magnitude error which correlates with galaxy luminosity) as a systematic. Otherwise the correlation can be strong, see for example Martí et al. (2014).

[12] We use 10 logarithmically distributed angular bins in $0.12 - 4$ deg. The 1σ limit corresponds to $\chi^2 > 11.53$, where the χ^2 is computed using the covariance of the $w^2_{\rm cross}/w_{\rm auto}$ estimates in the JK regions.

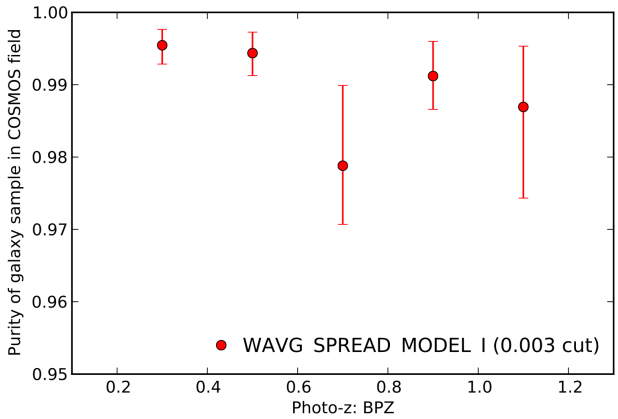

Figure 9. The purity of the DES galaxy sample in the COSMOS field in each of the BPZ tomographic redshift bins we use, determined using the default cut on WAVG_SPREAD_MODEL > 0.003. The stellar contamination is at most 2%.

sults not dominated by these corrections. This is discussed in more detail in Sec. 7 and Tables 2 and 3.

6.3 Stellar Contamination

Stellar contamination will affect the measured clustering signal, even if the cross-correlation between the galaxies and stars is negligible. This is due to the fact that rather than modulate the selection function of galaxies, stellar contamination introduces a separate population. In the limit where the stars are un-clustered and the stellar contamination is constant the observed galaxy density is, $\delta_o = (1 - f_{star})\delta_{gal}$ and thus

$$w_{\rm gal} = \frac{w_{\rm o}}{(1 - f_{\rm star})^2}, \tag{25}$$

i.e., the measured clustering is diluted by the square of the purity of the galaxy sample.

We derive a general expression for the clustering of a galaxy population, given some stellar contamination (full details are shown in Appendix D)

$$w_{gal} = (1 + f_{\rm star})^2 \left(w_{\rm o} - f_{\rm star}^2 w_{{\rm star},S} - \frac{f_{\rm star}^4}{(1 + f_{\rm star})^2} \right), \tag{26}$$

where in this formalism, the number density of contaminating stars is allowed to be a function of not just the total density of stars, but also the survey observing conditions.

The cross-correlation tests presented in previous sections did not find stellar contamination to be a significant systematic contaminant. This suggests that the term in Eq. 26 proportional to $w_{{\rm star},S}$ is consistent with zero. However, the terms $(1 + f_{\rm star})^2 w_{\rm o} - f_{\rm star}^4$ remain, and thus any stellar contamination must be accounted for.

We estimate the stellar contamination in each photometric redshift bin by using DES observations in the COSMOS field. This region contains space-based observations which identify matching DES sources as point-like or not (as detailed in Section 4.1). Figure 9 shows the purity of the DES sample in the COSMOS field and its uncertainty, that we estimate in each redshift bin for the classifier discussed in Sec. 4.1. The uncertainty includes the statistical Poisson variance, the error due to sample variance fluctuations of the galaxy sample, which impacts the purity too, and the fluctuations of the stellar sample in the SPTE area. These are large due to the presence of the Large Magellanic Cloud, and are included when extrapolating to the SPTE field. This extrapolation to the DES SPTE area is done by finding the density of highly probable stars in the COSMOS field with a $19 < i < 21$ cut and $|$WAVG_SPREAD_MODEL$)| < 0.001$ and we find that the average density is the same as that found in the DES SPTE area when applying the same cuts, to within 10% (below our error bars due to intrinsic fluctuations of the galaxy and stellar samples). We apply the correction given by Eq. (26) to the $w(\theta)$ measurements and propagate this uncertainty into measurements of the galaxy bias. The correction to the bias is largest in the $0.6 < z < 0.8$ bin, but is at most 2%.

7 RESULTS

We now discuss different aspects of our angular clustering measurements in the tomographic bins, in particular the corrections for systematics discussed above. Let us first note that in what follows we define the best-fit bias b and 1σ error σ_b in the standard way, i.e., given

$$\chi^2(b) = \sum_{\theta,\theta'} (\hat{w}(\theta) - b^2\, w_T(\theta)) {\rm Cov}^{-1}(\theta,\theta')(\hat{w}(\theta') - b^2\, w_T(\theta')) \tag{27}$$

we first find its minimum to define b and then vary it around this value such that $\Delta\chi^2 = 1$ with $\Delta\chi^2 \equiv \chi^2(b + \sigma_b) - \chi^2(b)$. The sum runs through a range of scales that will be specified accordingly while the covariance has been defined in Sec. 2.3 as the "mixed" approach, see Eq. (11). In Appendix B we discuss the best-fit bias results when different ways of estimating the covariance are considered. Note that we do not perform joint fits including multiple redshift bins, hence Eq. (27) does not include the covariance between different redshift bins. The $b^2\, w_T$ refers to the modeling of Eq. (2), where w_T is the underlying dark-matter clustering multiplied by a linear bias b to yield the model galaxy-galaxy correlation function. If w_T is computed starting from the linear theory power spectrum, we call this model "linear growth". If instead this is computed with the nonlinear prescription Halofit (Takahashi et al. 2012), but still maintaining a linear bias, we refer to it as the "linear bias" model. On large scales Halofit reduces to linear theory, while on small scales the effects of non-linear growth increase the model correlation above its linear value. The smallest scale at which w_T from Halofit equals w_T from linear theory will be referred to as "scale of linear growth". These are reported in the caption of Table 2 for each of the five photo-z bins.

Tables 2 and 3 show the result of fitting a "linear growth" model to the clustering measurements before any systematic correction is applied (second column, labeled "baseline mask"), after a subsequent masking of regions with the worst observing conditions, as discussed in Sec. 6.1 (third column, labeled "+ bad regions masking") and lastly, after the remaining correlation with relevant maps is corrected for as detailed in Sec. 6.2 (fourth column, "Gal-Syst cross correlations"). Table 2 corresponds to the BPZ photo-z catalog, while Table 3 to the TPZ one.

The corrections for systematic effects lower the clustering amplitudes and hence the derived best-fit biases. They are mainly important for low redshift, in particular the $0.2 < z < 0.4$ bin. Here the BPZ best-fit bias goes down by $\delta b/\sigma_b \sim 3$, while for TPZ the change is $\delta b/\sigma_b \sim 1.6$. For the $0.4 < z < 0.6$ bin the change in bias is $\mathcal{O}(\sigma_b)$. These two redshift ranges are important for the combination of galaxy-galaxy clustering with weak lensing observables, which typically use lens samples at low-z. Hence these corrections are likely to be necessary to produce an unbiased cosmological analysis using such a combination of probes, e.g., a spuriously large clustering amplitude could translate to a spuriously high measurement of σ_8. At higher redshifts, which we expect to be more suitable for future BAO studies due to larger volumes and better photo-z performance, we do not find significant observational systematic biases affecting our mea-

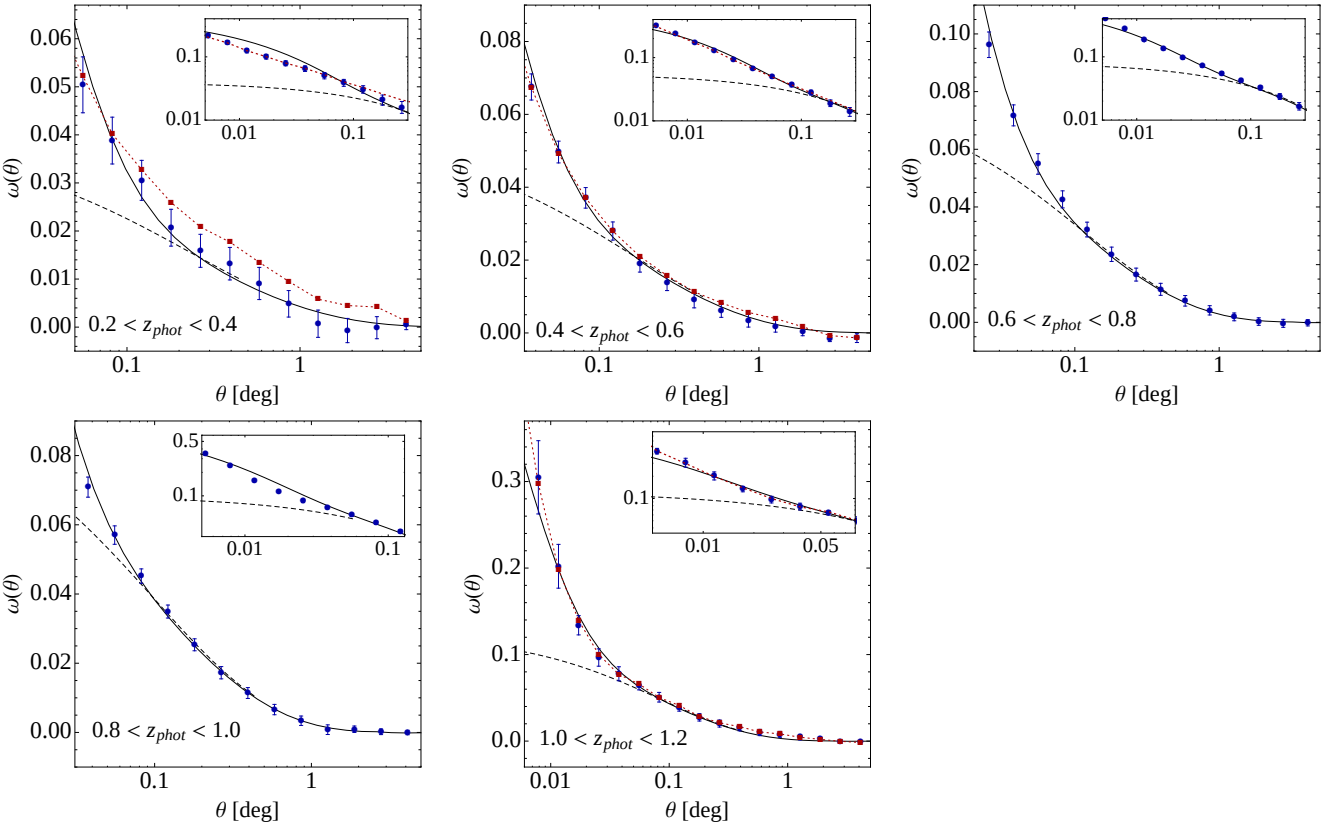

Figure 10. Angular auto-correlation functions, $w(\theta)$, in photo-z bins of the BPZ estimator for our flux limited sample ($i < 22.5$) spanning a broad range in redshift, from $z \sim 0.2$ to 1.2. Blue circles correspond to the measurements after correcting our sample for systematics as discussed in Sec. 6, while red squares is before such corrections (only shown when relevant). Solid lines correspond to a linear bias model applied to the non-linear matter $w(\theta)$ computed with `Halofit`. The best-fit bias displayed was obtained from fitting the range of scales shown in each case (main panels). Dashed lines correspond to linear theory, with the same value of the bias. Note how the simple "linear bias" model describes the clustering towards scales considerably smaller than the linear regime shown by dashed lines. The inset panels show the performance of the best-fit "linear bias" model towards smaller scales than the ones used in the fit (main panels), see text for a detailed discussion.

BPZ (template method): best-fit bias and 1σ error				
Photo-z Bin	Baseline Mask	+ Bad Area Masking	+ Gal-Syst Cross-Corr	χ^2/ d.o.f.
$0.2 < z < 0.4$	1.28 ± 0.07	1.20 ± 0.07	1.05 ± 0.07	$2.5/7$
$0.4 < z < 0.6$	1.25 ± 0.05	1.26 ± 0.05	1.23 ± 0.05	$8.3/8$
$0.6 < z < 0.8$	1.35 ± 0.04	–	–	$2.3/9$
$0.8 < z < 1.0$	1.54 ± 0.02	–	–	$10.3/10$
$1.0 < z < 1.2$	2.20 ± 0.07	2.17 ± 0.09	–	$3.6/10$

Table 2. Impact of the different corrections for observational systematic effects on the derived best-fit bias b for the tomographic bins selected with our template method. The baseline mask is described in Sec. 4.3 and corresponds to all regions where our sample is complete (i.e. 10σ depth $i >= 22.5$). The third column corresponds to an additional masking of regions with high values of potential systematic variables such as seeing, where we observe large decrements in galaxy density (as described in Sec. 6.1). The fourth column refers to further corrections to $w(\theta)$ after these masks, in cases where the data still correlates with maps for potential systematics (as discussed in Sec. 6.2). The fifth column reports the $\chi^2/\mathrm{d.o.f}$ after all corrections applied. Empty entries refer to cases where such corrections were not necessary. These values were obtained after fitting the "linear growth" model for scales $\theta > \theta_{min} = (0.26, 0.18, 0.12, 0.08, 0.06)$ deg., from first to last z -bin.

TPZ (machine learning method)				
Photo-z Bin	Baseline Mask	+ Bad Area Masking	+ Gal-Syst Cross-Corr	χ^2/ d.o.f.
$0.2 < z < 0.4$	1.18 ± 0.07	1.13 ± 0.08	1.07 ± 0.08	$2.1/7$
$0.4 < z < 0.6$	1.29 ± 0.04	1.30 ± 0.04	1.24 ± 0.04	$6.7/8$
$0.6 < z < 0.8$	1.34 ± 0.05	–	–	$14.5/9$
$0.8 < z < 1.0$	1.56 ± 0.03	–	–	$3.7/10$
$1.0 < z < 1.2$	1.97 ± 0.09	1.96 ± 0.06	–	$4.5/10$

Table 3. Same as Table 2 but for the tomographic bins selected with the machine learning method. Different algorithms for photometric redshift estimation use different data quantities (most notably, template based ones use magnitude errors while most machine learning do not). Thus one expects a different response to potential systematics.

surements, at least for the range of angular scales probed in with our current data.

Figure 10 shows the angular correlation function measurements in the five consecutive photo-z bins in which we have split our sample (for the BPZ sample). Blue circles show our measurements after applying all of the corrections for observational systematic effects that we have detailed, while the square red symbols show the measurements before any correction. As is clear from Fig. 10 these corrections do not affect the small scale clustering, and are most important at low redshifts. The dashed curve in each panel corresponds to a linear theory model ("linear growth"), while the solid curve uses the non-linear `Halofit` prescription ("linear bias"). Each model curve is for the bias given in Table 2, fit over the range of angular scales indicated in

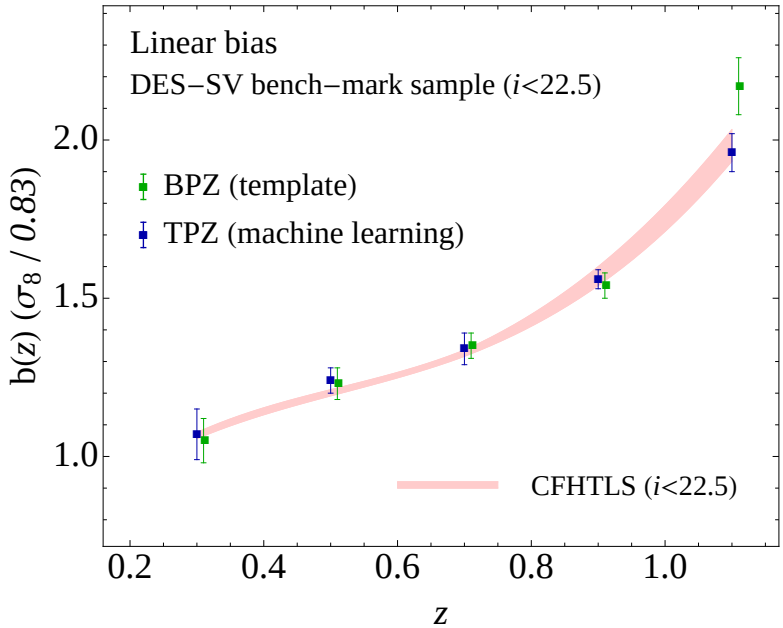

Figure 11. Comparison of the large-scale bias measured in a DES-SV flux limited sample ($i < 22.5$) to equivalent measurements from CFHTLS derived from Coupon et al. (2012). We present DES results for two different photometric redshift catalogs, one obtained using a template method (BPZ), another with a machine learning approach (TPZ). The overall agreement between the two DES samples as a function of redshift is better that 2 per cent for $z < 1$. At $z > 1$ is difference is not statistically significant ($\sim 2\sigma$). This represents a non-trivial test for DES-SV photometric redshift estimation. Our results are also in good agreement with those from CFHTLS, with $\chi^2/d.o.f = 4/5$ for TPZ and $8.7/5$ for BPZ, representing a cross-validation of data quality and sample selection.

its the caption. In Sec. 7.3 we discuss the angular range of validity of a linear bias prescription based on `Halofit`, therefore in Fig. 10 we have chosen to display the clustering measurements to scales that go beyond the linear regime.

There is good qualitative agreement between the linear bias model and the observed clustering at the scales shown in the main panels of Fig. 10 (we discuss this more in quantitative terms below in Sec. 7.3). This result is interesting because it implies that, at least for projected clustering in angular coordinates, the scale of non-linear biasing is different from the one of non-linear dark matter clustering. The latter is currently better understood, so in general terms this result is relevant. We will come back to it in Sec. 7.3.

We further note that at large scales ($\theta \gtrsim 2^\circ$) all the correlations tend to zero. This is a signal that systematic effects are under control, as systematic variations tend to introduce spurious large-scale power (Ross et al. 2011a; Ho et al. 2012; Leistedt et al. 2013)).

In the inset panels of Fig. 10, we show the clustering measurements at smaller angular scales and using a log scaling for $w(\theta)$. The model curves are merely extrapolations using the best-fit bias recovered on larger-scales (main panels). Qualitatively, this simple model does not depart strongly from the measurements. The data at these scales will be used to constrain the halo occupation distribution of DES SV galaxies in Sobreira et al. (in prep.). However, the general agreement signals that a more elaborate non-linear bias prescription based on perturbation theory (Fry & Gaztanaga 1993; McDonald 2006; Chan, Scoccimarro & Sheth 2012) might be adequate to describe the data in this high signal-to-noise regime. Such study is left for future work. Instead, in Sec. 7.3 we will show that the current data size and quality of DES-SV is able to distinguish the breakdown of a "linear bias" prescription in detail.

Lastly we note that we have chosen not to show the clustering for the TPZ sample in a manner analogous to Fig. 10, as the results look almost identical at the qualitative level.

7.1 Bias Evolution and Comparison to CFHTLS

The DES science verification data used in this paper (and in a series of papers accompanying it) is in several regards similar to that collected by the CFHTLS collaboration, for example in depth, photometry and area. CFHTLS has, to this point in time, been regarded as the state of the art for deep wide area photometric data. Hence in this section we compare our clustering measurements to those presented by Coupon et al. (2012).

Galaxies in Coupon et al. (2012) are selected according to SEXtractor `MAG_AUTO` magnitudes of $i < 22.5$, and thus comprise a very similar sample to the one presented in this paper. To facilitate the comparison further we have chosen our photo-z bins similarly to those in the clustering study of Coupon et al. (2012). Although photometric redshifts in CFHTLS were estimated using a different template method, LePhare (Ilbert et al. (2006), Coupon et al. (2009)) we have used a very similar set of templates to obtain our DES BPZ data (Sánchez et al. 2014). The comparison to a neural network photo-z catalog such as TPZ is novel in this regard.

The apparent magnitude sample $i < 22.5$ in Coupon et al. (2012) was then split and reported as several volume limited absolute magnitude "threshold" samples in each photo-z bin. Therefore in each redshift bin, our sample (also selected by $i < 22.5$) corresponds to the faintest absolute magnitude sample in Coupon et al. (2012), except for the fact that Coupon et al. (2012) imposed cuts to make the sample volume limited, where we have not. For example, for $0.4 < z < 0.6$ there is a group of galaxies brighter than $i = 22.5$ but fainter than $M_r - 5\log h = -18.8$ (see their Fig. 4). These galaxies are in our sample but not in the $M_r - 5\log_{10} h < -18.8$ sample of Coupon et al. (2012). In Appendix C we estimate the impact of these differences in selection when comparing $b(z)$ results. We find the differences are most relevant at high z, where the galaxies in the samples have luminosity thresholds, L, such that db/dL is large.

The bias evolution reported by Coupon et al. (2012), accounting

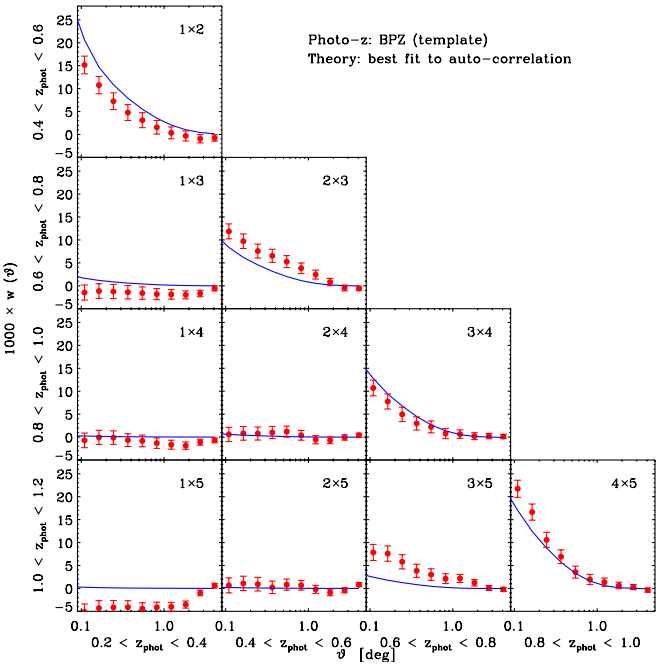

Figure 12. Full set of two-point cross-correlation functions between all of the BPZ redshift bins we use. The theory curves represent a Planck fiducial model, with linear bias values fitted to the auto-correlations displayed in Fig. 10 over scales $\theta > 0.2$ deg. Cases where this produces a match to the observed cross-correlations demonstrates the robustness of the dn/dz estimation.

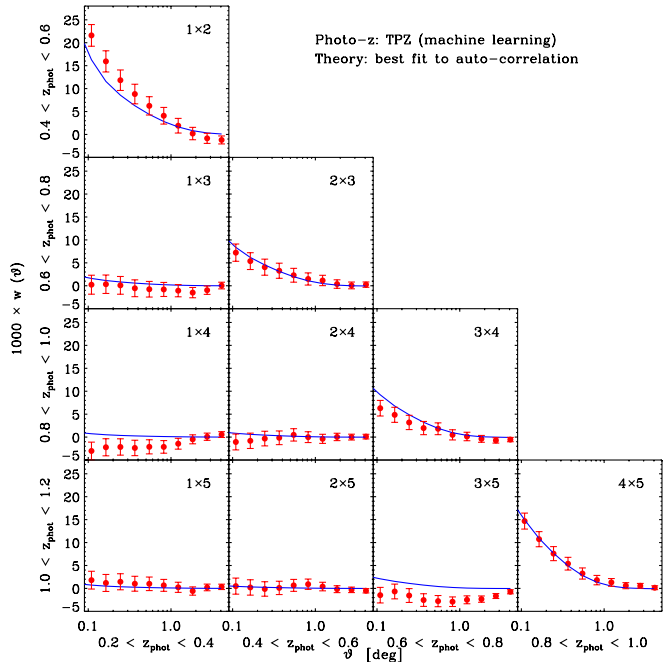

Figure 13. Same as Fig. 12 for the photometric catalog derived with TPZ.

for the differences in sample selection as described in Appendix C, is shown by a filled light-red region in Fig. 11 (where the width corresponds to the statistical errors bars reported by CFHTLS) [13]. The green squares with error bars correspond to our bias measurements for the BPZ sample presented in the previous section and in Table 2, while the blue squares correspond to our TPZ sample as reported in Table 3 (both for the cases where all corrections for observational systematics have been applied).

Overall, the match in the recovered bias as a function of redshift between our apparent magnitude sample and the corresponding one by CFHTLS is good, with a slight tension at the highest bin for BPZ. Note however that for CHFTLS the error bars depicted in Fig. 11 are statistical only (as reported by Coupon et al. (2012)), while in our case we include effects such as the reduction of effective area due to further masking of bad observing regions, or the impact of stellar contamination. If we define

$$\chi^2 = \sum_{\text{bin } j=1}^{5} (b^{(j)}_{\rm DES} - b^{(j)}_{\rm CFHTLS})^2/(\sigma^{(j)}_b)^2 \qquad (28)$$

and add the errors in quadrature, $\sigma_b^2 = \sigma^2_{b,\,\rm DES} + \sigma^2_{b,\,\rm CFHTLS}$ we find that $\chi^2(BPZ) = 8.7$ while $\chi^2(TPZ) = 4$ for five degrees of freedom in each case. Assuming that the DES and CFHTLS data are drawn from the same underlying Gaussian distribution, greater χ^2 would be expected in 55 per cent of cases for the TPZ comparison and 12 per cent of cases for the BPZ comparison. This means that the agreement is better than 1σ for TPZ and better than 2σ for BPZ.

We stress that the agreement for the bias evolution between the two photo-z catalogs is non-obvious because different photo-z codes could, in principle, select different types of galaxies in different redshift bins (note that the algorithms implemented in BPZ and TPZ are very different). This might explain the difference at the last z-bin if BPZ were selecting a rarer population with a higher bias. However we note that if the two samples were independent, the disagreement would be 2.1σ, and thus not remarkable. For $z < 1.0$, the two techniques return samples with bias values that agree to better than 2 per cent, suggesting that at these redshifts, systematic uncertainties associated with the photo-zs in this redshift range are well-controlled.

The increase in bias with redshift found in Fig. 11 is primarily a consequence of the effective luminosity threshold, as this threshold naturally increases with redshift for a flux limited sample. Only the high luminosity galaxies will be observed at high z, and they are more biased. This effect accentuates for $z \gtrsim 0.8$ as the majority of galaxies come from the exponential tail of the luminosity function. The $b(z)$ shown in Fig. 11 is therefore not physical in the sense that the sample is changing with redshift. In Appendix A we show that this evolution of galaxy bias as a function of redshift is also recovered in simulations based on Halo Occupation Distribution techniques to populate halos with galaxies.

7.2 Consistency from cross-correlations of photo-z bins

As a further test of potential systematics and photo-z we show in Figs. 12 and 13 the full set of cross-correlation measurements between photo-z bins for the BPZ and TPZ samples, respectively (see Christodoulou et al. (2012) for a related analysis using SDSS data). These measurements were done using a cumulative mask, for each photo-z estimator, combining the ones derived for each photo-z bin in Sec. 6.1.

For bins that are well-separated in redshift, we observe that their cross-correlations are consistent with zero, as expected from the redshift distributions in Fig. 3. One notable exception is the measured cross-correlation between bins $0.2 < z < 0.4$ and $1.0 < z < 1.2$ for BPZ, which is significantly negative. Understanding the cause of this is left for future work.

[13] We have re-scaled the CFHTLS measurements that assumed at $z = 0$ $\sigma_8 = 0.8$ to our cosmology $\sigma_8 = 0.83$

Solid lines in Figs. 12 and 13 correspond to our fiducial model, given by Eq. (2), for w_{ij} provided with the redshift distributions in photo-z bin i and j. They use a bias $b = \sqrt{b_i b_j}$ where b_i and b_j are the the best fit "linear growth" model bias from the auto-correlations. This is therefore a fully predictive model given the auto-correlations, and a test for outliers and the reconstructed dn/dz.

In Tables 4 and 5 we compare the predicted $\sqrt{b_i b_j}$ with bias values determined by directly fitting the cross-correlation functions over the same range of scales ($\theta > 0.2$ deg), which we call b_{ij}[14]. We estimate the error on the bias prediction $\sqrt{b_i b_j}$ by a direct propagation of the error on the auto-correlation biases, σ_{b_i} and σ_{b_j},

$$\sigma_{\sqrt{b_i b_j}} = (\sqrt{b_i b_j}/2)(\sigma_{b_i}/b_i + \sigma_{b_j}/b_j). \tag{29}$$

The fourth column displays the absolute difference between $\sqrt{b_i b_j}$ and b_{ij} divided by their error combined in quadrature. In the 20 total cases tested (considering both photo-z methods), we find only one case where $\Delta b/\sigma_b$ is greater than 2.21σ (BPZ bin 3×5) and only two cases where they are greater than 1.45σ (the additional case is TPZ bin 1×2). Considered in total, the cross-correlation results agree with the auto-correlation results at roughly the expected statistical level.

The picture recovered from Figs. 12 and 13 and Tables 4 and 5 is that outliers are not causing obvious inconsistencies in our analysis, and that to a large extent our estimated redshift distributions are correct. However, future DES data will require more work to assess the level of systematic uncertainty associated with photo$-z$ estimation.

7.3 On the Scale of Linear Bias

In Sec. 7 and Fig. 10 we showed that on large angular scales the clustering recovered in our data agrees well with the standard linear theory prescription: $b^2\, w_{\rm DM,\,Lin}(\theta)$ for all photo-z bins. We further noted that at smaller scales the assumption of linear bias appeared to still hold provided the description of the underlying dark matter clustering accounted for non-linear growth of structure (for which we used `Halofit` from Takahashi et al. 2012). In this section we explore in quantitative terms the scale at which this assumption leads to a significant data-theory discrepancy[15].

We define the transition from linear to non-linear dark matter clustering as the minimum scale such that $w_{\rm DM,\,Halofit} = w_{\rm DM,\,Lin}$. Given our photo-$z$ bins this occurs at $\theta_{\rm Linear\,Growth} \sim (0.26, 0.16, 0.11, 0.09, 0.08)$ deg. for bins one to five. This corresponds to roughly $4\,{\rm Mpc}\,h^{-1}$ of comoving separation across redshift given our cosmological model. The best-fit biases and χ^2 on linear scales ($\theta > \theta_{\rm Linear\,Growth}$) are provided in Table 2.

To state at what point the "linear bias" model becomes notably worse than the "linear growth" model we perform the following test. We first find the χ^2 using the data vector and covariance matrix for the angular bins where $\theta > \theta_{\rm Linear\,Growth}$ and the "linear growth" model. Then we add new data to the fit by decreasing the minimum angular scale considered, $\theta_{\rm min} < \theta_{\rm Linear\,Growth}$, and using a "linear bias" model we determine the difference in χ^2 values between the extended data vector and the "linear growth" one. A large $\Delta\chi^2$ (depending on the extra data-points, $\Delta({\rm d.o.f.})$) implies that a model with a single bias parameter is no longer a good description of the extended data vector. In quantitative terms, we define $\theta_{\rm min}$ such that,

$$\Delta\chi^2 - \Delta({\rm d.o.f}) = 4 \tag{30}$$

[14] Note that we use the same range of scales here for all redshift bins, so that a fair comparison is possible for all cross-correlations. The biases thus differ slightly than those that would be inferred from Table 2.

[15] In this section, we focus on the BPZ data for concreteness but the conclusions reached in this section do not significantly depend on this choice.

BPZ cross-correlations			
Bin i × j	b_{ij}	$\sqrt{b_i b_j}$	$\Delta b/\sigma_b$
1×2	0.98 ± 0.07	1.10 ± 0.05	1.45
1×3	0.00 ± 1.23	1.12 ± 0.05	0.91
1×4	0.01 ± 4.86	1.19 ± 0.05	0.24
1×5	0.00 ± 3.99	1.44 ± 0.07	0.36
2×3	1.25 ± 0.14	1.26 ± 0.05	0.04
2×4	0.00 ± 1.64	1.33 ± 0.05	0.81
2×5	0.04 ± 5.74	1.61 ± 0.08	0.27
3×4	1.22 ± 0.13	1.36 ± 0.06	0.97
3×5	3.27 ± 0.43	1.65 ± 0.08	3.74
4×5	1.96 ± 0.12	1.75 ± 0.08	1.39

Table 4. Comparison of best-fit biases b_{ij} obtained from fitting the two-point cross-correlation functions between photo-z bins i and j (displayed in Fig. 12) vs. the prediction for such signal from the best-fit to auto-correlations in z-bin i and z-bin j, given by $\sqrt{b_i b_j}$. The fourth column shows the absolute difference $\Delta b = |b_{ij} - \sqrt{b_i b_j}|$ divided by their error combined in quadrature.

TPZ cross-correlations			
Bin i × j	b_{ij}	$\sqrt{b_i b_j}$	$\Delta b/\sigma_b$
1×2	1.35 ± 0.08	1.13 ± 0.05	2.21
1×3	0.00 ± 1.91	1.12 ± 0.05	0.58
1×4	0.00 ± 1.79	1.24 ± 0.06	0.69
1×5	0.00 ± 3.28	1.36 ± 0.07	0.41
2×3	1.01 ± 0.21	1.25 ± 0.05	1.11
2×4	0.00 ± 1.40	1.39 ± 0.06	0.99
2×5	0.00 ± 2.35	1.53 ± 0.08	0.64
3×4	1.11 ± 0.20	1.37 ± 0.06	1.27
3×5	0.00 ± 2.14	1.51 ± 0.07	0.70
4×5	1.61 ± 0.15	1.68 ± 0.08	0.39

Table 5. Same as Table 4, but for the TPZ cross correlations displayed in Fig. 13.

as the point where a 2σ preference for an improved (but unknown) model for the bias as a function of scale exists. We thus define the largest angular scale at which this condition is met as our measurement of $\theta_{\rm Linear\,Bias}$. In what follows, we refer to the results of this test as the results of the 'χ^2 test'.

The results of the χ^2 test are shown in Fig. 14 for all z -bins. We find $\theta_{\rm Linear\,Bias} \sim (0.06, 0.05, 0.04, 0.037, 0.007)$ deg. for bins one to five. We consider the accuracy at which we detect the breakdown of the "linear bias" scheme as half[16] the precision in the measurement of the correlation function at $\theta_{\rm Linear\,Bias}$. For our current SV sample this is $\sim 2.5\%$ for the central photo-z bins, and $\sim 4\%$ for bins one and five.

In Fig. 15 (top panel) we plot the resulting $\theta_{\rm Linear\,Bias}$ as a function of redshift (solid purple curve) and compare it to $\theta_{\rm Linear\,Growth}$ (dashed orange curve). The scale where the "linear bias" model fails is approximately three times smaller than the scale where non-linear growth becomes important, except for the $1.0 < z < 1.2$ bin, where

[16] Because $w(\theta) \propto b^2$

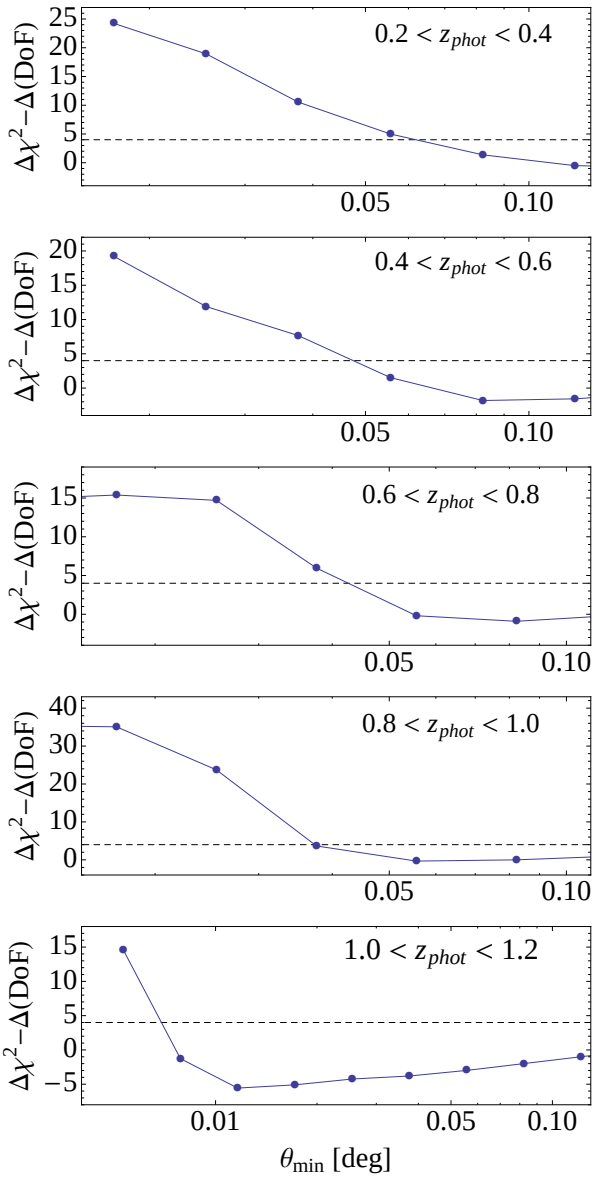

Figure 14. We estimate the smallest scale $\theta_{\rm min}$ at which a linear bias model is still a good description of the clustering by computing the χ^2 difference between a fit extending to $\theta_{\rm min}$ and a fit only on large linear scales. A $\Delta\chi^2 \sim 4 + \Delta(dof)$ roughly corresponds to 2σ evidence that the linear bias model fails to describe the data.

we find a factor of 10. The bottom panel of Fig. 15 shows the same information but translated to co-moving separations at the centre of each redshift bin, assuming our fiducial cosmological model. The scale of linear growth is roughly constant with redshift, $\sim 4\,h^{-1}\,\mathrm{Mpc}$, while the scale of linear bias is $\sim 1\,h^{-1}\,\mathrm{Mpc}$ (except in the last bin).

In Fig 16 we present, as a cross-check, the resulting best-fit bias as a function of the θ_{min} tested when obtaining the χ^2 test results. The dashed line in each panel shows the "linear growth" bias from Tables and 1 and 2. As we extend the fit down to smaller $\theta_{\rm min}$ we recover smaller error-bars, as expected. Further, the measured bias values remain consistent with the dashed line, until approximately the scale of Linear Bias determined using the χ^2 test. We find that in all redshift bins, the measured bias begins to be clearly under-estimated compared to the asymptotic large-scale value (this would be consistent with a negative 2nd-order bias parameter in the local Eulerian bias model) at close to the θ_{min} determined using the χ^2 test.

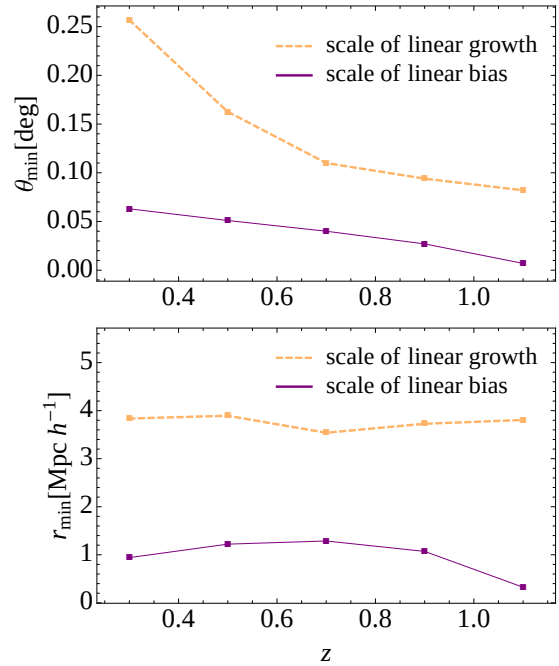

Figure 15. Top Panel: the evolution of the scale down to which the linear bias model (with non-linear dark matter) reproduces our angular clustering measurements, compared to the evolution of the linear bias + linear dark matter clustering. The former is valid to considerably smaller scales. Bottom Panel: Same information but translated into co-moving distances at the given redshift.

These results are interesting because they imply that forthcoming DES studies might be able to incorporate higher signal-to-noise data into cosmological analyses that use the full shape of the correlation function information with a simple bias model. We do stress that a more thorough analysis would involve allowing for a more complex model with non-linear and/or non-local bias terms and then investigate how the best-fit value for the linear order term in such scheme compares to the linear model results we present. This is however beyond the scope of the present paper and we postpone it to future work.

8 CONCLUSIONS

We present the first large-scale clustering analysis of a galaxy sample selected from the Science Verification Data of the Dark Energy Survey. The sample is selected as an apparent magnitude limited sample $i < 22.5$, with no color selection except for very conservative cuts to remove color outliers. The sample extends from $z_{\rm phot} = 0.2$ to $z_{\rm phot} = 1.2$. We performed our analysis in five tomographic bins of width $\Delta z_{\rm phot} = 0.2$.

This paper has three main foci:

(i) We perform a detailed analysis and amelioration of potential observational systematic effects. We use a set of maps for different variables that can modulate galaxy detection efficiency, which include varying observing conditions, stellar density, and Galactic dust. We analyze which of them are affecting our sample by means of angular cross correlations and a measurement of galaxy density as a function of the value of the potential systematic variable across the footprint. We then apply cuts to minimize these systematic correlations, with a small loss of statistical power. We then show how these two approaches complement and validate each other. These methods can be widely applied to future clustering studies, in-

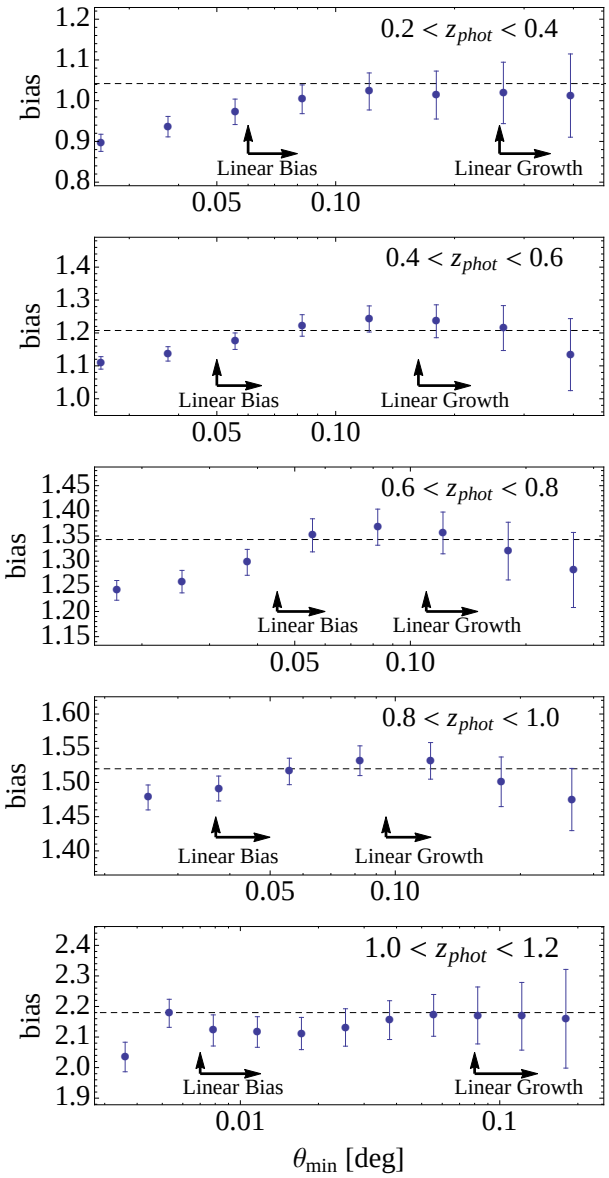

Figure 16. Best fit bias as a function of the minimum angular scale included in the fit. The inset label "Linear Growth" indicates the range of scales where a model with linear bias and linear matter clustering (hence "linear growth") applies in our data. The "Linear Bias" label corresponds to a model using nonlinear matter clustering (`Halofit` by Takahashi et al. (2012)) but keeping the bias linear.

cluding our detailed accounting of the effects of stellar contamination.

(ii) We evaluate the clustering in the five redshift bins comparing it against simple linear theory models. We find good agreement with these models and between the results we obtain from two separate photo-z methods (one template based, the other machine learning) of determining DES photometric redshifts. Both sets of measurements are also consistent with bias measurements, determined at the same redshifts for the same $i < 22.5$ flux limit, obtained by the CHFTLS survey. We measure the cross-correlation between these redshift bins as a test of the photo-z reconstruction, finding the cross-correlation matching predictions from the autocorrelation signal at the expected statistical level.

(iii) We explore the regime of validity of our simple models that involve either linear / non-linear dark matter clustering, and a linear bias term. We find that, in angular clustering, the scale at which a linear bias model is not consistent with our data is considerably smaller than the scale at which non-linear growth in the clustering of dark matter (as predicted by `Halofit` 2012) becomes important compared to the linear theory predictions. These results are relevant for probes that aim to combine weak lensing with large scale structure, as they suggest a linear bias model could be sufficient to scales as small as $1h^{-1}$Mpc.

The data we analyze is approximately 1/30th the size of the of final DES data set[17]. In the near future the methods presented here will be used and extended to larger DES samples suitable to measure the Baryon Acoustic Oscillation scale, constrain the shape of the matter power spectrum and measure the growth of structure. As part of this future analysis, we will use additional tools to robustly assess statistical and systematic uncertainties. We will make greater use of the `Balrog` tool presented in Suchyta et al. (2015), which will enable both a larger data sample and a more comprehensive assessment of systematic variations. Most importantly, we will use detailed mock catalogues and simulated data to provide robust and accurate covariance matrix estimates. These simulations will also be used to assess the systematic uncertainty introduced by the various systematic mitigation techniques and uncertainties in the photo-zs. Further, we will consider the impact of time-dependent variations in the data quality comprising the co-add. The work presented in this paper will help inform the creation of these simulations and we will continue to refine and apply the tools we have presented throughout this paper. In this sense, the methods and results we present represents a first step towards the goal of constraining cosmology with DES LSS measurements.

ACKNOWLEDGMENTS

We are grateful for the extraordinary contributions of our CTIO colleagues and the DECam Construction, Commissioning and Science Verification teams in achieving the excellent instrument and telescope conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the DES Data Management group.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The DES data management system is supported by the National Science Foundation under Grant Number AST-1138766.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgeticas, Medioambientales y

[17] The SV SPTE region of observations is 150 deg^2 and the final DES footprint will be 5000 deg^2

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

The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2012-39559, ESP2013-48274, FPA2013-47986, and Centro de Excelencia Severo Ochoa SEV-2012-0234. Research leading to these results has received funding from the European Research Council under the European Union Seventh Framework Programme (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. MC has been partially funded by AYA2013-44327. FS acknowledges financial support provided by CAPES under contract No. 3171-13-2. We thank Jean Coupon and Martin Kilbinger for useful discussions and help at different stages of this work.

This paper has gone through internal review by the DES collaboration. The DES publication number for this article is DES-2015-0055. The Fermilab pre-print number is FERMILAB-PUB-15-305.

REFERENCES

Ade P. A. R. et al., 2014, A&A, 571, A16
Benítez N., 2000, ApJ, 536, 571
Bertin E., 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, Gabriel C., Arviset C., Ponz D., Enrique S., eds., p. 112
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cabré A., Fosalba P., Gaztañaga E., Manera M., 2007, MNRAS, 381, 1347
Carlstrom J. E. et al., 2011, PASP, 123, 568
Carnero A., Sánchez E., Crocce M., Cabré A., Gaztañaga E., 2012, MNRAS, 419, 1689
Carrasco Kind M., Brunner R. J., 2013, MNRAS, 432, 1483
Carrasco Kind M., Brunner R. J., 2014, MNRAS, 442, 3380
Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, MNRAS, 447, 646
Chan K. C., Scoccimarro R., Sheth R. K., 2012, PRD, 85, 083509
Chang C. et al., 2014, ArXiv e-prints
Christodoulou L. et al., 2012, MNRAS, 425, 1527
Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926
Coleman G. D., Wu C.-C., Weedman D. W., 1980, ApJS, 43, 393
Colless M. et al., 2001, MNRAS, 328, 1039
Cooper M. C. et al., 2012, MNRAS, 425, 2116
Coupon J. et al., 2009, A&A, 500, 981
Coupon J. et al., 2012, Astronomy and Astrophysics, 542, A5
Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS, 414, 329
Crocce M., Castander F. J., Gaztanaga E., Fosalba P., Carretero J., 2013, ArXiv e-prints
Crocce M., Gaztañaga E., Cabré A., Carnero A., Sánchez E., 2011, MNRAS, 417, 2577
Dawson K. S. et al., 2013, AJ, 145, 10
de Jong J. T. A. et al., 2013, The Messenger, 154, 44
de Simoni F. et al., 2013, MNRAS, 435, 3017
Desai S. et al., 2012, ApJ, 757, 83
Diehl H. T., and the Dark Energy Survey Collaboration, 2012, Physics Procedia, 37, 1332
Dodelson S., 2003, Modern cosmology
Drinkwater M. J. et al., 2010, MNRAS, 401, 1429
Eriksen M., Gaztanaga E., 2014, ArXiv e-prints
Flaugher B. et al., 2015, ArXiv e-prints
Flaugher B. e. a., 2005, International Journal of Modern Physics A, 20, 3121
Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015a, MNRAS, 448, 2987
Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015b, MNRAS, 447, 1319
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Giannantonio T., et al., 2015, ArXiv e-prints
Glazebrook K., Peacock J. A., Collins C. A., Miller L., 1994, MNRAS, 266, 65
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
High F. W., Stubbs C. W., Rest A., Stalder B., Challis P., 2009, AJ, 138, 110
Ho S. et al., 2012, ApJ, 761, 14
Ilbert O. et al., 2006, A&A, 457, 841
Kaiser N., Tonry J. L., Luppino G. A., 2000, PASP, 112, 768
Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38
Kuehn K. et al., 2013, PASP, 125, 409
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Le Fèvre O. et al., 2005, A&A, 439, 845
Leauthaud A. et al., 2007, ApJS, 172, 219
Leistedt B., Peiris H. V., Mortlock D. J., Benoit-Lévy A., Pontzen A., 2013, MNRAS, 435, 1857
Leistedt B., et al., 2015, ArXiv e-prints
Lilly S. J. et al., 2007, ApJS, 172, 70
Martí P., Miquel R., Bauer A., Gaztañaga E., 2014, MNRAS, 437, 3490
McDonald P., 2006, PRD, 74, 103512
Miyazaki S. et al., 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 0
Mohr J. J. et al., 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 0
Monet D. G. et al., 2003, AJ, 125, 984
Myers A. D. et al., 2006, ApJ, 638, 622
Padmanabhan N. et al., 2007, MNRAS, 378, 852
Ross A. J., Brunner R. J., 2009, MNRAS, 399, 878
Ross A. J. et al., 2011a, MNRAS, 417, 1350
Ross A. J., Percival W. J., Crocce M., Cabré A., Gaztañaga E., 2011b, MNRAS, 415, 2193
Sánchez C. et al., 2014, MNRAS, 445, 1482
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Scranton R. et al., 2002, ApJ, 579, 48
Sevilla I. et al., 2011, ArXiv e-prints
Soumagnac M. T. et al., 2013, ArXiv e-prints
Suchyta E., et al., 2015, in prep.
Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Thomas S. A., Abdalla F. B., Lahav O., 2010, Physical Review Letters, 105, 031301
Tucker D. L. et al., 2007, in Astronomical Society of the Pacific Conference Series, Vol. 364, The Future of Photometric, Spectrophoto-

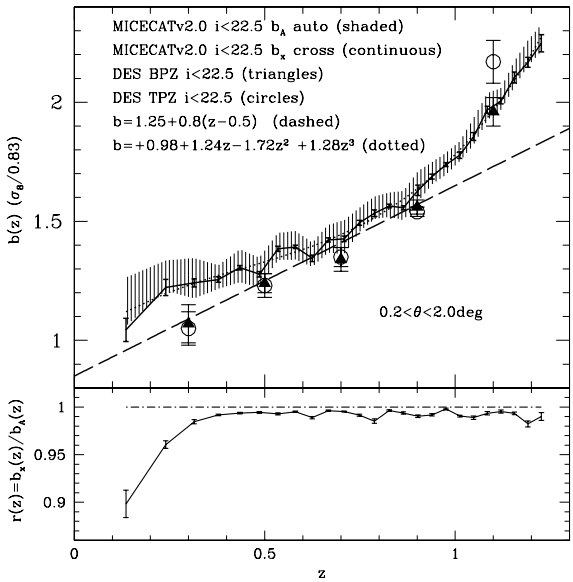

Figure 17. Top panel shows the evolution of galaxy bias with redshift in a flux limited sample ($i < 22.5$) similar to the one defined in this paper but selected from our MICECATv2.0 mock galaxy catalog. We display bias derived from galaxy auto-correlations $b_A = \sqrt{w_{gg}/w_{mm}}$ and galaxy-matter cross correlation $b_X = w_{gm}/w_{mm}$. The evolution is similar to the one in the DES-SV data. Bottom panel shows the ratio of the two estimates, i.e. the stochasticity coefficient $r = b_A/b_X$, which we found close to unity in the redshift range relevant for DES data.

metric and Polarimetric Standardization, Sterken C., ed., p. 187
Tyson J. A., Wittman D. M., Hennawi J. F., Spergel D. N., 2003, Nuclear Physics B Proceedings Supplements, 124, 21
York D. G. et al., 2000, AJ, 120, 1579

APPENDIX

A Testing the bias evolution against MICE bench-sample

Here we test the evolution of bias for our flux-limited sample using the MICECATv2.0, simulation which is an updated version from the MICECATv1.0 catalog (see Fosalba et al. (2015a), Crocce et al. (2013), Fosalba et al. (2015b), Carretero et al. (2015)) extended to lower mass halos and therefore less luminous galaxies. This version is complete to $i \simeq 24$ at all redshifts. Magnitudes are assigned to galaxies following a combination of Halo Occupation Distribution (HOD) and subhalo abundance matching (SHAM) prescriptions tuned to fit the local ($z \sim 0$) galaxy luminosity function and clustering as a function of luminosity and color (see Carretero et al. (2015)). To account for evolution in galaxy luminosity on top of the one induced by halo mass evolution, we transform the galaxy magnitudes using:

$$i_{evol} = i - 0.8[atan(1.5z) - 0.1489], \qquad (31)$$

which gives similar counts to those in COSMOS and other galaxy surveys (Castander et al, in prep.). We then restrict the sample to $i_{evol} < 22.5$ to compare to the DES SV bench-mark sample.

The top panel of Fig. 17 shows a comparison of the linear bias measured in MICECATv2.0 against the one in our BPZ and TPZ samples shown in Fig. 11. In the case of MICECATv2.0 we estimate the bias as $b_A = \sqrt{w_{gg}/w_{mm}}$, where w_{gg} is the measured galaxy angular correlation function and w_{mm} the corresponding measured matter-matter correlation in the same redshift bin. This is done using true redshifts to avoid redshift space distortions, which can have a large effect when using narrow redshift bins as in Fig. 17 (Eriksen & Gaztanaga 2014) but are negligible (given scales and errors) for $\Delta z = 0.2$ bins used in this paper (Crocce, Cabré & Gaztañaga 2011; Ross et al. 2011b). The MICECATv2.0 bias is fitted from the ratio in the range $0.2 < \theta < 2.0$ deg. The solid line is the bias measured from cross-correlating galaxies and mass directly, $b_X = w_{gm}/w_{mm}$. The shaded region tries to reproduce the expected sampling and shot-noise errors for DES SV in b_A. It corresponds to $\Delta b_A = (\Delta w_{gg}/w_{mm})/(2b_A)$, where matter is kept fixed and the error in the galaxy-galaxy correlation Δw_{gg} is scaled to 116 deg^2 and $\Delta z \simeq 0.2$, from the scatter in 24 JK regions over a larger 30x30 deg^2 simulation patch. The smaller error bars include sample variance cancelation between matter and galaxy fluctuations (i.e. the scatter in b_X from different regions). The dotted line shows a cubic fit in simulations to b_X and errors. The evolution with redshift is very similar to the one in the LSS bench-mark sample, in particular the fast rise at high redshift. At $z < 0.5$ DES-SV data seems to be 1σ lower than simulations. Larger areas are needed to decide if this is a significant discrepancy. The bottom panel of Fig. 17 shows the stochasticity coefficient: $r \equiv w_{gm}/\sqrt{w_{gg}w_{mm}} = b_X/b_A$. The error bars are scaled from the dispersion in r from JK regions: thus, they include sample variance cancelation and correspond to the 30x30 deg. patch used for this test. There seems to be some significant detection of r at the smallest redshifts, but this is most likely a reflection of the poor performance of JK errors for the smallest volumes. At $z > 0.3$ we find r to be in the range $r \simeq 0.98 - 1.00$, which is consistent with unity when errors are scaled to the SV area. This is relevant for analyses using probes that cross correlate the bench-mark galaxy sample defined here with other data sets such as CMB lensing (Giannantonio et al. 2015) or weak lensing maps.

Note that the bias is displayed in units of $(\sigma_8/0.83)$ to account for the difference in the MICE value of $\sigma_8 = 0.80$ and the one used in this paper ($\sigma = 0.83$). There is quite a good agreement in the values of $b(z)$ without need of any other adjustment. This illustrates the fact that the steep rise of bias at $z > 1$ is in fact expected using simple HOD and luminosity evolution models.

B Statistical error estimation

In Sec. 2.3 we introduced our approach to estimate the covariance matrix between angular clustering measurements, see Eq. (11). This combined a theory modeling for the off-diagonal elements with a Jackknife (JK) approach for the variance. The JK approach is intrinsically limited by the fact that typically (given the size of our footprint and data) is not possible to extract many statistically independent samples to construct a reliable full covariance. However it does account better than theory estimations for issues such as small scale mask effects, boundaries, non-linerities and/or contamination or systematics in the data. On the other hand, the theory estimation yields off diagonal elements that have no noise compared to their JK counter-parts. In this appendix we explore this choice in more detail.

A first test is to compare the error estimates on $w(\theta)$ measurements (i.e. the variance) that we obtain using the JK approach with the theory ones. For the theory estimate we used the best-fit bias over linear-scales, the exact footprint size and the exact shot-noise measured in the data (i.e. number density of objects in the photo-z bin after masking). The JK one employs 40 equal size JK regions, and we tested that the results that we now present do not depend on variations in this set-up (we have explored n_{JK} from 20 to 100). In Fig. 18 we show how these two estimates compare to each other. On large-scales ($\theta \gtrsim 0.1$ deg at low z, and $\theta \gtrsim 0.01$ deg at high z) they match each other well, except for the $0.8 < z < 1.0$ where the theory yields $\sim 20\% - 25\%$ higher errors in the range $0.1\,\mathrm{deg} < \theta < 1\,\mathrm{deg}$.

Photo-z Bin	Jack-knife Cov.	Theory Cov.	Mix Cov.
$0.2 < z < 0.4$	1.05 ± 0.07	1.04 ± 0.08	1.05 ± 0.07
$0.4 < z < 0.6$	1.22 ± 0.04	1.23 ± 0.06	1.23 ± 0.05
$0.6 < z < 0.8$	1.37 ± 0.03	1.36 ± 0.04	1.36 ± 0.04
$0.8 < z < 1.0$	1.55 ± 0.02	1.54 ± 0.04	1.54 ± 0.02
$1.0 < z < 1.2$	2.16 ± 0.12	2.16 ± 0.12	2.17 ± 0.09

Table 6. Dependence of best-fit bias on different methods to estimate the covariance matrix (see text for details). The bias is obtained from fitting the auto-correlations in Fig. 11 over a linear range of scales This paper uses the mixed approach for its main results.

A second test is to derive the best-fit bias using three possible methods for constructing the covariance matrix, a full theory approach, the JK method, or the combined one that we use in Sec. 2.3. This is shown in Table 6. The best-fit biases are not affected by the choice of covariance, while the 1σ errors vary, but not so much as to affect any of our conclusions. At smaller scales, the JK uncertainties are significantly larger than the theoretical ones; this is likely due to both non-linear and small-scale mask effects.

C Galaxy Bias in CFHTLS

Galaxies in Coupon et al. (2012) are first selected according to `SExtractor` `MAG_AUTO` magnitudes of $i < 22.5$, similar to the sample presented in this paper. Their flux limited sample was then split in several volume limited absolute magnitude "threshold" samples in each photo-z bin. Therefore in each redshift bin, our sample corresponds to the faintest absolute magnitude sample in Coupon et al. (2012), except for the fact that our sample is not volume limited.

For example over the redshift range $0.4 < z < 0.6$ the faintest volume limited sample is defined as brighter than $M_r - 5\log h = -18.8$ because this is the absolute magnitude corresponding to apparent magnitude $i < 22.5$ for the farthest galaxies at $z = 0.6$. At $z = 0.4$ the absolute magnitude corresponding to $i < 22.5$ is one magnitude fainter, $M_r - 5\log h = -17.8$. Therefore as one spans the whole redshift range of the sample from $z = 0.4$ to $z = 0.6$ there is a set of galaxies brighter than $M_r - 5\log h = -17.8$ but fainter than $M_r - 5\log h = -18.8$, that do not enter the volume limited sample in $0.4 < z < 0.6$ but do enter the flux limited one. And this holds for all photo-z bins: there is always one magnitude difference between the faintest absolute magnitude from the lowest to the highest redshift of each photo-z bin (Coupon et al. 2012).

To directly compare our bias measurements to those in Coupon et al. (2012) we need to know how bias depends on luminosity to then estimate the contribution to the bias of this "extra" set of galaxies, as

$$b_{\rm flux\,lim,\,bin\,N} = \int_0^1 b_{\rm vol\,lim}(< M_r + x)\, dx \qquad (32)$$

where M_r is the magnitude threshold defining the faintest sample for the photo-z bin N. Here $x = 1$ would correspond to the lowest redshift bound of such bin, while $x = 0$ the highest.

At low redshift the difference discussed above does not impact the derived bias strongly because these galaxies are faint enough that have biases close to one and are in a regime of very weak dependence with luminosity. However as we move to higher redshifts, the faintest samples move to $M_r - 5\log_{10} h \lesssim -21$ where bias evolves strongly with luminosity (e.g. see Fig. 18 in Coupon et al. (2012)).

To account for this effect so as to properly compare our bias results to those of CFHTLS we did as follows. We collected all the best-fit biases reported by Coupon et al. (2012) in their Table B.1. This is shown in Fig. 19. We have found that these measurements are very well fit by a simple linear relation with r-band luminosity

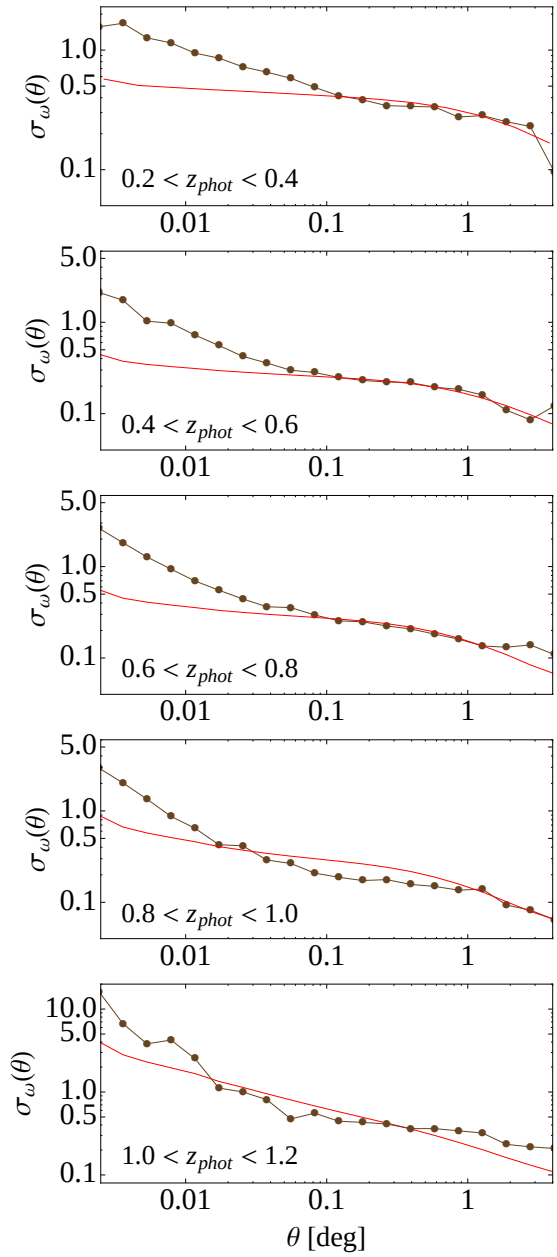

Figure 18. Comparison of Jackknife resampling errors (points) and theoretically derived ones (lines) for our five tomographic bins (BPZ photo-z). For the former we used 40 equal size regions. The theory estimation uses the best-fit bias, the projected number density and the area of the footprint (and its based on the nonlinear matter power spectra). The agreement is very good on scales ($\theta \gtrsim 1$ deg at low-z and $\theta \gtrsim 0.1$ deg at high-z), while on smaller scales the structure of the data as captured by the JK resampling yields larger variance.

$b(M_r) = b_0(z) + a_0 L_r$ with constant $a_0 = 0.13/L_\star^{z=0.1}$. This result is in fact very similar to the one discussed by Coupon et al. (2012), although we do not normalize by the characteristic luminosity at the given redshift. We then fit b_0 at every photo-z bin, including $1.0 < z < 1.2$ where only one data-point is given by Coupon et al. (2012). We find $b_0 = (1.08, 1.18, 1.25, 1.35, 1.4)$ at photo-z bins centered at $(0.3, 0.5, 0.7, 0.9, 1.1)$. These results are displayed in Fig. 19.

Once provided with a fit for $b(< M_r)$ we computed the bias conversion as in Eq. (32). Results are detailed in Table 7. The fourth column displays the values that we use in Sec. 7 and Fig. 11. Because the effect is not strong and the bias range not large, the results from an

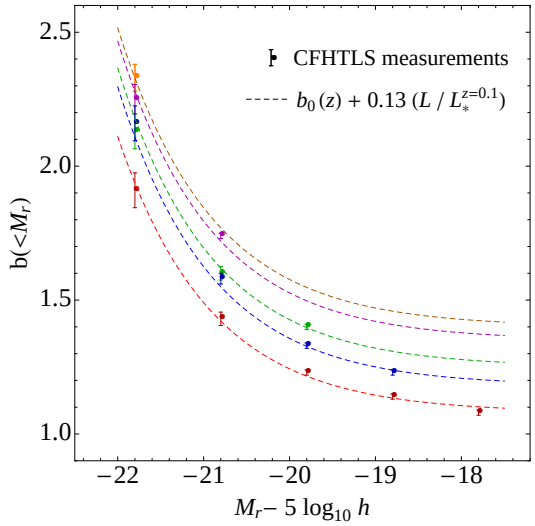

Figure 19. Fit to the bias as a function of luminosity presented by Coupon et al. (2012) for the 5 bins also used throughout our paper (starting from $0.2 < z < 0.4$ at the bottom). The evolution can be simply described as depending linearly with luminosity (see inset label).

Photo-z Bin	$M_r - 5\log h$	$b_{\rm vol\,lim}(< M_r)$	$b_{\rm flux\,lim}(< M_r)$
$0.2 < z < 0.4$	-17.8	1.08 ± 0.01	1.09
$0.4 < z < 0.6$	-18.8	1.23 ± 0.01	1.22
$0.6 < z < 0.8$	-19.8	1.40 ± 0.01	1.35
$0.8 < z < 1.0$	-20.8	1.74 ± 0.03	1.59
$1.0 < z < 1.2$	-21.8	$2.33^{+0.05}_{-0.06}$	2.00

Table 7. Translation of the biases defined by Coupon et al. (2012) where samples are defined in terms of volume limited quantities to our case where we have a flux-limited sample across the full redshift range (the parent catalog in both cases is similar, $i - band$ MAG_AUTO < 22.5). Here M_r is the absolute magnitude corresponding to the faintest sample in each photo-z bin as defined by Coupon et al. (2012), $b_{\rm vol\,lim}$ are the bias values reported in their Table B.1., and $b_{\rm flux\,lim}$ the ones directly comparable to our results.

integration as in Eq. (32) are equivalent to simply computing the bias for samples shifted half by a magnitude (i.e. $b_{\rm flux\,lim}(< -17.8+0.5)$ for $0.2 < z < 0.4$, then $b_{\rm flux\,lim}(< -18.8+0.5)$ for $0.4 < z < 0.6$ and so on).

D Stellar Contamination

Building from the results of, e.g., Myers et al. (2006); Ross et al. (2011a); Ho et al. (2012), we derive the effect that stellar contamination has on the measured clustering of a galaxy sample. We assume the observed number of galaxies in some location, N_o, is a function of the true number of galaxies, the number of stars, N_{st}, that could be mis-classified as a galaxy, and some function of the survey conditions, S. If the only source that modulates the observed number of galaxies is stellar contamination, then

$$N_o = N_{gal} + N_{st}F(S), \tag{33}$$

where $N_{st}F(S)$ is the number of stars mis-classified as galaxies at a particular location. The fractional stellar contamination, f_{st}, over the footprint is

$$f_{st} = \frac{\langle N_{st} f(S)\rangle}{\langle N_o\rangle} \tag{34}$$

and

$$\langle N_o\rangle = (1+f_{st})\langle N_{gal}\rangle. \tag{35}$$

Given the over-density

$$\delta = N/\langle N\rangle - 1, \tag{36}$$

we have

$$\delta_o = \frac{1}{1+f_{st}}\frac{N_{gal}}{\langle N_{gal}\rangle} + \frac{1}{1+f_{st}}\frac{N_{st}F(S)}{\langle N_{gal}\rangle} - 1, \tag{37}$$

which can be re-written in terms of the galaxy over-density

$$\delta_o = \frac{\delta_{gal} - f_{st}}{1+f_{st}} + \frac{1}{1+f_{st}}\frac{N_{st}F(S)}{\langle N_{gal}\rangle}. \tag{38}$$

and then the over-density of the stellar contamination, $\delta_{st,S} = \frac{N_{st}F(S)}{\langle N_{st}F(S)\rangle} - 1$

$$\delta_o = \frac{\delta_{gal} - f_{st}}{1+f_{st}} + f_{st}(\delta_{st,S}+1) = \frac{\delta_{gal}}{1+f_{st}} + f_{st}\delta_{st,S} + \frac{f_{st}^2}{1+f_{st}}, \tag{39}$$

after recognizing that $\frac{\langle N_{st,S}F(S)\rangle}{(1+f_{st})\langle N_{gal}\rangle} = f_{st}$.

The measured angular correlation function, $w_o = \langle\delta_o\delta_o\rangle$, is

$$w_o = \frac{w_{gal}}{(1+f_{st})^2} + f_{st}^2 w_{st,S} + \frac{f_{st}^4}{(1+f_{st})^2}, \tag{40}$$

(assuming no true correlation between galaxies and stars). Therefore

$$w_{gal} = (1+f_{st})^2\left(w_o - f_{st}^2 w_{st,S} - \frac{f_{st}^4}{(1+f_{st})^2}\right) \tag{41}$$

The above requires no approximations, but is limited to the specific case of contamination. One must be able to estimate $w_{st,S}$, but we expect this can be approximated as w_{st} in many cases. Further, estimating the stellar contamination as a function of survey conditions can be done empirically and using forward-modeling tools (the potential exists when using similar tools as, e.g., Leistedt et al. (2015); Suchyta et al. (2015)).

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

AFFILIATIONS

[1] *Institut de Ciències de l'Espai, IEEC-CSIC, Campus UAB, Carrer de Can Magrans, s/n, 08193 Bellaterra, Barcelona, Spain*
[2] *Institut de Física d'Altes Energies, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Barcelona, Spain*
[3] *Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[4] *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*
[5] *Department of Astronomy, University of Illinois, 1002 W. Green Street, Urbana, IL 61801, USA*
[6] *Kavli Institute for Cosmology Cambridge and Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[7] *Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*
[8] *Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[9] *Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[10] *National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*
[11] *Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro,*

RJ - 20921-400, Brazil
[12]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[13]*Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland*
[14]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[15]*Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK*
[16]*Instituto de Física Teórica and ICTP SAIFR, Universidade Estadual Paulista, Rua Dr. Bento T. Ferraz 271, Sao Paulo, SP 01140-070, Brazil*
[17]*Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA*
[18]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[19]*Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile*
[20]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[21]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[22]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[23]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[24]*Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching, Germany*
[25]*Faculty of Physics, Ludwig-Maximilians University, Scheinerstr. 1, 81679 Munich, Germany*
[26]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*
[27]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[28]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[29]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany*
[30]*Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany*
[31]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[32]*Australian Astronomical Observatory, North Ryde, NSW 2113, Australia*
[33]*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*
[34]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, CEP 05314-970, São Paulo, SP, Brazil*
[35]*Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*
[36]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, UK*
[37]*Instituto de Física, UFRGS, Caixa Postal 15051, Porto Alegre, RS - 91501-970, Brazil*
[38]*Department of Physics, University of Illinois, 1110 W. Green St., Urbana, IL 61801, USA*
[39]*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA*
[40]*Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA*
[41]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*
[42]*Centre for Theoretical Cosmology, DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*
[43]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa*