Galaxy clustering with photometric surveys using PDF redshift information

Jacobo Asorey[1]*, Matias Carrasco Kind[2,3], Ignacio Sevilla-Noarbe[2,4], Robert J. Brunner[1,2,3], Jon J. Thaler[1,2]

[1] *Department of Physics, University of Illinois, Urbana, IL 61801 USA*
[2] *Department of Astronomy, University of Illinois, Urbana, IL 61801 USA*
[3] *National Center for Supercomputing Applications, Urbana, IL 61801 USA*
[4] *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*

11 December 2015

ABSTRACT

Photometric surveys produce large-area maps of the galaxy distribution, but with less accurate redshift information than is obtained from spectroscopic methods. Modern photometric redshift (photo-z) algorithms use galaxy magnitudes, or colors, that are obtained through multi-band imaging to produce a probability density function (PDF) for each galaxy in the map. We used simulated data to study the effect of using different photo-z estimators to assign galaxies to redshift bins in order to compare their effects on clustering and galaxy bias measurements. We found that if we use the entire PDF, rather than a single-point (mean or mode) estimate, the deviations are less biased, especially when using narrow redshift bins. When the redshift bin widths are $\Delta z = 0.1$, the use of the entire PDF reduces the typical measurement bias from 5%, when using single point estimates, to 3%.

Key words: galaxy clustering; angular correlations; photometric redshift surveys

1 INTRODUCTION

The analysis of the three-dimensional distribution of galaxies has become one of the major probes used to understand the history of the Universe and the growth of matter perturbations. Spectroscopic surveys such as WiggleZ[1] (Drinkwater et al. 2010), BOSS[2] (Dawson et al. 2013) and VVDS[3] (Le Fèvre et al. 2004) have obtained precise maps of this distribution and many studies have increased our knowledge about the expansion history of the Universe and the growth of structures. However, targeting galaxies and obtaining spectra is a slow and costly process; therefore, past and current spectroscopic surveys have been limited to relatively low redshift and a reduced number of galaxies.

Multi-band imaging of wide areas of the sky is complementary to spectroscopic surveys. These photometric surveys, such as the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS)[4], Dark Energy Survey (DES[5], Dark Energy Survey Collaboration 2005) and the Large Synoptic Survey Telescope (LSST[6], Ivezić et al. 2008), enable lower accuracy redshift estimation from the colours of millions of galaxies without being affected by spectroscopic selection effects.

Photometric redshifts (photo-z) are estimated by using multi-band photometry as inputs to one or more different techniques that map galaxy photometric properties into a redshift. These techniques can broadly be classified into two categories; the first is known as template-based methods (Benitez 2000; Ilbert et al. 2006), in which a set of calibrated galaxy spectral energy distributions (SEDs) is fit to the photometric data to find the one that best represents the observed fluxes. The second category use a spectroscopic training set and machine learning algorithms, such as artificial neural networks (Collister & Lahav 2004), boosted

decision trees (Gerdes et al. 2010), or prediction trees and random forests (Carrasco Kind & Brunner 2013), to generate a photo-z PDF estimate.

Probability density functions of various astronomical measurements have been used in cosmological analyses, for example, luminosity functions (Sheth 2007), weak lensing (Mandelbaum et al. (2008)), cluster identification (van Breukelen & Clewley 2009), and the real-space clustering of quasars (Myers, White & Ball 2009). However, a systematic analysis of the use of photometric redshift PDFs in galaxy clustering has not been performed, mostly due to the lack of reliable PDF estimation and its computational cost.

In this work, we study how the angular clustering of galaxies depends on the chosen photo-z estimate and on the photo-z bin width, by using realistic simulations to compare clustering measurements based on photometric and spectroscopic redshifts. We also show how to include the full probability density when estimating the angular correlation functions, and we study the impact of using different photo-z PDF statistics (e.g., mean, mode, and median) on estimating the galaxy bias. In Section 2 we describe the methodology followed in the paper. We describe the angular clustering results and the fitting to galaxy bias in Section 3, and we discuss these results in Section 4. Finally, we summarize the main conclusions in Section 5.

2 METHODOLOGY

The standard approach to analyse galaxy clustering in photometric surveys beings with subdividing the catalogue into sub-samples selected by 'top-hat' photo-z redshift bins. The photo-z value used to determine if a galaxy is in one or another bin is usually a point estimate of the photo-z PDF. In this paper, we quantify how angular clustering analyses are affected by the choice of the specific photometric redshift estimator, including one that uses the full PDF information instead of photo-z point estimates. We address this by measuring the clustering of a subsample of the DES-BCC Aardvark simulation mock galaxy catalogue (Busha et al. 2013). We compared the clustering measurements given by different photo-z estimators and when considering different redshift bin widths. As a specific test, we quantify these differences by fitting theoretical galaxy correlation functions in order to estimate the linear galaxy bias, (Kaiser 1984), as a metric to evaluate which photo-z estimator is more reliable.

2.1 Simulation Data

The mock galaxy catalogue considered here is the Aardvark v1.0 catalogue from the Blind Cosmology Challenge (BCC) simulations, developed for the DES. The catalogue is created from three ΛCDM N-body dark matter simulations, with sizes of 1050 Mpc/h, 2600 Mpc/h, 4000 Mpc/h and 1400^3, 2048^3 and 2048^3 particles, respectively. They were created using Gadget-2 (Springel 2005) and initial conditions given by CAMB (Lewis et al. 2000) and 2LPT (Crocce, Pueblas & Scoccimarro 2006). The algorithm that populates the dark matter halos with galaxies, ADDGALS (Busha et

al. 2013), follows a prescription based on SubHalo Abundance Matching (SHAM) techniques, (Conroy, Wechsler & Kravtsov 2006; Busha et al. 2013; Reddick et al. 2013). The final catalogue is complete down to $r < 25$ and covers 1/4 of the sky. Galaxy properties such as colour or luminosity are assigned by matching a spectroscopic training sample from the SDSS DR6 value added catalogue (Blanton et al. 2005) at low redshift. This training is extrapolated to higher redshift matching the colour distribution to SDSS DR8 (Aihara et al. 2011) and DEEP2 (Newman et al. 2013) photometric data. Then, the output catalogue includes DES colors and errors for each galaxy of the catalogue. These catalogues have been compared with real data by the DES collaboration, (Chang at al. 2015; Leistedt et al. 2015).

The full BCC-Aardvark-v1.0c catalogue covers $10,313$ square degrees to the full DES depth, and includes a total of 1.36×10^9 galaxies. The simulated catalogue is stored in files according to Healpix[7] (Górski et al. 2005) pixels of $n_{side} = 8$. We chose a contiguous area of the simulation by using 24 pixels, which corresponds to an area of about $1,200$ square degrees on the sphere in order to have a significant sampling of small scales. For our study, we have selected the galaxy sample according to a magnitude limited cut of $g < 24$. This cut corresponds to a selection in the g-band of signal-to-noise greater than 20 in the simulation, which incorporates the DES observed photometric error. The total number of galaxies in the catalogue after applying the magnitude cut is around 30 million galaxies.

2.2 Photometric redshift code

We have used the publicly available code TPZ[8], (Carrasco Kind & Brunner 2013) to estimate the galaxy redshift probability distributions. TPZ is a parallel code that estimates photo-z PDFs using prediction trees and random forests. A prediction tree is constructed by splitting the data in recursive branches until a convergence criterion is reached. In order to construct more robust PDFs, the code uses the random forest technique in which N_T bootstrap samples of the training set are created and prediction trees are generated for the N_T samples. In order to include the measurement errors, e.g., magnitude errors, N_R training samples are created by perturbing the training set according to the errors of the measurement variables. Finally, the PDF of each galaxy in the sample is created by combining the prediction trees. TPZ was one of the algorithms used in the DES Science Verification Data (Sánchez et al. 2014), and produced one of the best performances for that dataset.

We have considered 10^5 galaxies as a training set, up to the full depth; and we, therefore, use less than 1% of the available data for this purpose. The training set galaxies were confined to a region of 54 square degrees.

As defined in Carrasco Kind & Brunner 2013, the concentration of individual galaxy PDFs, $p(z)$, is output by TPZ as a PDF concentration parameter called zConf. This parameter is defined as the integrated probability between

---

[7] http://healpix.jpl.nasa.gov/
[8] http://lcdm.astro.illinois.edu/code/mlz.html

$\bar{z} \pm \sigma_{TPZ}(1 + z_{phz})$ and measures the narrowness of the PDF. In this case, we selected $\sigma_{TPZ} = 0.075$, which is similar to the $1 - \sigma$ confidence interval of the PDFs. We can select different quality cuts by using this parameter, which is related to the BPZ ODDS parameter (Benitez 2000).

TPZ is a particular method of the MLZ framework. MLZ incorporates a Bayesian combination of techniques that estimate photometric redshift PDFs, including both template based methods and unsupervised machine learning algorithms(Carrasco Kind & Brunner 2014b), and also enables an efficient storage of the PDFs by using a sparse representation basis(Carrasco Kind & Brunner 2014a). For simplicity, we only used TPZ for the photometric redshifts in this paper, which is justified by the excellent results produced by TPZ on the DES Science Verification Data (Sánchez et al. 2014), and by the fact that we want to study the impact of using photo-z PDFs in clustering as produced by a single technique (to simplify the resulting analysis). TPZ has been used, together with other codes listed in Sánchez et al. 2014, in several DES Science Verification Data studies (Bonnett et al. 2015; Crocce et al. 2015; Giannantonio et al. 2015; Dark Energy Survey Collaboration 2015).

2.3 Survey configuration: Photo-z binning

Galaxy clustering analysis in photometric surveys are usually done by measuring angular correlations of galaxy samples selected in different redshift bins. We divide the full redshift range, which in this paper we restrict to the range $0.2 < z < 1.4$, into N_z redshift bins of width Δz in order to reduce the extent of the projection of radial information for 2D clustering analysis. As shown in Asorey et al. (2012) and Eriksen & Gaztañaga (2015), the optimal photometric redshift bins are given by shells of about twice the size of the photometric redshift standard deviation. In this paper, we consider different configurations: $\Delta z = \{0.1, 0.15, 0.2, 0.3\}$ in order to study the evolution of photometric clustering with bin width. The true redshift distribution of galaxies, n(z), is shown in Figure 1, together with the redshift distribution obtained by stacking TPZ PDFs. We also show (red) the n(z) of the spectroscopic training set. For this paper, we were only interested in a patch of the sky that covers $1,200$ square degrees, which allows us to measure the small scale clustering and to study how it depends on the different photo-z statistical quantities.

In Figure 2, we present the evolution with redshift of the photometric redshift error, σ, for the BCC sample of galaxies, with $g < 24$, given by:

$$\sigma^2 = \int p(z)(z - \bar{z})^2 \qquad (1)$$

where $\bar{z}$ is the mean redshift, defined in Equation (2) below, and we compare it with the different redshift bin widths that we have considered in this paper. Although the optimal choice would be $\Delta z = 0.15$ or 0.2, we also consider extreme cases with $\Delta z = 0.1$ or 0.3 to extend the analysis of the dependence of photo-z clustering on this quantity. In Figure 3, we show the normalized dispersion between the true redshift and the mean redshift.
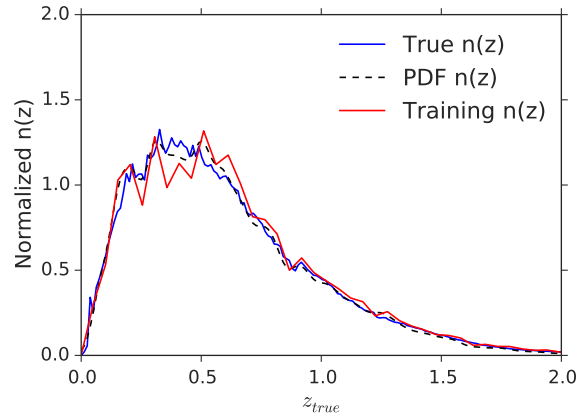

Figure 1. In blue, the galaxy redshift distribution for 31 million galaxies with $g < 24$ in the BCC Aadvark 1.0 catalogue in the selected region of $1,200$ square degrees. The dashed black line shows the result of stacking the individual PDFs of all the galaxies in the sample. In red, we show the true n(z) for the galaxies in the training set used in the photo-z machine learning algorithm.
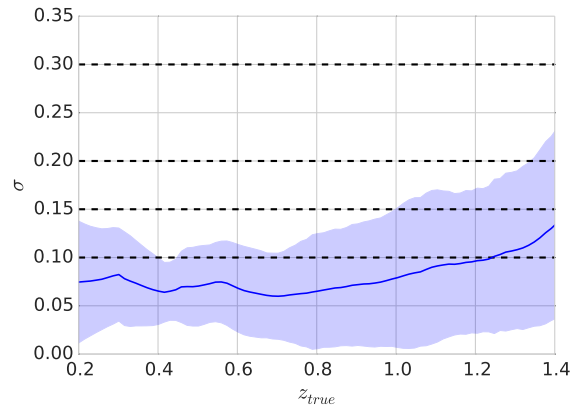

Figure 2. Dependence of the mean photo-z error, given by the mean value of σ in Equation (1) for galaxies in small true redshift bins, on the true redshift. Standard deviations are given by the square root of the variance of the photo-z errors in each bin. Dashed lines correspond to the width of the different bin configurations treated in this paper, in order to compare the bin widths with the photo-z dispersion for the sample considered.

2.4 Photo-z estimators per galaxy

2.4.1 Single point statistics

Once we have computed photo-z PDFs with TPZ, we estimate single statistical summary quantities. In this study, we focus on the mode redshift, $\hat{z}$ and the mean redshift, $\bar{z}$.

We define the mean as the first moment of the PDF, $p(z)$:

$$\bar{z} = \int z p(z) dz \qquad (2)$$

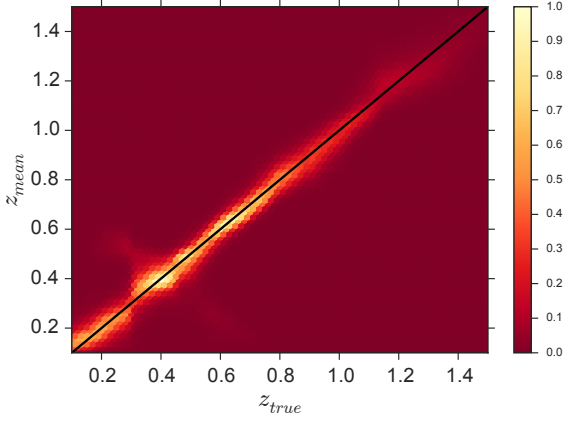The mode redshift is the redshift with highest probability

Figure 3. Normalized dispersion of mean photometric redshifts with respect to the true redshift. The color code corresponds to the relative number of galaxies with respect to the 1:1 relation (black line) between true redshift and mean photometric redshift.

Figure 4. Example of the definitions of the different photometric redshift estimators, where the mode redshift is shown by the vertical red dotted line and the mean redshift by the black dashed line. The blue region corresponds to the part of the PDF that is between the photometric bin $0.5 < z < 0.8$ (true redshift is z=0.603), which is shown as a hatched area. The PDF weight defined in eq. (4) would be the fraction of the total area below the continuous line that is contained in the blue region.

in the PDF, $p(z)$:

$$p(\hat{z}) = p_{max} \qquad (3)$$

Another summarization or single point estimate that we consider in the paper is the Monte Carlo sampling redshift, z_{MC} (Wittman 2009). The Monte Carlo photo-z is the redshift that corresponds to the value of the cumulative distribution function given by a random number between (0,1].We also evaluated the median redshift, but we decided not to include it in the final results as it is similar to the other point estimates, thus for clarity we decided to reduce the analysis to the chosen single point estimates. We created a catalogue for each redshift bin considered by selecting the galaxies with the single point estimate in the range covered by the redshift bin.

2.4.2   Photo-z weights

The proposed technique to incorporate the PDF information in our analysis consists on doing number counts in redshift bins according to a weight for each galaxy in each radial shell, where the weight is given by the probability that the galaxy lies in the corresponding redshift bin. Because the TPZ PDF output is discretized, we define the galaxy weight as:

$$f_z = \sum_i p(z_i) dz \qquad (4)$$

where we only consider redshifts z_i that belong to the redshift bin in consideration. According to this definition, a galaxy may have weights in different redshift bins, where the total weight of the galaxy in the whole redshift space is $f_{tot} = \sum_{j=1}^{N_z} f_{z_j} = 1$. We measured the galaxy clustering by using the weights for the galaxy counts. The case that involves the photo-z single point estimates (mean, mode) is equivalent to setting the weight to $f_z = 1$ for all galaxies selected in the corresponding redshift bin.

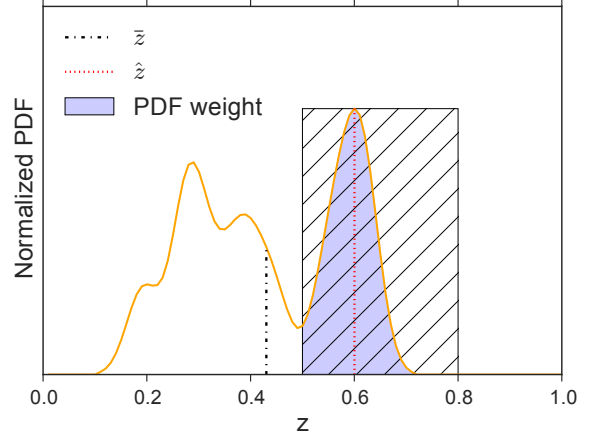We defined threshold cuts, $p_{threshold}$, in a similar way

as in Mandelbaum et al. (2008), as the process of determining if a galaxy lies within a redshift bin or not when using weights. Thus, a galaxy α in redshift bin j would only be incorporated if

$$f_z > p_{threshold} \qquad (5)$$

When PDFs are broad and contain multiple peaks, we might be introducing noise in each redshift bin from galaxies that are not in the bin but have a non-negligible weight. This can be addressed by applying the threshold cut.

In Figure 4, we present a graphic example of the different photometric redshift estimators that we use. We intentionally selected a PDF with a most frequent (mode) redshift, given by eq. (3) within the photometric bin $0.5 < z < 0.8$ but where the mean, defined in eq. (2) is in a different redshift bin. In blue, we show the portion of the PDF between $0.5 < z < 0.8$ that corresponds to the weight of the galaxy in that redshift bin. Of course the PDF displayed in Figure 4 is an extreme case. For this particular PDF, zConf=0.37, while near z=0.6 the mean PDF quality parameter is zConf~ 0.95. In appendix A, we discuss the statistical properties of the mean and the mode and the overall quality of the photometric sample PDFs .

With photo-z PDFs we can easily obtain the photometric sample redshift distribution, $n(z)$, in each bin by stacking all the individual $p(z)$ of the selected galaxies.

$$N(z)_{pdf} = \sum p_i(z) f_{z_i} \qquad (6)$$

In this paper, we have considered this definition as the default estimation of $n(z)$ by setting $f_z = 1$ for the galaxies selected according to single point statistics in a redshift bin. We can also determine the true $n(z)$ measured by the distribution of the true redshifts of this simulated sample.

Throughout this paper, when we refer to true values, we are considering the latter definition.

2.5 Two point angular correlation function estimators

2.5.1 Pixel based estimator

We computed the angular correlations by using pixel maps of the galaxy density field. These maps are created by using `Healpix` for each redshift bin and for each photometric redshift estimator, with $n_{side} = 1024$, corresponding to a minimum angular resolution of 0.06 degrees. For the definition of the estimator, see Scranton et al. (2002), Crocce, Cabré, & Gaztañaga (2011), and Wang, Brunner & Dolence (2013). The angular correlation, is:

$$w(\theta) = \frac{1}{N_{pairs}} \sum_{ij} \delta_i \delta_j \qquad (7)$$

where N_{pairs} is the number of pixel pairs at an angle θ. We defined the density contrast as $\delta_i = (n_i - \bar{n})/\bar{n}$ where n_i is the number of counts in pixel i and $\bar{n}$ the mean number density of galaxies. When selecting galaxies in terms of the PDFs, the total number of counts in every pixel is the sum of the weights of all the galaxies in the pixel, i.e.: $n_i = \sum_{gal \in i} f_z$ and $\bar{n} = \sum_i \sum_{gal \in i} f_z$.

In order to include errors on our measurements, we considered jackknife samples, dividing the survey area into N_{JK} regions, each about 3 square degrees. The covariance matrix, therefore, is given by:

$$C_{\theta_i,\theta_j} = \frac{N_{JK}-1}{N_{JK}} \sum_{k=1}^{N_{JK}} (w_k(\theta_i)-w(\theta_i))(w_k(\theta_j)-w(\theta_j)) \quad (8)$$

which is the same definition used in Scranton et al. (2002); Norberg et al. (2009); Wang, Brunner & Dolence (2013). For the galaxy bias fitting, we adopt the mixed approach used in Crocce et al. 2015, where the correlation matrix between diagonal elements and off-diagonal elements of the covariance matrix is calculated by using theoretical angular power spectra that are rescaled by the variances given by the jackknife errors in order to determine the covariance matrix of the angular correlations.

2.5.2 Direct pair counting estimator

An alternative method to measure angular correlations consists of using pair counts. In order to estimate the angular correlation functions, we used the Landy-Szalay estimator, (Landy & Szalay 1993),

$$w(\theta) = \frac{DD - 2DR + RR}{RR} \qquad (9)$$

where DD is the number of galaxy-galaxy pairs, DR is the galaxy-random pairs and RR the random random pairs within θ and $\theta + \delta\theta$. Random catalogues are created by throwing points in the survey footprint following a uniform density. These are appropriately normalized to the total number of counts. When counting pairs, each galaxy was

weighted according to Equation (4). We computed the point-to-point angular correlation functions by using the publicly available tree code[9], explained and used in Dolence & Brunner (2008); Wang, Brunner & Dolence (2013). We compare in section 3 the point-to-point estimator with the pixel based estimator.

2.6 Theoretical modeling

The angular auto-correlation within a given redshift bin is given by:

$$w(\theta) = \int dr_1 \phi(r_1) \int dr_2 \phi(r_2) \xi(r_1, r_2, \theta) \qquad (10)$$

where the spatial correlation function $\xi(r_1, r_2, \theta)$ encodes the 3D information of the density field that we are projecting. The window functions, ϕ, are a combination of the galaxy redshift distribution, $n(z)$, the galaxy bias, $b(z)$, and the linear growth rate of structure, $D(z)$, in such a way that $\phi(z(r)) = n(z)b(z)D(z)$, where we assumed the linear local bias model (Kaiser 1984):

$$\delta_g = b_g \delta \qquad (11)$$

We define one bias parameter per redshift bin.

We used CAMB (Lewis et al. 2000) to obtain the linear power spectrum with Halofit (Smith et al. 2003; Takahashi et al. 2012) in order to include non-linearities at small scales. We Fourier transform this angular power spectrum in order to compute the 3D angular correlations required by Equation (10). For this paper, we considered a flat ΛCDM model driven by the simulation cosmological parameters when computing the theoretical correlation function. We also included linear redshift space distortions as a series of multipoles following (Kaiser 1987; Hamilton 1992).

3 RESULTS

3.1 Comparison between direct pair counting and pixel-based estimators

As shown in Wang, Brunner & Dolence (2013), pixel-based and point-to-point pair count methods yield similar results for 2-point angular clustering. However, this previous work only considered unweighted pair counts in any given bin, i.e., weights of $f_z = 1$. In Figure 5, we extend this previous result to compare the results for both pair count and pixel-based methods when considering galaxy weights in the redshift bin $1.0 < z < 1.2$ and a threshold $p_{threshold} = 0.1$. As shown in this figure, over the angular range $0.1 < \theta < 1.0$, we find a good agreement. The number of jackknife regions is different in the two cases, however, being $N_{JK} = 32$ for the point-to-point case and $N_{JK} = 384$ when considering the pixel-based estimator. Given this agreement, we opted to use the computationally simpler pixel-based estimator in the rest of the paper.
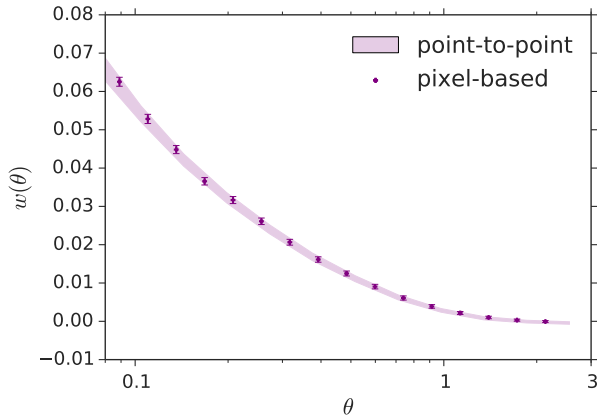
[9] http://lcdm.astro.illinois.edu/code/tpacf.html

Figure 5. A comparison, demonstrating good agreement, between the pixel based clustering measurement (purple points) and the point-to-point clustering measurement (purple shadow include measurements within the error bars) for the photometric redshift bin $1.0 < z < 1.2$ when considering galaxy weights and a threshold on the weights of $f_z > 0.1$.

3.2 Clustering amplitude

We now focus our analysis on the relative amplitude between the angular clustering signal using different photo-z selection criteria and the true redshift clustering. This allows us to directly study how the different statistical representations of photometric redshift change the measurement signal.

In the analysis, we divide the redshift range $0.2 < z < 1.4$ into different numbers of bins in order to consider different bin configurations. As a result, a comparison between individual redshift bins for different bin configurations may consider different redshift regions for different configuration. For this reason, we present clustering results for bins with different widths in symmetrical manner about a given central redshift. As discussed previously, we restrict our analysis to four different bin widths. In Figure 6, we present the ratios of the photometric redshift clustering with respect to the true redshift clustering, for different statistical estimators and a redshift bin centered at $z = 1$.

In the left panels, we show the results for a broad bin of $\Delta z = 0.3$ in the redshift range $0.85 < z < 1.15$. As expected, for all algorithms, the amplitude of the clustering of the photometric sample is smaller than when using the true redshifts, since the errors on a photometric redshift estimate will suppress this inherent clustering. Notice that the clustering measurements when using the mode, $\hat{z}$, and mean, $\bar{z}$, are similar. Any small differences are due to the fact that each selection produces a different $n(z)$ when individual PDFs are not symmetric, like the one shown in Figure 4.

In the bottom panels we show the clustering ratios with respect to the true clustering when using PDF weights with different probability thresholds. To aid in the comparison, we also present the ratio for $\hat{z}$ with the red dashed line. This allows us to both clean our sample, as if using a quality cut, and sample more narrow redshift bins. The thresholds considered are $p_{threshold} = 0.0, 0.1$. The signal depends on

the cut on the selection of weighted galaxies. The stronger the cut, the cleaner the sample and the clustering amplitude increases, as well as the intrinsic $n(z)$. But this may also bias our results as we may be changing the average galaxy types in the sample, as discussed in Martí et al. (2014). Here we considered a low threshold that is non-negligible to compare with the full PDF case.

We consider narrower bin configurations in order to test what happens as we approach the intrinsic photo-z dispersion error, summarized in Figure 2. We see in the case when $\Delta z = 0.2$ that the amplitude of the true redshift clustering increases because the signal is proportional to $n(z)^2$, as explained by Equation (10), which in the top hat case means that it is inversely proportional to $(\Delta z)^2$. The bin considered in this case for galaxy selection is $0.9 < z < 1.1$. We see that the differences between the mode and the mean are bigger than in the previous case with $\Delta z = 0.3$. This may be a result of the fact that when we consider bins much bigger than the intrinsic separations, the differences between photo-z single statistic estimators are smaller. As we extend the comparison to smaller widths, $\Delta z = 0.15$ ($0.925 < z < 1.075$), the ratio between the angular clustering signal for the photometric samples and the sample with true redshift becomes smaller. This is in agreement with the trend we saw before, as $w(\theta) \propto (1/\Delta z)^2$ for true redshift clustering, while the photo-z dispersion keeps the corresponding signal diluted in the radial direction.

The case with $\Delta z = 0.1$ ($0.95 < z < 1.05$) shows the same trend than the previous cases with bigger bin widths. The ratio of the photometric redshift signals to the true redshift signal continues to decrease. The results with mean and mode estimators tend to converge as the bin width approaches the photo-z dispersion error.

We found that the evolution of the clustering amplitude with bin width evolves differently for the different estimators. We show in Figure 7 the clustering amplitude evolution at $\theta = 0.1$ for the angular correlations shown in Figure 6. For photometric redshift estimators, the clustering signal increases by about 50% when increasing the numbers of bins by a factor of 4. Therefore, increasing the number of bins is not an efficient process for any photo-z estimator, at least from the point of view of a clustering measurement, especially for bins with a width smaller than $\Delta z = 0.15$, which is twice the mean photo-z error, as shown in Figure 2.

3.3 Bias measurement

3.3.1 PDF redshift distributions

We next evaluate how the selection of galaxies in radial shells when using different photo-z statistics affects the information on the linear galaxy bias b_g, as defined in Equation (11). Fitting the galaxy bias, or any cosmological parameter, can help us to calibrate the effect of different photo-z statistics. We only used angular auto-correlation functions and one galaxy bias parameter per redshift bin. The total redshift range considered is $0.2 < z < 1.4$, and all bins have the same width for each configuration. The angular range considered was set to cover the co-moving coordinates range $10h^{-1}Mpc < r < 60h^{-1}Mpc$, which corresponds to different
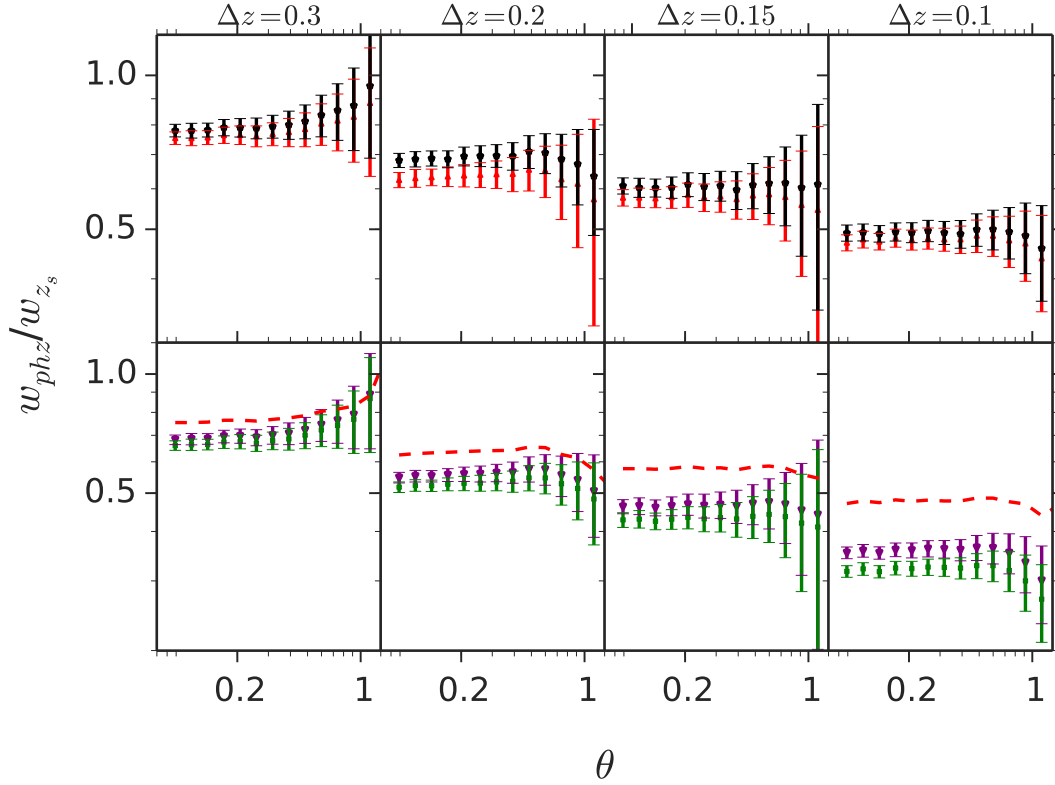
Figure 6. This comparison between angular clustering measurements that use photometric redshifts and true redshifts. We consider different redshift estimators for redshift bins centered at $z = 1$ but with different widths. (Top) The results when computed by using the mode (red triangles) and the mean (black stars). (Bottom) The results when using PDF weights, with threshold $p_{threshold} = 0$ (green squares), and when considering thresholds to the weights of $p_{threshold} = 0.1$ (purple triangles). In both panels, the different columns correspond to different bin widths. In the bottom row, we show the results for $\hat{z}$ with a red dashed line in order to compare with the upper panels. The ratios with respect to the true redshift results decrease with the bin width and are different depending on the considered photo-z statistics used.
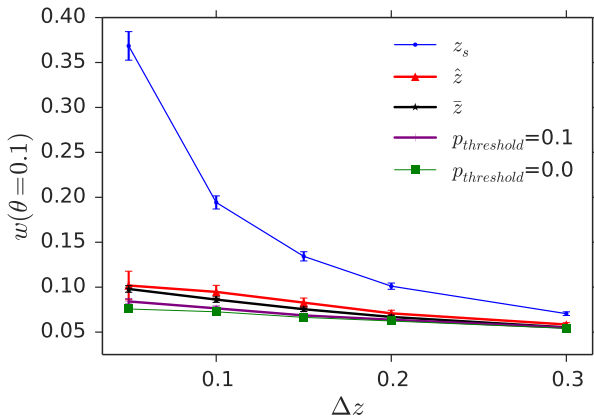

Figure 7. The evolution of the clustering amplitude at $\theta = 0.1$ for a redshift bin centered at $z = 0.5$ with a given bin width, Δz. We show the evolution for the true redshift and for different photometric redshift statistics in order to quantify when the clustering signal saturates with bin width.

angular ranges in each redshift bin. The minimum scale was selected by testing at which scale the linear growth model for the spatial correlation departs from a non-linear model. Notice that this is a conservative cut when compared with the cuts used in (Crocce et al. 2015). This corresponds to $\theta_{min} = 0.8$ degrees at the lowest redshift and $\theta_{min} = 0.19$ degrees at highest redshift bin.

The comparison between the different photo-z selection methods is done by comparing each galaxy bias measurement with the true result, which is determined by using the true redshift distribution of the selected galaxies in order to do a fair comparison. First, we show in Figure 8 the galaxy bias measurement made by using different photometric redshift statistics and the bias measurement done by selecting galaxies according to the spectroscopic redshift. Notice that the spectroscopic sample in each bin is different than the photometric samples considered, but since this is accounted for in the bias measurement, we can study how the photo-z statistic measurements compare with the spectroscopic one.

In panel *a* of Figure 8, we show the evolution of galaxy bias for the broad $\Delta z = 0.3$ bin configuration. We only show results for the true redshift, z_s, the mode redshift, $\hat{z}$, the
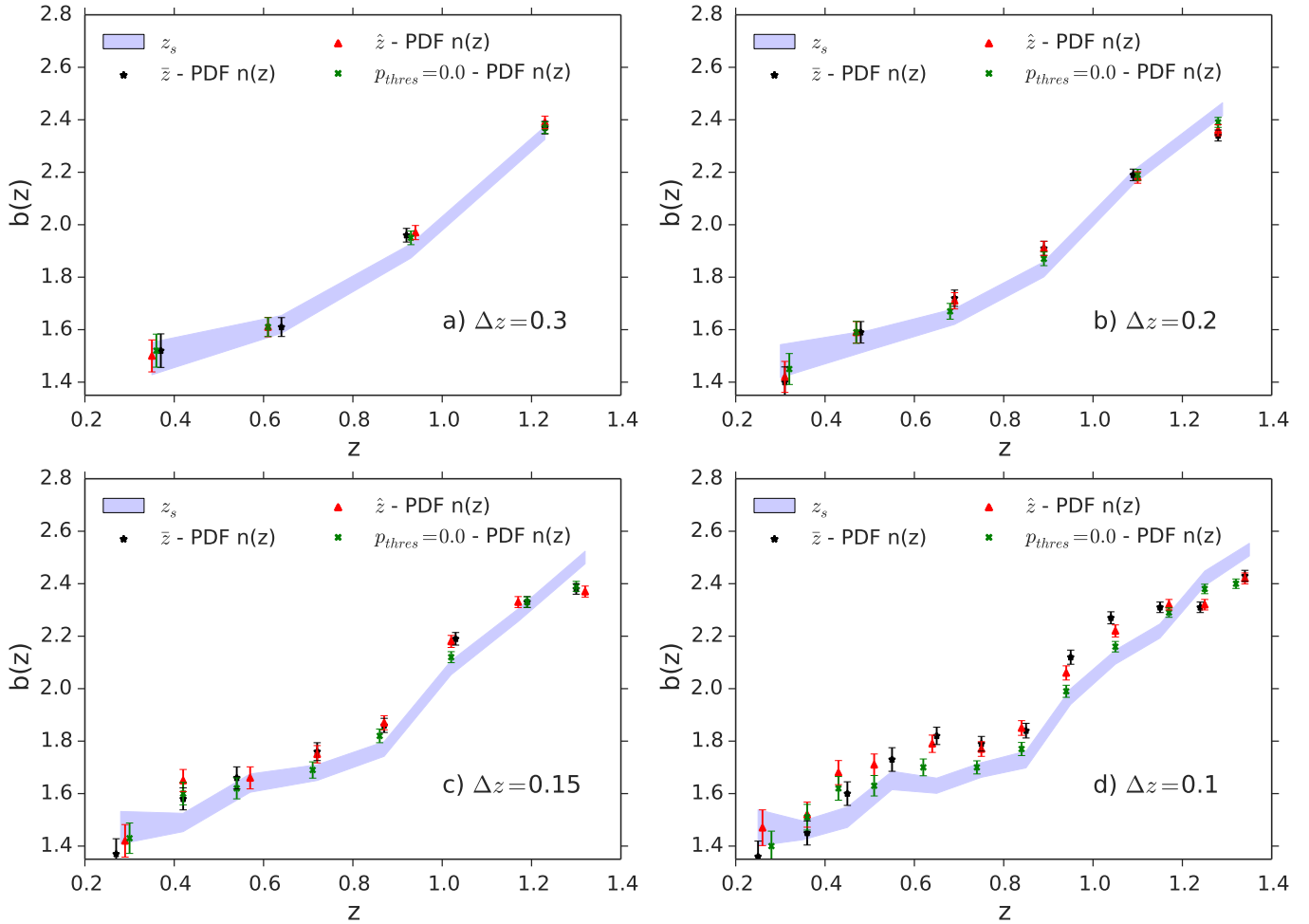
Figure 8. Bias evolution for different redshift bin configurations: The evolution with redshift of the linear galaxy bias when dividing the full sample into redshift bins. Results are shown both for the different bin widths and the different photometric redshift statistics, described in section 2.4: spectroscopic redshift results (blue shadow), mean photo-z (black star), mode photo-z (red triangle), and photo-z PDF weights (green cross). In each bin, the spectroscopic sample is different than the corresponding photo-z sample, thus we cannot directly compare them. The x-axis position is given by the mean redshift in each bin according to the $n(z)$, which is given by stacking the photo-z PDFs.

mean redshift, $\bar{z}$, and the PDF weighted samples for clarity. We do not show here the results when applying photometric redshift quality cuts. The measurements are similar and the slightly different values for the different estimators are within the statistical error bars. This is reasonable as we are considering a broad redshift bins in this panel, and the differences between different photometric samples redshift distributions are thus small.

The same trend is observed when considering $N_z = 6$ redshift bins, as shown in panel b on Figure 8. This case corresponds to bins with $\Delta z = 0.2$ width, which is larger than twice the photo-z dispersion, and, therefore, photometric redshift effects are still not the biggest issue. The evolution of linear galaxy bias with redshift resembles the plot shown in Figure 17 of Crocce et al. 2015 for a MICE-CATv2.0 (Carretero et al. 2015) sample, as the ratio $b_g(z = 1.1)/b_g(z = 0.3)$ is similar for both simulations.

We show in the lower left plot in Figure 8 the measurement of the bias for the different redshift estimators when considering $N_z = 8$ redshift bins. In this case, we begin to observe bigger differences between the case when using photometric redshifts and the case when the bias was obtained by using spectroscopic redshifts, especially when compared to the previous cases that used larger bin widths.

Finally, we show in the bottom right panel of Figure 8 the bias evolution when using 12 bins of width $\Delta z = 0.1$. The differences between the photo-z galaxy bias results and the spectroscopic bias measurement are larger than for the previous cases with broader bin configurations. The closest result to the spectroscopic results value of the bias is obtained when using the full PDF information ($p_{threshold} = 0$), especially at intermediate redshifts.

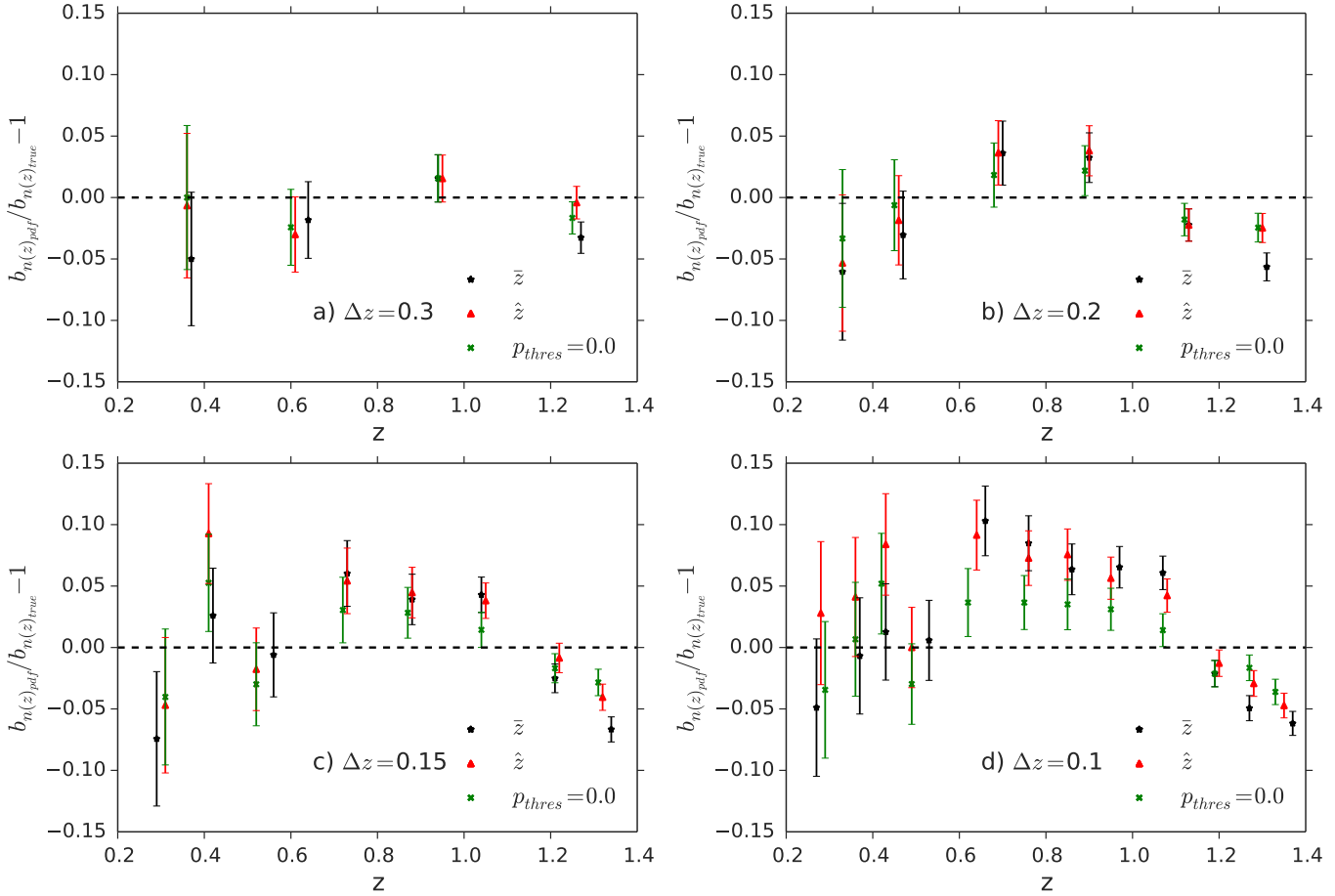We show in Figure 9 the measurement bias between the method that uses PDF stacking to estimate n(z) and the

Figure 9. Relative bias between PDF stacking and the true redshift distribution: The relative differences in galaxy bias measurements when measuring the redshift distribution by using PDF stacking with respect to the bias measurements given by using the true redshift distribution. This computation is done for the three photo-z selection methods in each of the four redshift bin widths considered in this paper. Notice that for each method we are using the sample photometric sample in each bin, but we are fitting the galaxy bias by using different redshift distributions.

true value, given by the n(z) measured directly from the true redshifts from the simulation of the photometric sample. We show in the top left panel the relative differences when considering $\Delta z = 0.3$ for the three methods, finding small deviations with respect to the true results with minor differences between the different selection techniques. These observed differences exist because the PDF stacking technique is not perfectly reconstructing the true n(z) of the population sample in the tomographic bins. In this case, the differences are small because the bin width is broad and photo-z systematics in the n(z) are smaller.

When we decrease the bin width to $\Delta z = 0.2$, the differences grow, as shown in the top right panel of Figure 9. The three methods are still producing similar results, and, because the bins are still too broad, the relative bias is zero, within the error bars. When the configuration changes to bins with widths $\Delta z = 0.15$, differences start to become more apparent and the PDF weighting method begins to differ from the single point estimate estimators. For the narrowest bin width configuration considered, $\Delta z = 0.1$, the

differences at intermediate redshifts are larger than 5% for single point estimators, whereas for PDF weighted galaxy samples, these differences are around 3%. We include a table in Appendix B that presents all galaxy bias measurements and the relative differences with the true results.

In order to summarize and quantify these results, we show in Figure 10 the mean value of the mean absolute deviation between each selection method and the true result for each bin width. We found that for the largest bin widths, the differences are around 1% and are similar for the three photo-z selection statistics: mean, mode, and pdf weighting. However, for the narrower bins, the deviation when we consider summary statistics is around 5%, while it is 3% when using the photo-z PDF galaxy weighting method.

3.4 Reducing the redshift bin catalogue size

Using full PDF information in galaxy clustering produces less biased measurements than point estimate photo-z methods, but it also increases the amount of information that we
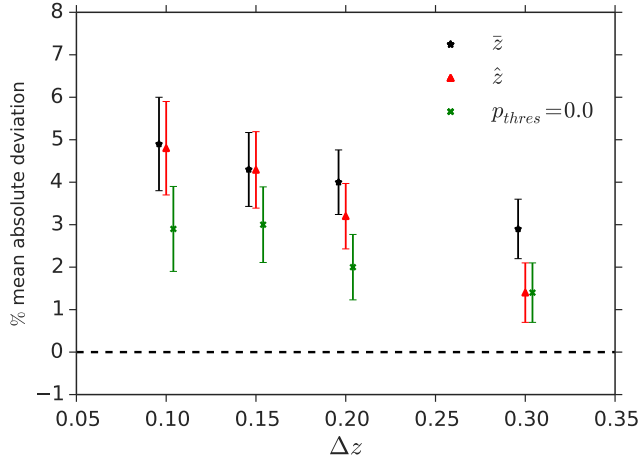
Figure 10. Evolution with redshift bin width of the percentile mean absolute deviation with the true redshift result for the different photometric redshift statistics. We artificially shift the x-axis values in order to more clearly show the results from the different measures. Notice the accumulated measurement bias when using photo-z redshifts, which is smaller when using PDF weights in the clustering measurements, especially for the narrower bin configurations.

Figure 11. Bias evolution in Monte Carlo sampling redshift shells of width $\Delta z = 0.15$: Bias measurement in 8 redshift bins defined by the Monte Carlo sampling redshifts (orange). We consider the cases with $n(z)$ estimated by using PDF stacking and the true $n(z)$. We also compare the galaxy bias fitting with the PDF weights results from figure 8 (green). The standard performance of Monte Carlo redshifts is similar to the results given by PDF weights.

have to store in each redshift bin galaxy sample. We also studied how a Monte Carlo sampling of the PDF, in order to define a point estimate that encloses more of the PDF than the mean or the mode, or applying a threshold cut based on the amount of PDF in each bin compares to the full PDF inclusion method.

3.4.1 Monte Carlo sampling redshift

We extended the previous analysis to include a Monte Carlo sampling redshift, z_{MC}, which assigns a redshift value based on the cumulative distribution function for each galaxy. We make our previous galaxy bias measurement in the different redshift bins according to z_{MC}. In Figure 11 we show the best fit galaxy bias for galaxies selected according to z_{MC} in $N_z = 8$ redshift bins of $\Delta z = 0.15$. We observe that in this case the results are similar to the results given by PDF weights (for example, the results from panel c of Figure 8), both when using PDF sampling or the true redshift distributions. This is expected, as we are using the probabilistic information to determine the Monte Carlo sampling redshifts.

3.4.2 Quality cuts

The effect of sparse PDFs with multiple peaks can introduce significant noise into our PDF weighting scheme. Although it is not the main interest of this paper, we considered a case in which we applied a threshold cut $f_z > (p_{threshold} = 0.1)$ in order to select galaxies in the different bins. The effect is a combination of a quality cut and a cut on galaxies that are not in the bin but whose tails are inside the bin, which
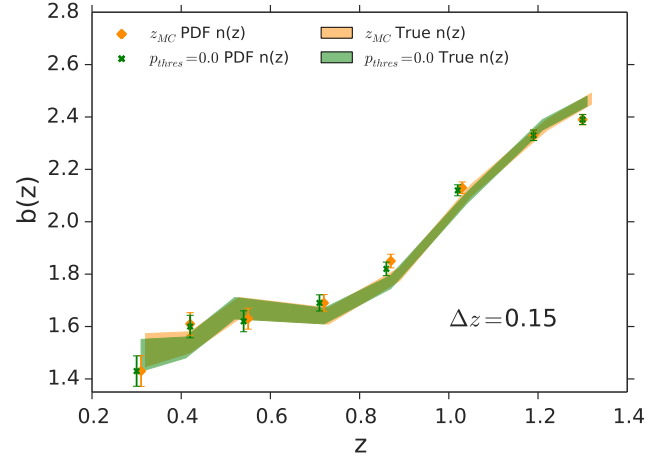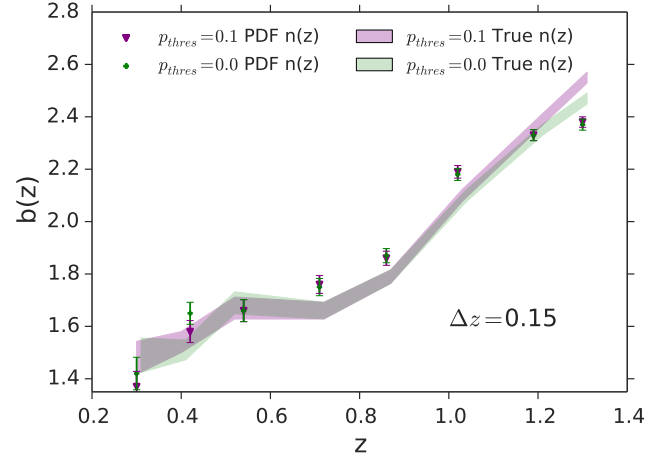

Figure 12. PDF threshold cuts: A comparison between the effect of applying a threshold cut on the selection process with photo-z PDF weights for the configuration in a redshift bin width of $\Delta z = 0.15$. Both samples produce similar galaxy bias measurements.

produces bigger maps. In Figure 12, we present a comparison between a photo-z sample selected according to full photo-z PDFs for a configuration with bin width $\Delta z = 0.15$ and a a sample selected by applying a threshold to the photo-z PDF weights of $f_z > 0.1$. We found that the results are similar, supporting the idea of applying threshold cuts to reduce the size of the galaxy density in each pixel in the map, although cuts to a sample have to be applied carefully in order to avoid introducing selection biases to the sample, see, e.g., Martí et al. (2014). We observe this effect in Figure 12, as
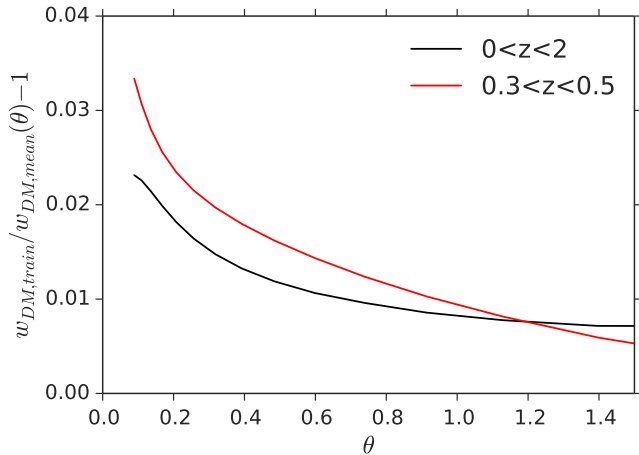
Figure 13. Training set sample variance: Relative differences between the amplitude of the theoretical dark matter angular correlations when using the spectroscopic n(z) of the training set and the mean n(z) of ten samples with the same number of spectroscopic objects as the training set but distributed across the catalogue area. We considered both the full redshift range $0 < z < 2$ (black) and a redshift bin $0.3 < z < 0.5$ (red). We find a relative bias in the angular range $0.1 < \theta < 1$ smaller than 2%, which propagates to an error smaller than 1% on the galaxy bias, which is lower than the observed galaxy bias described in this paper.

the true results for both samples are not exactly equivalent. A detailed study of using quality cuts from PDF information is outside the scope of this paper.

3.5 Systematics

3.5.1 Training set sample variance

Since we use one galaxy sample from one particular pixel of the simulation for our photo-z training set, we wanted to demonstrate that the choice of this one pixel did not bias our results. As a result, we compared the spectroscopic n(z) for the galaxies in our training sample with the mean spectroscopic n(z) from ten randomly selected galaxy samples from the entire area, to demonstrate that our results were not dependent on the choice of a particular pixel. We are not, in this paper, exploring the more traditional concept of photometric redshift 'sample variance' as discussed, for example, in Cunha et al. 2012.

We show in Figure 13 the relative difference between theoretical angular correlations, computed by using Equation (10), when we used the training set n(z) (red line in Figure 1) and the mean n(z) of ten different random samples extracted from the catalogue with the same number of galaxies than the training set, which is similar to the blue solid line in Figure 1. We found relative differences smaller than 2% over the redshift range for all angles, $0.1 < \theta < 1$. This implies a relative difference smaller than 1% for the galaxy bias, which is lower than the differences observed for our different photo-z statistics.
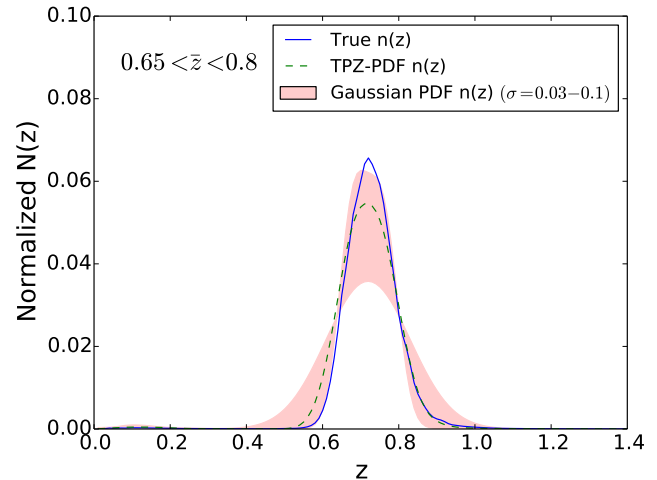

Figure 14. Redshift distribution reconstruction: A comparison between the PDF stacking n(z) (dashed green) and the true redshift distribution (blue) obtained by selecting galaxies with mean photometric redshift in the redshift bin $0.65 < z < 0.8$. The red region covers the space between the n(z) obtained by stacking gaussian PDFs for the galaxy sample with standard deviation within $\sigma_{gauss} = 0.03 - 0.1$. The differences between redshift distributions are contained within the photo-z error.

3.5.2 True redshift distribution reconstruction

As observed in section 3.3, the main difference between the photometric redshift and the true redshift galaxy bias measurement is a result of the failure to recover the true redshift distribution, $n(z)$. As an example, in Figure 14, we show the difference between the true redshift distribution (blue line) and the PDF stacking PDF (green dashed line) for galaxies selected with mean redshift within $0.65 < z < 0.8$. The red shadowed region shows the range between n(z) created by stacking gaussian PDFs with standard deviations in the range $\sigma_{gauss} = 0.03$ to $\sigma_{gauss} = 0.1$. We see that the difference between the measured n(z) is within the accuracy of the photo-z catalogue, shown in Figure 2. We explore this result in more detail in Appendix C, where we look at the differences between the n(z) obtained from stacking the photo-z PDF of galaxies selected in redshift shells according to their mean redshift and the true distribution of the same sample when using both different bin configurations and different redshift ranges.

4 DISCUSSION

In this paper, we have studied how the angular galaxy clustering obtained from photometric populations depends on the different statistical estimators used to assign galaxies to specific redshift bins. The primary estimators that we have considered are the mean and the mode of a galaxy's photo-z PDF. We found differences between the different estimators, in part, since they produce different galaxy samples in each top hat photometric redshift bin. As a result, the clustering signal is different when using either the mean or the mode.

We also included the full PDF information in our clustering analysis by weighting each galaxy according to the integrated probability that the galaxy actually resided within each redshift bin. This clustering signal is smaller than the clustering signal from single point estimates samples. If we apply a threshold cut of $p_{threshold} = 0.1$, the clustering amplitude increases. This is explained by the fact that when we consider a larger threshold the corresponding $n(z)$ is narrower than when considering all galaxies with non negligible weights in a redshift bin. However, we also may be sampling different type of galaxies, since we are only selecting galaxies with higher probability to lie within the bin.

We extended the comparison between the different photometric redshift statistical representations to a cosmological parameter estimation analysis by measuring the linear galaxy bias in different redshift bins. We find that, in general, the photo-z estimators produce similar results, especially when considering broad bins. We find that there is a relative bias with respect to the true galaxy bias results, since the PDF stacking redshift distributions in each bin differ from the true redshift distributions. For narrow bins, the selection method given by PDF weights produces less biased differences with respect to the true results. The mean deviation for a bin configuration with width $\Delta z = 0.1$ is 3% when using PDF weights, while it is 5% when using summary or single point estimate statistics. Thus, the use of photo-z PDF weights to select galaxies in tomographic redshift bins in order to measure the galaxy clustering in a photometric survey produces more robust results than using single point estimates. We can use the methodology presented in this paper to calibrate the effect of assigning galaxies to photo-z bins to ensure that the model parameters from simulations mimic the real data catalogues. This also applies to other photo-z methods that estimate PDFs (Sánchez et al. 2014; Bonnett et al. 2015; Leistedt et al. 2015) as they will have similar behaviours

Creating maps with PDF weights involves much larger data sets than catalogues of galaxies selected only by redshift. One way to reduce the amount of data is to apply a cautious PDF quality cut by using a threshold when considering PDF weights. We found similar results to the full PDF results, although any cut on a sample has to be carefully tested. Another way to reduce the size of the catalogues, while still retaining a certain level of the PDF information, is by using Monte Carlo Sampling point estimates. We found that Monte Carlo Sampling estimators produce similar results to our PDF weight results.

The effect of choosing different photometric redshift training samples from the simulation on the calculation of the galaxy bias measurement is smaller than 1%, which is lower than the effects due to the different photo-z statistics used in this paper. Likewise, the differences between stacking photo-z PDFs to compute the redshift distribution and the true redshift distribution are also within the photo-z errors.

5   CONCLUSIONS

With photometric surveys, we can accumulate much larger galaxy samples in less time than with spectroscopic surveys.

However, the lack of true redshifts restricts the quality of any radial information on such a survey, as photometric redshift are produced from multi-band imaging.

Therefore, we need to set a statistical definition of a photometric redshift in order to identify which tomographic redshift bin contains a given galaxy. The search for an optimal definition is the main goal of this paper. The core analysis of this paper consisted of defining a new photo-z selection method that includes the full photo-z PDF information by weighting each galaxy in the redshift bin with the probability that the galaxy lies in that bin, and to compare this result with methods based on single statistical estimates such as the mean or the mode of the photo-z PDF.

We found, using mock galaxy catalogues and a machine learning photo-z code, that if we use single point statistics, like the mode or the mean, there is an offset on the galaxy bias measurements. These bias measurements are obtained either by measuring photometric redshift distributions by stacking the individual photo-z PDFs or from the true redshift distribution of the same galaxies. This shift must be taken into account when considering similar large scale structure analyses that leverage galaxies drawn from photometric surveys. This corrective effect can be estimated by applying a similar method to measure the offset in the determination of the cosmological measurement of interest by using simulations in similar conditions to the expected photometric data. In our case, we used the galaxy bias as the metric to test different photo-z statistics, and we found that, for single point statistics, the cumulative deviation is a 5% for a bin configuration with width $\Delta z = 0.1$.

Our results are closer to the ground truth if we weight the contribution of each galaxy to a photo-z bin according to the amount of their photo-z PDF in each redshift bin. This approach, on the other hand, produces a difference of 3% in the $\Delta z = 0.1$. Therefore, and especially for narrow photometric top hat bins, PDF weighting is more optimal than simply using summary statistic photometric redshifts.

REFERENCES

Aihara H, Allende Prieto C., An D., Anderson S. F. , Aubourg E. ,Balbinot E., Beers T. C., Berlind A. A., Bickerton S. J., Bizyaev D. , et al., 2011, ApJS **193**, 29
Asorey J., Crocce M, Gaztañaga E., Lewis A., 2012, MNRAS **427**, 1891
Benitez N., 2000, ApJ **536**, 571
Blanton M. R., Schlegel D. J., Strauss M. A., et al., 2005, AJ, **129**, 2562
Bonnett C. et al., 2015, arXiv e-print:1507.05909
Busha M. T., Wechsler R. H., Becker M. R., Erickson B., Evrard A. E., 2013, AAS Meeting Abstract, **221,341**
Carrasco Kind M., Brunner R. J., 2013, MNRAS **432**, 1483
Carrasco Kind M., Brunner R. J., 2014, MNRAS **441**, 3550
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
Chang C., Busha M. T., Wechsler R. H., Refregier A.; Amara A., Rykoff E., Becker M. R., Bruderer C., 2015, ApJ, **801**, 73
Collister A. A., Lahav. O., 2004, PASP **116**, 345
Conroy C., Wechsler R. H., Kravtsov A. V, 2006, ApJ **647**, 201
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS **373**, 369
Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS **414**, 329
Crocce M. et al., 2015, arXiv e-print:1507.05360
Cunha C. E., Huterer D., Busha M. T., Wechsler R. H., 2012, MNRAS, **423**, 909
Dawson et al., 2013, ApJ, bf 145, id. 10
Dark Energy Survey Collaboration, 2005, arXiv e-print:astro-ph/0510346
Dark Energy Survey Collaboration, 2015, arXiv e-print:1507.05552
Dolence, J. C., Brunner, R. J., 2008, Proceedings of the 9th LCI International Conference on High-Performance Clustered Computing
Drinkwater M. J. et al., 2010, MNRAS, bf 401, 1429
Eriksen M., Gaztañaga E., 2015, arXiv e-print:1502.03972]
Gerdes D. W., Sypniewski A. J., McKay T. A., Hao J., Weis M. R., Wechsler R. H., Busha M. T., 2010, ApJ, **715**, 823
Giannantonio T., et al., 2015, arXiv e-print:1507.05551
Górski, K. M.; Hivon, E.; Banday, A. J.; Wandelt, B. D.; Hansen, F. K.; Reinecke, M.; Bartelmann, M., 2005, ApJ **622**, 759
Hamilton A. J .S, 1992, ApJ, **385**, L5
Ilbert O., 2006, AA **457** , 841
Ivezić Z. et al., 2008, arXiv e-print:0805.2366
Kaiser N., 1984, MNRAS, **284**, L9
Kaiser N., 1987, MNRAS, **227**, 1
Landy S. D., Szalay A. S,, 1993, ApJ **412**, 64
Le Fèvre O. et al., 2004, A&A, bf 417, 839
Leistedt B., el al. 2015, arXiv e-print:1507.05647
Lewis A., Challinor A., Lasenby A., 2000, ApJ **538**, 473
Mandelbaum R., et al., 2008, MNRAS, **386**, 781
Martí P., Miquel R., Bauer A., Gaztañaga E., 2014, MNRAS, **437**, 3490
Myers A. D., White M., Ball N. M., 2009, MNRAS, **399**, 2279
Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, MNRAS, **396**, 19
Newman J. A., Cooper M. C., Davis M. , Faber S. M., Coil A. L., Guhathakurta P. , Koo D. C. , Phillips A. C., Conroy C., Dutton A. A., et al., 2013, ApJS, **208**, 5
Reddick R. M. ,Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, **771**, 30
Sánchez C. et al., 2014, MNRAS **445**, 1482
Scranton R., et al., 2002, ApJ **579**, 48
Sheth R. K., 2007, ApJ **378**, 709

Springel V., 2005, MNRAS **364**, 1105
Smith R. E. et al., 2003, ApJ **341**, 1311
Takahashi. R., Sato M., Nishimichi T., Taruya A. Oguri M., 2012, ApJ, **761**, 152
van Breukelen C., Clewley L., O. 2009, MNRAS **395**, 1845
Wang Y. Brunner R. J., Dolence J.C., 2013, MNRAS, **432**, 1961
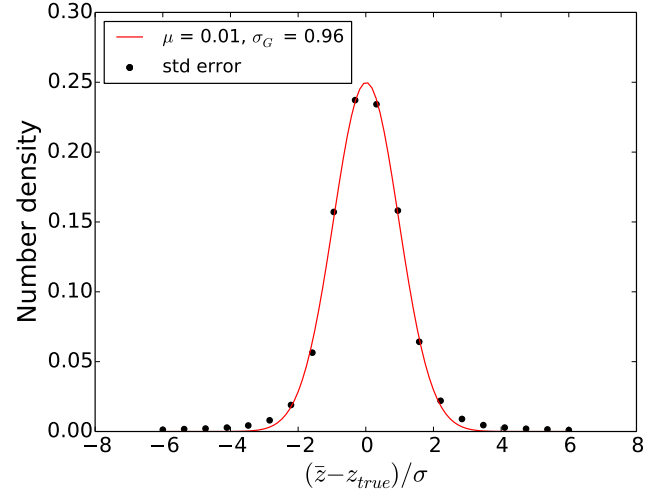Wittman D., 2009, ApJL, **700**, L174

Figure A1. Photometric standardized error for the mean: The photometric standardized error computed from the mean of each individual galaxy's photo-z PDF compared to the best fit Gaussian, shown with the solid red line (mean μ and error σ_G).

	$1-\sigma$	$2-\sigma$
$\bar{z}$	70%	93%
$\hat{z}$	69%	90%

Table A1. Proportion of photometric redshifts inside $1-\sigma$ and $2-\sigma$ level confidence intervals for the mean, $\bar{z}$ and the mode, $\hat{z}$, photometric redshifts. In the ideal case, they are 68% and 95%.

APPENDIX A: ERROR DISTRIBUTION

In order to check the robustness of the galaxy photo-z PDFs that we used in the paper, we estimated the distribution of the photometric standardized error of the photo-z BCC galaxies used in the paper, $(z_{phot} - z_{true}/\sigma)$. The standard deviation, σ, is given by Equation (1). In Figure A1, we show the results using the mean redshifts. We observe that the simple error estimate is close to the unbiased estimate ($\mu = 0$, $\sigma_G = 1$). We can also consider the mode redshift, as shown in Figure A2, where we see that the distribution of the modes tends to be more concentrated than the distribution of the means.

We also tested how photometric redshifts are distributed according to the confidence intervals by estimating the number of galaxies with photometric redshifts inside $1-\sigma$ and $2-\sigma$ levels, which is shown in Table A1. We see that the distribution of mean values in confidence intervals is close to the expected 68% and 95% distributions. When
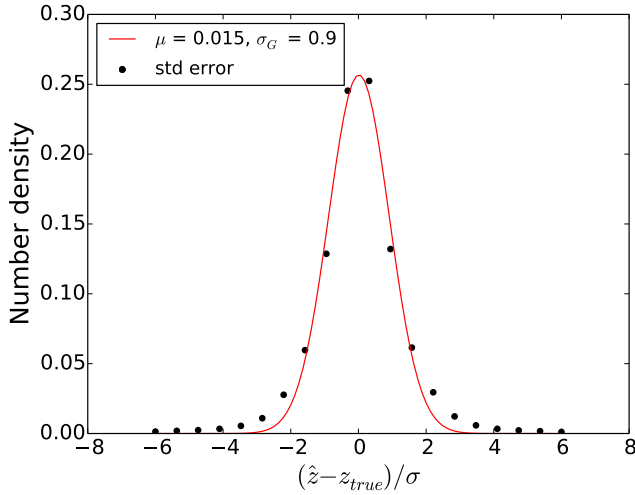
Figure A2. Photometric standardized error for the mode: The photometric standardized error computed from the mode of each individual galaxy's photo-z PDF compared to the best fit Gaussian. As modes are defined by the peak of each PDF, the distribution tends to be more concentrated than the distribution of mean PDF values.

considering the mode, the values are more concentrated as we are considering the peaks of each individual PDF.

APPENDIX B: GALAXY BIAS RESULTS

In table B1, we show the galaxy bias fits for the different bin configurations and the three photometric redshift methods: mean, mode, and PDF, used in this paper to select galaxies in tomographic redshift bins. In each case, we stack the galaxy photo-x PDFs to compute the redshift distribution. We also present the goodness of fit for each fit and the relative difference with the appropriate true measurement.

APPENDIX C: TRUE REDSHIFT DISTRIBUTIONS

As shown in section 3.3, the galaxy bias measurements obtained from n(z) given by PDF stacking are different than the true measurements. This is caused by the difference between the true redshift distribution of the photo-z galaxy sample and the PDF stacking $n(z)$. As an example, we show in Figure C1 the differences between the true redshift $n(z)$ and the PDF stacking $n(z)$ for galaxies selected according to the mean redshift for a given set of top hat redshift bins at low redshift. We note that the tails of the PDF stacking n(z) are longer than the true redshift distribution. This disagreement is expected and has been observed when using template based and machine learning algorithms that incorporate PDFs. We expanded the comparison to intermediate redshift (Figure C2) and high redshift (Figure C3), observing similar differences.

This paper has been typeset from a TeX/ LaTeX file prepared by the author.

	$\bar{z}$			$\hat{z}$			$p(z)$		
	PDF-stacking		Relative difference	PDF-stacking		Relative difference	PDF-stacking		Relative difference
Photo-z bin	Galaxy bias	χ^2	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))
$0.2 < z < 0.5$	1.52 ± 0.064	12.3/7	-0.05 ± 0.057	1.50 ± 0.061	12.1/7	-0.01 ± 0.057	1.52 ± 0.063	11.2/7	0 ± 0.06
$0.5 < z < 0.8$	1.61 ± 0.036	0.99/7	-0.018 ± 0.032	1.61 ± 0.037	0.92/7	-0.03 ± 0.031	1.61 ± 0.036	0.58/7	-0.02 ± 0.03
$0.8 < z < 1.1$	1.96 ± 0.026	7.6/7	0.016 ± 0.020	1.97 ± 0.027	8.3/7	0.015 ± 0.019	1.95 ± 0.026	7.4/7	0.016 ± 0.019
$1.1 < z < 1.4$	2.37 ± 0.024	5.71/7	-0.033 ± 0.015	2.39 ± 0.024	8.9/7	0 ± 0.014	2.37 ± 0.022	7/7	-0.017 ± 0.013
Photo-z bin	Galaxy bias	χ^2	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))
$0.2 < z < 0.4$	1.40 ± 0.058	14.6/7	-0.06 ± 0.055	1.42 ± 0.059	14.2/7	-0.05 ± 0.056	1.45 ± 0.059	12.9/7	-0.03 ± 0.056
$0.4 < z < 0.6$	1.59 ± 0.041	4.1/7	-0.03 ± 0.035	1.59 ± 0.041	4.4/7	-0.02 ± 0.027	1.59 ± 0.042	4.5/7	-0.01 ± 0.037
$0.6 < z < 0.8$	1.72 ± 0.031	2.6/8	0.04 ± 0.026	1.71 ± 0.031	2.8/8	0.04 ± 0.027	1.67 ± 0.03	2.9/8	0.02 ± 0.026
$0.8 < z < 1.0$	1.91 ± 0.027	4/7	0.03 ± 0.021	1.91 ± 0.027	3.7/7	0.04 ± 0.02	1.87 ± 0.026	4.3/7	0.02 ± 0.02
$1.0 < z < 1.2$	2.19 ± 0.022	12/8	-0.02 ± 0.013	2.18 ± 0.022	11.1/8	-0.02 ± 0.014	2.19 ± 0.021	12.9/8	-0.02 ± 0.013
$1.2 < z < 1.4$	2.34 ± 0.021	3.9/8	-0.06 ± 0.011	2.36 ± 0.021	5.5/8	-0.02 ± 0.012	2.39 ± 0.019	6.9/8	-0.02 ± 0.011
Photo-z bin	Galaxy bias	χ^2	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))
$0.2 < z < 0.35$	1.37 ± 0.058	12.5/7	-0.07 ± 0.056	1.42 ± 0.062	12.7/7	-0.05 ± 0.059	1.43 ± 0.058	12.6/7	-0.04 ± 0.055
$0.35 < z < 0.5$	1.58 ± 0.042	8/8	0.026 ± 0.038	1.65 ± 0.042	7.8/8	0.09 ± 0.039	1.60 ± 0.043	6.5/8	-0.05 ± 0.04
$0.5 < z < 0.65$	1.66 ± 0.042	3.6/7	0.01 ± 0.035	1.66 ± 0.042	1.9/7	-0.01 ± 0.035	1.62 ± 0.04	1.2/7	-0.03 ± 0.033
$0.65 < z < 0.8$	1.76 ± 0.034	2.9/7	0.06 ± 0.029	1.75 ± 0.033	2.7/7	0.05 ± 0.28	1.69 ± 0.031	2.7/7	-0.03 ± 0.027
$0.8 < z < 0.95$	1.86 ± 0.027	2.3/7	0.04 ± 0.021	1.87 ± 0.027	2.2/7	0.04 ± 0.021	1.82 ± 0.026	/7	0.028 ± 0.021
$0.95 < z < 1.1$	2.19 ± 0.024	16/8	0.043 ± 0.016	2.18 ± 0.023	16.7/8	0.04 ± 0.016	2.12 ± 0.021	17.2/8	0.014 ± 0.014
$1.1 < z < 1.25$	2.33 ± 0.021	8/7	-0.025 ± 0.013	2.33 ± 0.021	11.4/7	0.01 ± 0.013	2.33 ± 0.02	8.5/7	-0.02 ± 0.011
$1.25 < z < 1.4$	2.38 ± 0.020	3.9/8	-0.07 ± 0.011	2.37 ± 0.021	4.2/8	-0.04 ± 0.011	2.39 ± 0.019	6.8/8	-0.03 ± 0.011
Photo-z bin	Galaxy bias	χ^2	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))	Galaxy bias	χ^2/dof	Comparison True (n(z))
$0.2 < z < 0.3$	1.36 ± 0.059	8.7/7	-0.05 ± 0.059	1.47 ± 0.068	8.1/7	0.028 ± 0.068	1.4 ± 0.057	8.2/7	-0.03 ± 0.056
$0.3 < z < 0.4$	1.45 ± 0.046	6.2/8	-0.01 ± 0.044	1.52 ± 0.048	5.7/8	0.04 ± 0.046	1.51 ± 0.049	/8	0.01 ± 0.046
$0.4 < z < 0.5$	1.60 ± 0.045	7.9/7	0.01 ± 0.04	1.68 ± 0.046	7.3/7	0.08 ± 0.042	1.62 ± 0.045	6.2/7	0.05 ± 0.041
$0.5 < z < 0.6$	1.73 ± 0.045	3.8/7	0.01 ± 0.037	1.71 ± 0.041	1.7/7	0 ± 0.033	1.63 ± 0.039	1.4/7	-0.03 ± 0.033
$0.6 < z < 0.7$	1.82 ± 0.033	2.5/8	0.1 ± 0.028	1.79 ± 0.033	2.7/8	0.09 ± 0.028	1.70 ± 0.032	1.6/8	0.04 ± 0.028
$0.7 < z < 0.8$	1.79 ± 0.028	12.3/8	0.08 ± 0.025	1.77 ± 0.028	13.1/8	0.07 ± 0.024	1.70 ± 0.025	8.7/8	0.04 ± 0.022
$0.8 < z < 0.9$	1.84 ± 0.028	0.85/7	0.06 ± 0.023	1.85 ± 0.028	0.8/7	0.076 ± 0.023	1.77 ± 0.025	2/7	0.04 ± 0.021
$0.9 < z < 1.0$	2.12 ± 0.027	7.3/7	0.07 ± 0.019	2.06 ± 0.027	6.7/7	0.06 ± 0.020	1.99 ± 0.023	9.9/7	0.03 ± 0.017
$1.0 < z < 1.1$	2.27 ± 0.023	12.2/8	0.06 ± 0.015	2.22 ± 0.023	14.2/8	0.04 ± 0.015	2.16 ± 0.02	16.2/7	0.014 ± 0.013
$1.1 < z < 1.2$	2.31 ± 0.02	12.6/7	-0.02 ± 0.012	2.32 ± 0.02	16.7/7	-0.01 ± 0.012	2.29 ± 0.018	9.4/7	-0.02 ± 0.011
$1.2 < z < 1.3$	2.31 ± 0.02	3.5/8	-0.05 ± 0.011	2.32 ± 0.02	4.6/8	0.03 ± 0.011	2.38 ± 0.018	7.2/8	-0.02 ± 0.010
$1.3 < z < 1.4$	2.43 ± 0.021	4.1/8	-0.06 ± 0.011	2.42 ± 0.021	5.1/8	-0.05 ± 0.012	2.40 ± 0.018	6.8/8	-0.04 ± 0.01

Table B1. Galaxy bias measurements for photometric samples selected according to the mean, $\bar{z}$, mode, $\hat{z}$, or PDF weighted galaxies in redshift bins, given different bin configurations.
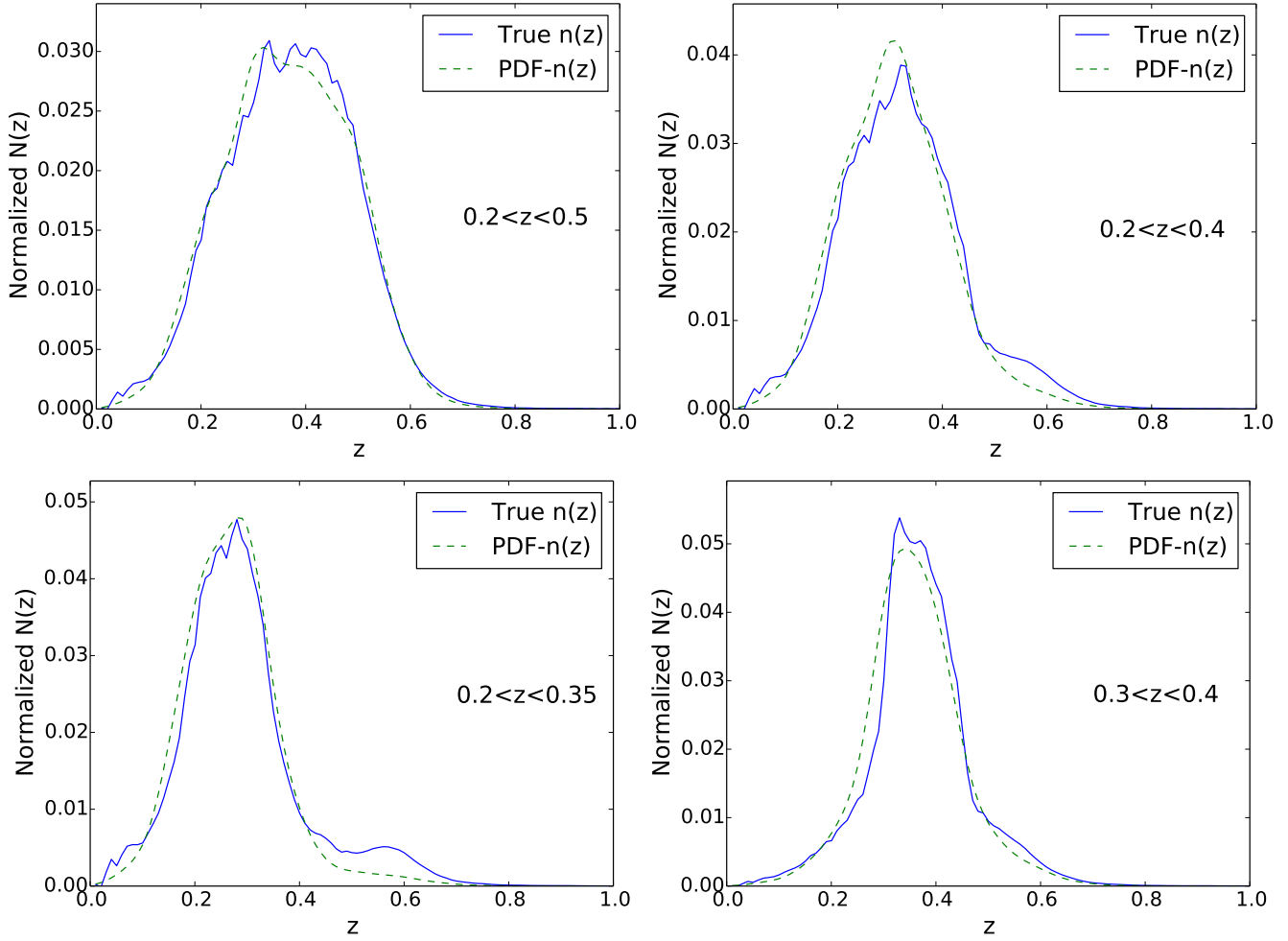
Figure C1. <u>Low redshift</u>: A comparison of the redshift distributions for the true redshift distribution of galaxies selected according to mean photo-z redshift and the PDF stacking redshift distribution for the same galaxies over the lowest redshift range of the true galaxy sample.
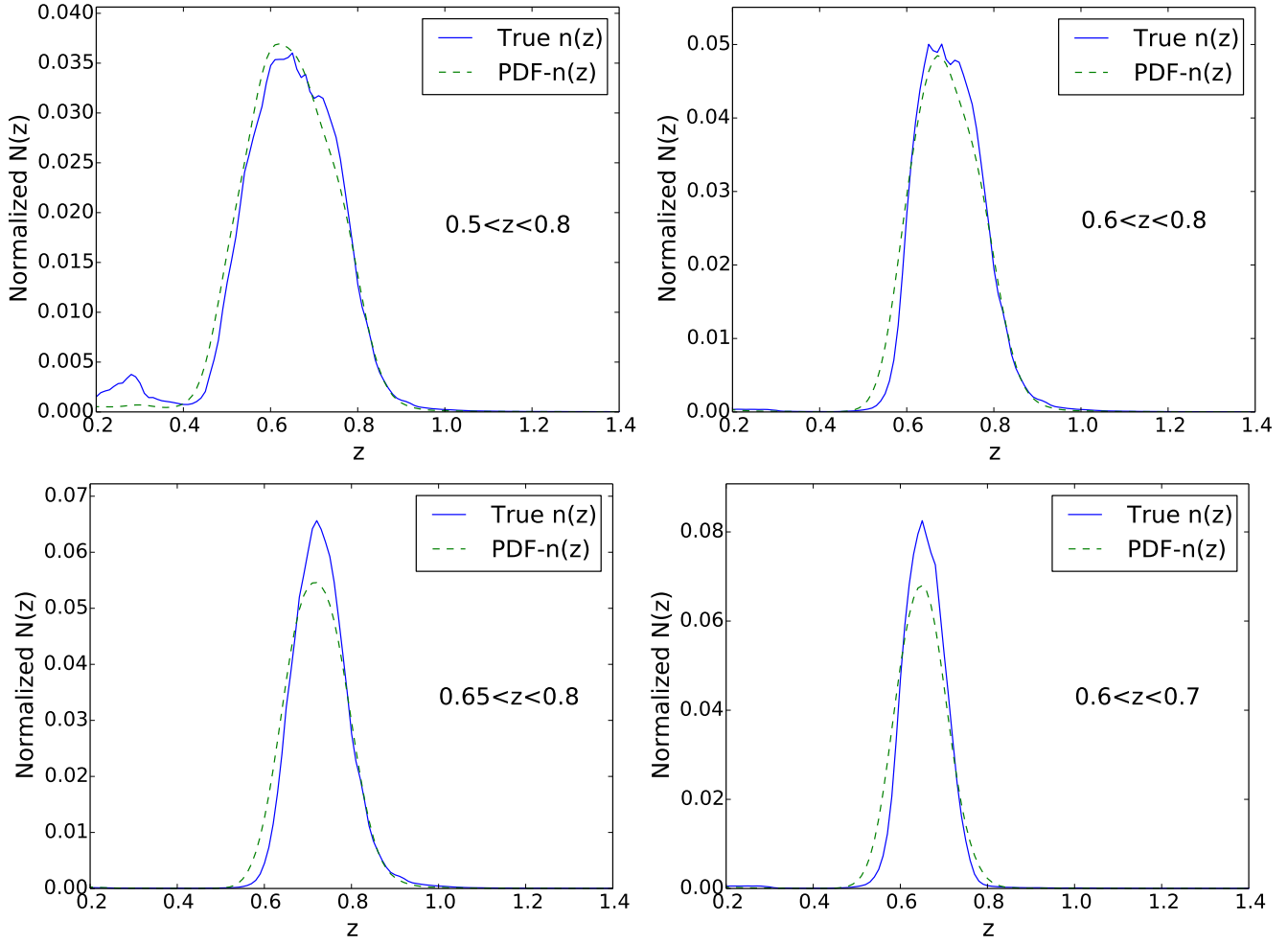
Figure C2. Intermediate redshift: A comparison between the true redshift distribution of galaxies selected according to mean photo-z redshift and the PDF stacking redshift distribution for the same galaxies over the intermediate redshift range of the true galaxy sample.
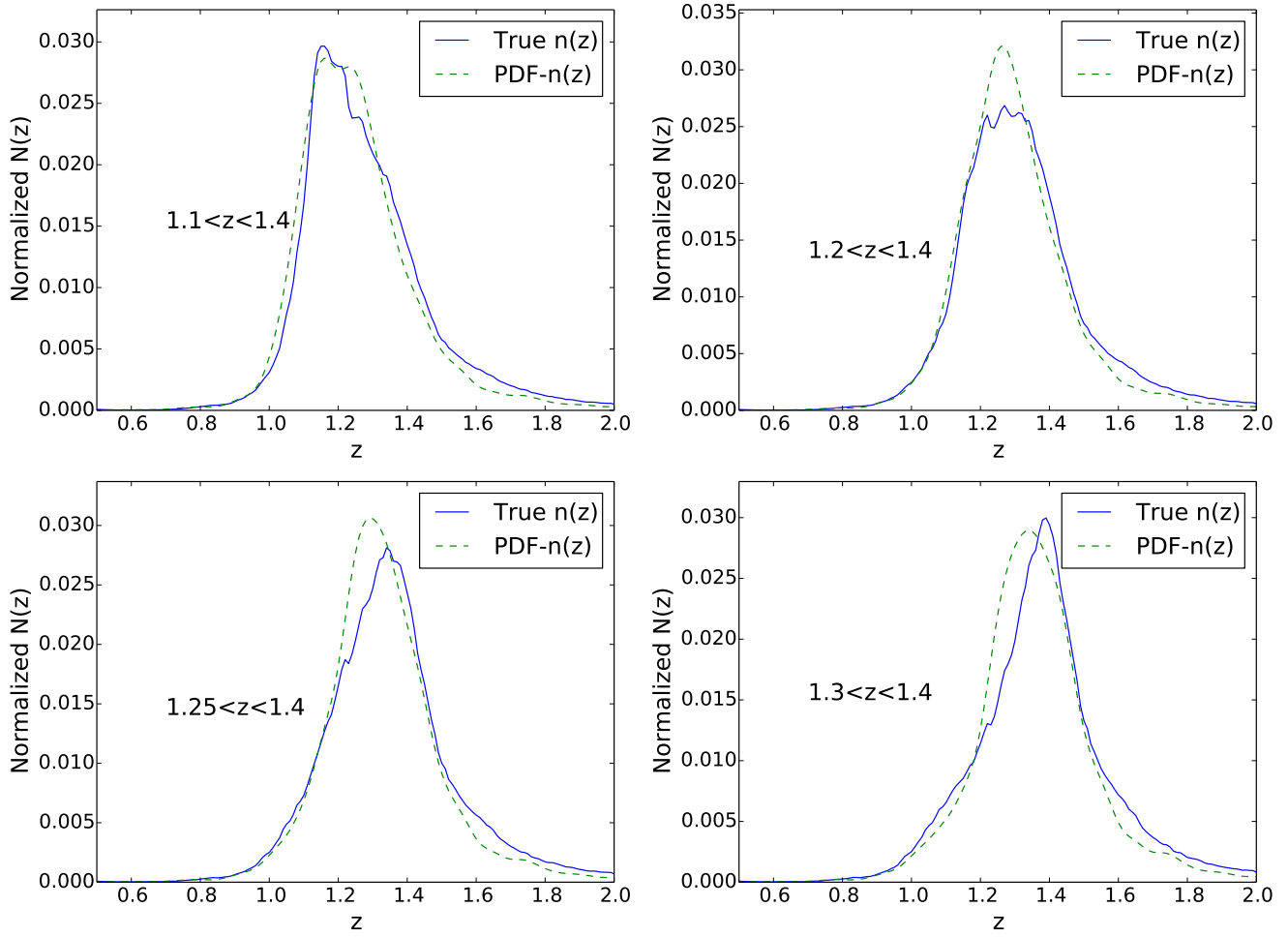
Figure C3. <u>High redshift</u>: A comparison between the true redshift distribution of galaxies selected according to mean photo-z redshift compared with the distribution given by the PDF stacking of the same sample over the highest redshift range of the true galaxy sample.