Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **463,** 3737–3754 (2016)
Advance Access publication 2016 September 8

doi:10.1093/mnras/stw2241

CFHTLenS and RCSLenS: testing photometric redshift distributions using angular cross-correlations with spectroscopic galaxy surveys

A. Choi,[1]★ C. Heymans,[1] C. Blake,[2] H. Hildebrandt,[3] C. A. J. Duncan,[1] T. Erben,[3] R. Nakajima,[3] L. Van Waerbeke[4] and M. Viola[5]

[1]*Scottish Universities Physics Alliance, Institute for Astronomy,University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[2]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[3]*Argelander-Institut für Astronomie, Auf dem Hügel 71, D-53121 Bonn, Germany*
[4]*University of British Columbia, Department of Physics and Astronomy, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*
[5]*Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*

Accepted 2016 September 2. Received 2016 September 2; in original form 2015 December 11

ABSTRACT

We determine the accuracy of galaxy redshift distributions as estimated from photometric redshift probability distributions $p(z)$. Our method utilizes measurements of the angular cross-correlation between photometric galaxies and an overlapping sample of galaxies with spectroscopic redshifts. We describe the redshift leakage from a galaxy photometric redshift bin j into a spectroscopic redshift bin i using the sum of the $p(z)$ for the galaxies residing in bin j. We can then predict the angular cross-correlation between photometric and spectroscopic galaxies due to intrinsic galaxy clustering when $i \neq j$ as a function of the measured angular cross-correlation when $i = j$. We also account for enhanced clustering arising from lensing magnification using a halo model. The comparison of this prediction with the measured signal provides a consistency check on the validity of using the summed $p(z)$ to determine galaxy redshift distributions in cosmological analyses, as advocated by the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS). We present an analysis of the photometric redshifts measured by CFHTLenS, which overlaps the Baryon Oscillation Spectroscopic Survey (BOSS). We also analyse the Red-sequence Cluster Lensing Survey, which overlaps both BOSS and the WiggleZ Dark Energy Survey. We find that the summed $p(z)$ from both surveys are generally biased with respect to the true underlying distributions. If unaccounted for, this bias would lead to errors in cosmological parameter estimation from CFHTLenS by less than ∼4 per cent. For photometric redshift bins which spatially overlap in 3D with our spectroscopic sample, we determine redshift bias corrections which can be used in future cosmological analyses that rely on accurate galaxy redshift distributions.

Key words: gravitational lensing: weak – methods: analytical – techniques: photometric – surveys – galaxies: distances and redshifts.

1 INTRODUCTION

Cosmological parameter estimation often relies on highly accurate knowledge of the underlying 3D spatial distributions of the galaxies used in the analysis. The most direct way to estimate the distributions in the redshift dimension is to measure the redshifts of all galaxies of interest using high-resolution information from spectroscopy, but this is not only costly but potentially incomplete due to the difficulties of measuring secure redshifts for certain populations of galaxies (Cunha et al. 2014; Masters et al. 2015). Photometric redshift estimation provides a lower resolution and less expensive tool for constraining redshift distributions. While fundamentally limited by the available filters, relevant model and assumptions (such as template choice or how representative the training set is), summed redshift probability distribution functions, $p(z)$, provide estimates of the underlying redshift distributions that are less biased than counts of single-point estimates when compared against spectroscopic information (Mandelbaum et al. 2008; Cunha et al. 2009; Wittman 2009; Abrahamse et al. 2011; Nakajima et al. 2012; Sheldon et al. 2012; Benjamin et al. 2013).

Various techniques for calibration of these errors and the distribution of photometric galaxies have been promoted (for a review, see Newman et al. 2015), falling roughly into three categories.

★ E-mail: choi@roe.ac.uk

(i) Direct calibration requires a complete and representative training sample with which to re-weight the photometric galaxies or compare them on an individual basis (Bordoloi et al. 2012; Bonnett et al. 2016). The spectroscopic data sets that are used to characterize photometric redshift scatter and bias are often only complete to a magnitude that is much brighter than the magnitude of a typical galaxy used in cosmological analyses.

(ii) Reconstruction methods utilize the fact that there is an excess probability of pairs of galaxies (relative to a Poisson distribution) that are truly physically correlated in 3D. The clustering information is then used to infer the true underlying redshift distributions (Schneider et al. 2006; Newman 2008; Matthews & Newman 2010; Schulz 2010; McQuinn & White 2013; Ménard et al. 2013; Schmidt et al. 2013; de Putter, Doré & Das 2014; Rahman et al. 2015). This strategy needs a spectroscopic sample to span the full redshift range that can be incomplete in terms of the galaxy properties. Cross-correlations measure the combination of the galaxy bias times the redshift probability distribution, and thus all of these methods require additional constraints on the galaxy bias of the given sample to break the degeneracy. Newman (2008) propose an iterative procedure to account for evolution in the galaxy bias based on the autocorrelations of the spectroscopic and photometric samples, respectively, although the effectiveness of the correction depends on the shape of the redshift distribution as well as the linearity of the galaxy bias evolution (for further discussion, see Newman 2008; Schmidt et al. 2013). In principle, the cross-correlation strategies have a key advantage over standard direct calibration methods as they do not rely as strongly on the completeness of the sample with available spectroscopic redshifts provided that the spectroscopic redshifts cover a 3D space overlapping with the photometric redshifts.

(iii) Verification methods use the cross-clustering signal between galaxies in different redshift bins to indicate the degree of contamination between those redshift bins and test for consistency with other estimates of the redshift distribution such as summed $p(z)$. Erben et al. (2009) and Benjamin et al. (2010, 2013) investigated the photometric–photometric case, and here we extend the formalism to photometric–spectroscopic samples. The advantage of this approach over others is that it can yield constraints on catastrophic outliers even when the spectroscopic sample does not extend over the full redshift range under consideration.

There are three types of photometric redshift error. Two of these, random scatter and systematic bias, move galaxies into adjacent or closely neighbouring redshift bins. A third kind occurs when systematic bias sends galaxies into distant redshift bins and is commonly referred to as a 'catastrophic outlier'. With the angular cross-correlation analyses used in reconstruction and verification methods, a detection of a strong clustering signal between a low-redshift spectroscopic sample and a high-redshift photometric sample is an indication that catastrophic outliers exist in the photometric sample. However, one must also consider astrophysical effects caused by lensing magnification, which most previous works have ignored. Magnification can be thought of as both a contaminant to redshift recovery via cross-correlations (Bernstein & Huterer 2010) and an informative signal in its own right (Scranton et al. 2005; Hildebrandt, van Waerbeke & Erben 2009; Ménard et al. 2010; Morrison et al. 2012; Duncan et al. 2014), containing the imprint of galaxy evolution and cosmological processes. Lensing of light by foreground structures (de)magnifies images of background galaxies. Hence, at a fixed apparent magnitude, the number density behind a massive foreground galaxy will change, which will be seen in an angular cross-correlation signal. A cross-correlation signal can contain contributions from both magnification and catastrophic outliers, thus necessitating a careful investigation of the magnification effects before a clear interpretation can be made about the presence of catastrophic outliers. Moessner & Jain (1998) present the theory behind the lensing and intrinsic contributions to the total observed clustering signal as well as investigate the cosmological dependence.

In this paper, we outline formalism for testing consistency between the estimated redshift probability distributions of a photometric sample and angular cross-correlations between the photometric sample and a spatially overlapping spectroscopic sample (Section 2). Our verification method fully accounts for the unknown galaxy bias, assuming that the average galaxy bias of an outlier population at a given photometric redshift does not significantly deviate from the average galaxy bias of the main population at the same photometric redshift (see Section 2.3 for the details). This is a notable advantage compared with the aforementioned clustering-based reconstruction approaches for which the redshift probability distribution is completely degenerate with galaxy bias (before additional corrections). We model the effects of the magnification component using the halo model. We then apply this test to a $\sim$66 deg^2 region where there are spatially overlapping samples of photometric galaxies imaged by the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS) and galaxies with spectroscopic redshifts from the Baryon Oscillation Spectroscopic Survey (BOSS) survey and a $\sim$200 deg^2 region with overlap between the Red Sequence Cluster Lensing Survey (RCSLenS) and both the WiggleZ and BOSS surveys. The considerable amount of spectroscopic overlap makes CFHTLenS and RCSLenS ideal data sets on which to test cross-correlation techniques. We describe these surveys and the catalogue production pipeline in Section 3. In Section 4, we present the angular cross-correlation measurements and predictions based on our models for the intrinsic and magnification clustering and discuss the level of consistency. We conclude in Section 5. In the appendix, we describe validation tests of our method on mock galaxy catalogues, provide further details of the halo modelling used, and check for systematics in the object catalogues.

For the modelling of the magnification signal only, we assume cosmological parameters from Planck Collaboration et al. (2014), with $\Omega_{\rm m} = 0.315$, $\Omega_\Lambda = 0.685$, $\sigma_8 = 0.829$, $n_{\rm s} = 0.9603$ and $\Omega_{\rm b}h^2 = 0.022\,05$.

2 FORMALISM

In this section, we present the formalism for the angular cross-correlations that are the focus of this work, beginning with a general discussion of galaxy clustering and the physical processes that contribute to the signal. We then relate the redshift probability distributions to the intrinsic angular cross-correlations between spectroscopic and photometric samples and describe how we use the halo model to construct predictions for the angular cross-correlations arising from lensing magnification.

2.1 Angular correlations and magnification

The clustering signal $w(\theta)$ is a two-point angular correlation function, or the excess probability of finding a pair of objects in a solid angle $\mathrm{d}\Omega$ and angular separation θ, such that the probability element is given by

$$\mathrm{d}P = N[1 + w(\theta)]\mathrm{d}\Omega\,, \tag{1}$$

where N is the mean number of galaxies per unit steradian (Peebles 1973). For two different galaxy samples at mean redshifts $\langle z_1 \rangle$ and $\langle z_2 \rangle$,

$$\mathrm{d}P_{1,2} = N_1 N_2 [1 + w_{1,2}(\theta)]\, \mathrm{d}\Omega_1\, \mathrm{d}\Omega_2\,. \quad (2)$$

In practice, we use the Landy–Szalay estimator (Landy & Szalay 1993) for $w_{1,2}(\theta)$ given by

$$w_{1,2}(\theta) = \frac{(D_1 D_2)_\theta}{(R_1 R_2)_\theta} \frac{N_{R,1} N_{R,2}}{N_1 N_2} - \frac{(D_1 R_2)_\theta}{(R_1 R_2)_\theta} \frac{N_{R,2}}{N_1} - \frac{(D_2 R_1)_\theta}{(R_1 R_2)_\theta} \frac{N_{R,1}}{N_2} + 1\,. \quad (3)$$

$(D_1 D_2)_\theta$ is the number of pairs with one galaxy in data sample 1 and the other in data sample 2 as a function of the angular separation θ. Similarly, $(D_1 R_2)_\theta$ is the number of pairs with one galaxy in data sample 1 and the other in a random sample 2, which is constructed to reflect the same selection properties like masks and geometry as the corresponding data sample 2. $(R_1 R_2)_\theta$ is the number of pairs with one galaxy in random sample 1 and the other in random sample 2. It is necessary to use random catalogues that are many times more highly sampled than the data catalogues in order to minimize the noise contributed by including the additional random pair counts. Therefore, each of the terms in equation (3) must be normalized by factors involving N_1, N_2, $N_{R,1}$, $N_{R,2}$ representing the number of galaxies in data sample 1, the number in data sample 2, the number in random sample 1, and the number in random sample 2, respectively.

We can also write $w_{1,2}(\theta)$ in terms of differential number densities

$$w_{1,2}(\theta) = \langle \delta n_1(\phi) \delta n_2(\phi + \theta) \rangle_\phi\,, \quad (4)$$

where

$$\delta n_i(\phi) \equiv \frac{n_i(\phi) - \bar{n}_i}{\bar{n}_i} = \delta n_i^g(\phi) + \delta n_i^\mu(\phi)\,, \quad (5)$$

and $n_i(\phi)$ is the number density of galaxies belonging to a sample at redshift $\langle z_i \rangle$ observed at position angle ϕ, and $\bar{n}_i$ is the average density of the ith sample. The superscript g indicates the intrinsic galaxy component and the superscript μ indicates the magnification component. Combining equations (5) and (4) yields

$$w_{1,2}(\theta) = \langle \delta n_1^g(\phi) \delta n_2^g(\phi + \theta) \rangle_\phi + \langle \delta n_1^g(\phi) \delta n_2^\mu(\phi + \theta) \rangle_\phi + \langle \delta n_1^\mu(\phi) \delta n_2^g(\phi + \theta) \rangle_\phi + \langle \delta n_1^\mu(\phi) \delta n_2^\mu(\phi + \theta) \rangle_\phi\,. \quad (6)$$

$w_{1,2}(\theta)$ contains four terms. The first is the intrinsic galaxy clustering due to gravity. In the limit that sample 1 and sample 2 do not overlap, this term disappears. Any redshift overlap between samples 1 and 2 increases the strength of this term. The second and third components arise from the lensing magnification and depend on the amount of overlap. $w_{1,2}(\theta)$ is dominated by the second term when $\langle z_1 \rangle < \langle z_2 \rangle$ and there is no redshift overlap between samples 1 and 2. The fourth term is due to pure matter–matter correlations, which we will ignore for the remainder of this work as it is sub-dominant to the other terms in every case considered here (Heavens & Joachimi 2011; Duncan et al. 2014).

Focusing on the first two terms, respectively, and assuming linear bias:

$$\langle \delta n_1^g(\phi) \delta n_2^g(\phi + \theta) \rangle_\phi = b_1 b_2 \int_0^{\chi_H} \mathrm{d}\chi\, \eta_1(\chi) \eta_2(\chi) \times \int_0^\infty \frac{k\, \mathrm{d}k}{2\pi} P_\delta(k, \chi) J_0(\chi k \theta)\,, \quad (7)$$

where b_i is the galaxy bias of sample i, χ is the comoving distance, P_δ is the 3D dark matter power spectrum, η_i is the comoving distance distribution of sample i and J_0 is the zeroth order Bessel function.

$$\langle \delta n_1^g(\phi) \delta n_2^\mu(\phi + \theta) \rangle_\phi = b_1 (\alpha - 1) \int_0^{\chi_H} \mathrm{d}\chi\, \eta_1(\chi) K(\chi) \times \int_0^\infty \frac{k\, \mathrm{d}k}{2\pi} P_\delta(k, \chi) J_0(\chi k \theta)\,, \quad (8)$$

where α is the slope of the magnitude number counts, defined formally as

$$\alpha(r) = 2.5 \frac{\mathrm{d} \log_{10} n(> r)}{\mathrm{d}r}\,, \quad (9)$$

with n the observed galaxy number density and r the galaxy magnitude. α is measured using the appropriate detection band (r-band for RCSLenS and i-band for CFHTLenS). $K(\chi)$ is the lensing kernel-weighted distribution of background sources $\eta_2(\chi)$ defined as,

$$K(\chi) = \frac{3 H_0^2 \Omega_\mathrm{m}}{c^2} \frac{\chi}{a} \int_\chi^{\chi_H} \mathrm{d}\chi' \eta_2(\chi') \frac{\chi' - \chi}{\chi'}\,, \quad (10)$$

with H_0 the *Hubble* constant today, a the scalefactor and assuming a flat universe.

Magnification affects the clustering in two ways. First, it raises the flux of a magnified galaxy such that the galaxy count might be boosted within the flux limit of a survey. Secondly, it increases the observed solid angle around a magnifying galaxy. The net effect, given by equation (8),[1] depends on the slope of the luminosity function of the background objects, α, defined in equation (9). See Scranton et al. (2005), Hildebrandt et al. (2009), Ménard et al. (2010) and Morrison et al. (2012) for measurements of the magnification signal via number counts. We measure α for RCSLenS and CFHTLenS by calculating the slope of the cumulative number densities of galaxies as a function of limiting magnitude and binned by photometric redshift (Duncan et al. 2014).

We use the halo model implemented in the PYTHON package CHOMP[2] to generate theoretical predictions for the magnification contribution to $w(\theta)$ for BOSS galaxies. We take the BOSS halo occupation distribution (HOD) parameters determined in Parejko et al. (2013) for the LOWZ sample and Miyatake et al. (2013) for the CMASS sample. We do not estimate predictions for the magnification around WiggleZ galaxies due to the difficulty of obtaining an HOD description of this sample, which is not volume-limited. The WiggleZ galaxies have lower masses, so we expect the amplitude of the signal to be negligible in the one-halo regime. The signal might be more comparable in the regime where the two-halo term dominates (roughly the largest two to three θ bins), but the S/N of the data over these last few bins does not warrant the modelling as we will see in Section 4.3. Appendix B contains further details of the halo model and the input HOD parameters to the estimated magnification signal.

2.2 Contamination from photometric redshift errors

In this section, we extend the formalism presented in Benjamin et al. (2010, 2013), who considered cross-correlations between photometric redshift bins. Here, we examine the case of cross-correlations

[1] The magnification contribution to the clustering signal of off-diagonal redshift bin combinations can be positive or negative.
[2] http://code/google.com/p/chomp

between spectroscopic and photometric redshift bins. We define the following quantities.

(i) Each galaxy has a single-point best-fitting photometric redshift z_B and photometric redshift probability distribution $p(z)$.

(ii) $N_j^{\rm O}$ is the observed number of galaxies with single-point photometric redshifts that place them in bin j.

(iii) When discussing aggregate or summed $p(z)$, we use the notation $\Phi_j(z)$ as the sum of the $p(z)$ for all galaxies in photometric redshift bin j,

$$\Phi_j(z) = \sum_{k=1}^{N_j^O} p_{kj}(z)\,. \tag{11}$$

(iv) $N_{ij}^{\rm T}$ is the true number of galaxies which would have a spectroscopic redshift in bin i (if spectroscopy had been measured for those galaxies) and placed in photometric redshift bin j. $N_{ii}^{\rm T}$ is then the number of galaxies that have both spectroscopic and photometric redshifts placing them in bin i.

(v) $w_{ij}^{\rm SP,O}$ is the observed clustering signal between galaxies with photometric redshifts in bin j and galaxies in the spectroscopic sample with spectroscopic redshifts in bin i and includes all terms from equation (6).

(vi) $w_{ij}^{\rm SP,T}$ is the true clustering signal between galaxies in the full photometric sample with spectroscopic redshifts in bin j and those in the spectroscopic sample with spectroscopic redshifts in bin i.

If $i \neq j$ (which we will often refer to as 'off-diagonal'), in the absence of lensing magnification, the true cross-clustering signal $w_{ij}^{\rm SP,T} = 0$. Consider the pairwise case of the first two bins. When $i = j$ and the spectroscopic and photometric redshift bins are specified to be the same range ('diagonal'), the observed cross-clustering signal is related to the true cross-clustering signal. The true signal is scaled by the ratio of the true number of galaxies assigned to photometric bin $j = 1$, which actually have a true redshift in the same spectroscopic bin $i = 1$, to the total observed number of galaxies in photometric bin $j = 1$. When the spectroscopic and photometric redshift bins are different (off-diagonal, e.g. $i = 1$ and $j = 2$) but still overlapping due to photometric redshift scatter and outliers, the observed cross-clustering signal will be the true diagonal cross-clustering signal multiplied by the ratio of the true number of galaxies assigned to photometric bin $j = 2$, which actually have a true redshift in spectroscopic bin $i = 1$, to the total observed number of galaxies in photometric bin $j = 2$. Written using the quantities defined above and generalizing to the case of multiple bins,[3]

$$w_{ij}^{\rm SP,O} = \frac{w_{ii}^{\rm SP,T} N_{ij}^{\rm T}}{N_j^{\rm O}}\,. \tag{12}$$

2.3 Procedure for estimating the clustering signal using $\Phi_j(z)$

We outline the steps to model the intrinsic clustering signal (i.e. equation 7) of off-diagonal redshift bin combinations using photometric redshift error distributions. We use $E(x)$ to denote the estimator for the quantity x defined in Section 2.2.

[3] Equation (4) from Benjamin et al. (2013) describes the observed cross-clustering signal between two photometric redshift bins i and j. In this analysis, one bin is always spectroscopic, so that the fractional leak from the spectroscopic bin 1 to the photometric bin 2 is always zero. In this case, the formalism presented here is consistent with Benjamin et al. (2013).

(i) Step 0: create spectroscopic and photometric samples using redshift bin limits based on available spectra. In this work, we use the same bin limits for both spectroscopic and photometric bins due to signal-to-noise considerations. In principle, however, the spectroscopic information has much higher resolution and can potentially be binned more finely.

(ii) Step 1: measure the observed number of galaxies in each photometric bin $N_j^{\rm O}$ by counting the single-point photometric redshift estimates z_B.

(iii) Step 2: measure the auto- and cross-correlations between spectroscopic and photometric bins and corresponding covariance matrices. In this work, we use jack-knife resampling to estimate the covariance matrices.

(iv) *Step 3:* estimate $E(N_{ij}^{\rm T})$ by summing the $p(z)$ corresponding to galaxies selected to be in a photometric redshift bin j using z_B to obtain $\Phi_j(z)$ (equation 11) and then integrating $\Phi_j(z)$ over the limits of true-z bin i. Note that $E(N_{ij}^{\rm T})$ summed over all bins i must equal $N_j^{\rm O}$. In practice, we implement this constraint by ensuring that $\Phi_j(z)$ integrated from the minimum to the maximum of the full redshift range normalizes to $N_j^{\rm O}$.

(v) *Step 4:* use the $E(N_{ij}^{\rm T})$ to predict the observed cross-correlation, which can be derived from equation (12) as

$$E\left(w_{ij}^{\rm SP,O}\right) = w_{ii}^{\rm SP,O} \frac{N_i^{\rm O}}{N_j^{\rm O}} E\left(\frac{N_{ij}^{\rm T}}{N_{ii}^{\rm T}}\right)\,. \tag{13}$$

(vi) *Step 5:* compare the predicted $\mathrm{E}(w_{ij}^{\rm SP,O})$ and observed cross-correlation $w_{ij}^{\rm SP,O}$. If they agree, $\Phi_j(z)$ is a good estimate of the redshift distribution of the galaxy sample.

We present a validation of this method on an idealized mock galaxy catalogue in Appendix A.

By constructing the estimate in this way, we bypass having to constrain the true galaxy bias of the spectroscopic sample in each bin, b_i, and the galaxy bias of our photometric sample in each bin, b_j, since they are included within the measurement of the diagonal cross-correlations as can be seen in equation (7). However, we must assume that b_j has either no or slow evolution over the full extent of the distribution in each photometric bin and, more importantly, that catastrophic outliers are a random sample of the bin's galaxy population, such that the average galaxy bias of the outliers is the same as the average bias of the whole population. We revisit these assumptions in Section 3.3.

2.4 Goodness of fit

To compare the predicted model and observed cross-correlation signal, we calculate a goodness of fit as measured by χ^2. We define the χ^2 of the fit of the model prediction to the data as

$$\chi^2 = (\boldsymbol{d} - \boldsymbol{m})^T \boldsymbol{C}^{-1} (\boldsymbol{d} - \boldsymbol{m})\,, \tag{14}$$

where $\boldsymbol{d}$ is a vector of length $N_z \times N_\theta$ containing the measured $w_{ij}(\theta_q)$ with $i = \{1, \ldots, N_{\rm s}\}, j = \{1, \ldots, N_{\rm p}\}$ and $q = \{1, \ldots, N_\theta\}$. $N_{\rm s}$ and $N_{\rm p}$ are the number of spectroscopic and photometric redshift bins, respectively. $N_z = N_{\rm s} \times N_{\rm p} - \min(N_{\rm s}, N_{\rm p})$ is the total number of spectroscopic-photometric redshift bin combinations with $i \neq j$. N_θ is the total number of angular scales for which $w_{ij}(\theta_q)$ is calculated. $\boldsymbol{m}$ contains the model for $w_{ij}(\theta_q)$ as given by equation (13). This model depends on $w_{ii}(\theta_q)$, which has an associated error. We assume the two terms $w_{ij}(\theta_q)$ and $w_{ii}(\theta_q)$ are independent and add the corresponding covariance matrices, such that the total covariance matrix $\boldsymbol{C}_{ij} = \mathrm{Cov}\left[\mathrm{w_{ij}}(\theta_{\rm q}), \mathrm{w_{ij}}(\theta_{\rm r})\right] + \mathrm{f_{ij}^2 Cov}\left[\mathrm{w_{ii}}(\theta_{\rm q}), \mathrm{w_{ii}}(\theta_{\rm r})\right]$

is the propagated jack-knife resampled covariance between θ scales q and r, and $f_{ij} = (N_i^{\rm O}/N_j^{\rm O})E(N_{ij}^{\rm T}/N_{ii}^{\rm T})$ is the pre-factor in equation (13). Note that the spectroscopic bins are sufficiently broad such that radial correlation between bins, and thus the covariance between different spectroscopic i bins in this analysis, is negligible. Throughout this work, we use angular scales in the range 1 arcmin $< \theta <$ 35 arcmin, where the limits are chosen to mitigate the impact of scale-dependent galaxy biases while still utilizing the signal at intermediate scales.

2.5 Modifying photometric redshift distributions

As we will see in Section 4, the models predicted from the $\Phi_j(z)$ are often a poor fit to the data. We therefore investigate two methods to modify the redshift distributions for galaxy samples binned by photometric redshift.

The first modification takes the $\Phi_j(z)$ and allows for a shift along the z-dimension in the overall distribution. This represents a correction for a bias error in the measured $\Phi_j(z)$, whilst maintaining the level of scatter and catastrophic outliers as specified by the $\Phi_j(z)$. This is modelled using one free parameter per photometric redshift bin j, Δz_j. When the probability gets shifted to negative redshifts, we re-normalize $\Phi_j(z)$ by the integrated $\Phi_j(z < 0)$. This shifting approach is similar to that adopted by the cosmological tomographic shear analysis of The Dark Energy Survey Collaboration et al. (2015), who allow for an independent shift of the estimated photometric redshift distribution as a nuisance parameter.

The second modification models the redshift distributions in each photometric redshift bin as a Gaussian. This model has two free parameters per photometric redshift bin j with a mean μ_{zj} and standard deviation σ_{zj} and are defined on a baseline redshift range with $z > 0$. This test allows us to directly determine the photometric redshift bias and scatter in each bin, independent of the BPZ $\Phi_j(z)$. The limitation of this single Gaussian model definition, however, is that it sets all catastrophic outliers to zero.

We determine fits for these free parameters for photometric redshift bins by jointly fitting the data using Monte Carlo Markov Chain (MCMC) sampling, as implemented with the PYTHON code EMCEE[4] (Foreman-Mackey et al. 2013), which is an implementation of the affine invariant methods by Goodman & Weare (2010). For a given combination of spectroscopic/photometric redshift surveys, we minimize the negative log likelihood calculated jointly for every ij cross-correlation as a function of either the additive shifts to the $\Phi_j(z)$, or the Gaussian model parameters. In the MCMC analysis of the data, we do not include the effects of magnification as in most cases, the contribution is at the percent level on the smallest scales used (1 arcmin).

In Appendix A, we demonstrate that our methodology to determine the redshift distribution offsets is valid for the idealized case of Gaussian errors in the photometric redshift distributions.

3 DATA

In this section, we provide details about the photometric and spectroscopic data sets we use in this study and an investigation into the dependence of photometric redshift errors on galaxy type.

[4] http://dan.iel.fm/emcee/current

3.1 Photometric surveys

We utilize two deep, wide, and high-resolution photometric surveys observed by MegaCam on the 3.6-m Canada–France–Hawaii Telescope (CFHT). The CFHTLenS (Heymans et al. 2012; Erben et al. 2013)[5] is based on 154 deg^2 of *ugriz* imaging from the wide component of the CFHT Legacy Survey. Heymans et al. (2012) provide an overview of the pipeline, and details about the data analysis can be found in the following: Erben et al. (2013) describes the data reduction with automated masking; the photometry was Gaussianized to homogenize the point spread functions among the different filters, object catalogues were created with SEXTRACTOR (Bertin & Arnouts 1996; Erben et al. 2013), and photometric redshifts were estimated using BPZ (Benítez 2000; Hildebrandt et al. 2012); galaxy shapes were estimated using Bayesian model fitting with *Lens*fit (Miller et al. 2013). The mean seeing is 0.72 arcsec (r band) and 0.68 arcsec (i band), and the median redshift is $z_{\rm m} = 0.7$.

The RCSLenS[6] is based on the second Red-sequence Cluster Survey (RCS2), comprising nearly 800 deg^2 of multicolour imaging 1–2 mag deeper than the Sloan Digital Sky Survey (SDSS) (Gilbank et al. 2011). The resolution is lensing-quality with a median seeing in the r band of 0.7 arcsec. A total of 513 deg^2 is imaged in multiple bands with *griz*, allowing photometric redshifts and $p(z)$ to be estimated. The images were processed and the object catalogues were created with the same methods applied to CFHTLenS (Hildebrandt et al. 2016). The number distribution of magnitudes as a function of unweighted objects led to choices of magnitude cut-offs of $i = 24.7$ for CFHTLenS sources and $r = 23.7$ for RCSLenS sources. These numbers correspond to the 5σ detection limit in a 2.0 arcsec aperture (Erben et al. 2013).

Fig. 1 shows the normalized $\Phi_j(z)$ corresponding to the eight photometric redshift bins considered in this analysis with limits given by: [0.15, 0.29], [0.29, 0.43], [0.43, 0.57], [0.57, 0.7], [0.7, 0.9], [0.9, 1.1], [1.1, 1.3], [1.3, 3.5]. The choice of bin edges was motivated by the spectroscopic samples used (i.e. using the LOWZ and CMASS partitions described in Section 3.2), and $z = 1.3$ was the cut-off redshift for previous CFHTLenS lensing analyses, motivated by the lack of near-infrared photometry to constrain higher redshifts. The top panel shows the $\Phi_j(z)$ for the entire CFHTLenS area overlapping with BOSS Data Release (DR10), while the lower panel shows the $\Phi_j(z)$ for the RCSLenS overlap with BOSS DR10 and WiggleZ. Compared with CFHTLenS, the RCSLenS $\Phi_j(z)$ are noticeably multimodal, with flatter tails extending to adjacent bins, which reflects poorer photometric redshift accuracy due to the limited filter coverage, most critically the missing u band.

As the calculation of $w(\theta)$ requires pair counts with random positions (see equation 3), we generate random catalogues for each field, taking into account edges and masks.

3.2 Spectroscopic surveys

We briefly summarize the spectroscopic data sets used in this analysis but refer the interested reader to Blake et al. (2016) for further details and statistics regarding the WiggleZ–RCSLenS, BOSS–CFHTLenS, BOSS–RCSLenS overlap regions, as they use nearly identical WiggleZ and BOSS spectroscopic samples.

[5] http://www.cfhtlens.org
[6] http://www.rcslens.org

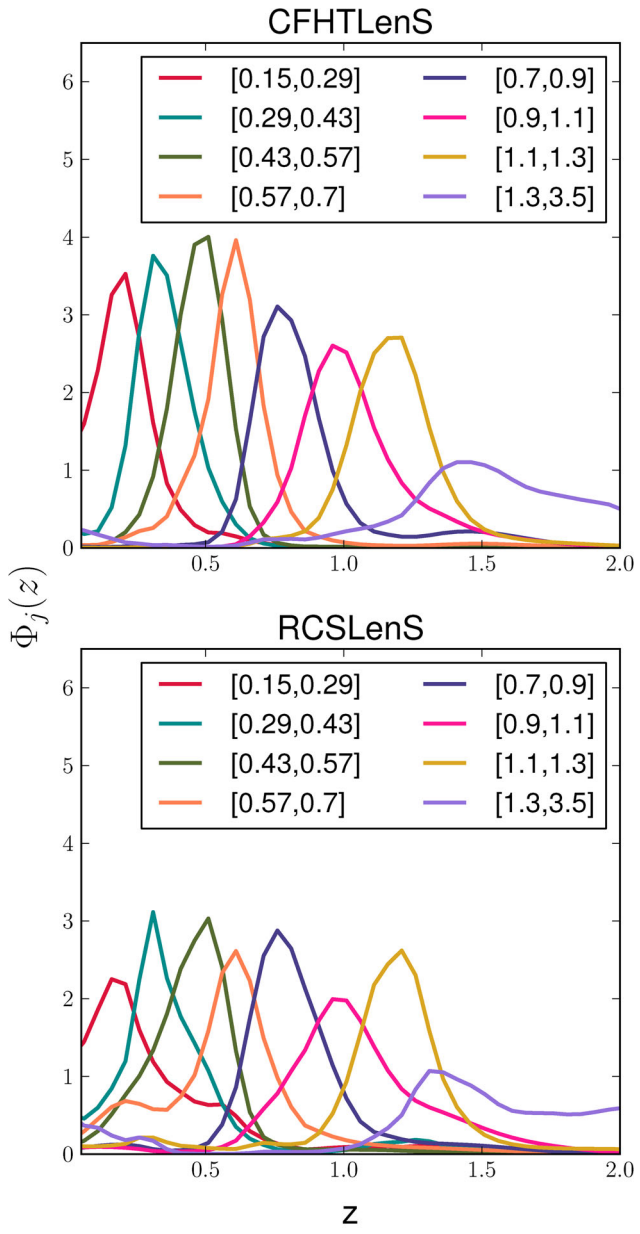

Figure 1. Summed probability redshift distributions $\Phi_j(z)$ for galaxies with single-point photometric redshifts z_B in different bins j. The top panel corresponds to the 66 deg^2 of unmasked overlap between CFHTLenS and BOSS DR10. The bottom panel corresponds to the 184 deg^2 where there is unmasked overlap between RCSLenS and BOSS DR10.

BOSS

The BOSS (Eisenstein et al. 2011) is a spectroscopic survey of massive galaxies and quasars selected from SDSS photometry carried out at the Sloan Telescope at Apache Point Observatory in Sunspot, New Mexico, USA. DR10 contains BOSS spectra taken through 2012 July and comprises 927 844 galaxy spectra over 6373.2 deg^2 (Ahn et al. 2014). BOSS galaxies were selected using colour and magnitude cuts and are divided into the 'LOWZ' sample with red galaxies $z < 0.43$ and the 'CMASS' sample which is designed to be approximately stellar mass-limited for $z > 0.43$. There is a total of 66.3 deg^2 of total unmasked overlap with CFHTLenS (W1 and W4, 2830 LOWZ galaxies, 5567 CMASS galaxies) and a total of 183.9 deg^2 of total unmasked overlap with RCSLenS (six fields, 9214 LOWZ galaxies, and 18 156 CMASS galaxies). See table 1 of Blake et al. (2016) for numbers corresponding to each field. The six RCSLenS fields are labelled as 0047, 0133, 1514, 1645, 2143 and 2329. The field named 1303 has a very small number of galaxies with spectroscopic redshifts, and we exclude it from this analysis. We trim the catalogues to restrict them to the overlap regions (i.e. no BOSS or CFHTLenS/RCSLenS galaxies falling outside of the overlap are included in the analysis).

BOSS galaxies are assigned completeness weights as in equation (18) of Anderson et al. (2014) in order to correct for the effects of redshift failures, fibre collisions and other known systematics, and we use these weights in our determination of the cross-correlations of the galaxies. Specifically, the pair counts in equation (3) are weighted by the completeness weights.

WiggleZ

The WiggleZ Dark Energy Survey is a spectroscopic survey of bright emission line galaxies with median redshift ∼0.6 carried out at the Anglo-Australian Telescope in Siding Spring, Australia (Drinkwater et al. 2010). WiggleZ galaxies were selected using colour and magnitude cuts from a combination of optical and UV imaging. There is a total of 175.1 deg^2 of total unmasked overlap with RCSLenS imaging that has four bands and photometric redshift estimates. 76 900 galaxies covering the range $0.1 < z_s < 0.9$ reside in this region, which is comprised of five different RCSLenS fields (0047, 0310, 2143, 2329, 2338; for numbers corresponding to each field, see table 1 of Blake et al. 2016). Note that we use WiggleZ galaxies at higher redshifts compared to Blake et al. (2016), who cut their samples to $z_s < 0.7$ to match BOSS. 1514 contains a small number of galaxies with spectroscopic redshifts, and the geometry of the overlap with the RCSLenS data is irregular and patchy (the pointings are non-contiguous). The estimated covariance matrices are not positive-definite, and we exclude 1514 in this analysis. We again trim the catalogues to restrict them to the overlap regions (i.e. no WiggleZ or RCSLenS galaxies falling outside of the overlap are included in the analysis).

VVDS, VIPERS and DEEP2 EGS

CFHTLenS overlaps with three small area, but densely sampled deep spectroscopic surveys. There are ∼2.6^2 of spectroscopic overlap with the VIMOS VLT Deep Survey (VVDS; Le Fèvre et al. 2005). VVDS selects objects with $17.5 \leq i \leq 24$ and is ∼90 per cent complete down to $i < 23$. There are ∼0.6 deg^2 of spectroscopic overlap with the extended Groth Strip EGS DEEP2 survey (Davis et al. 2007; Newman et al. 2013). DEEP2 selects objects with $18.5 \leq R_{\rm AB} \leq 24.1$ with ∼60 per cent of objects with $i < 23$. Finally, there are 23.1 deg^2 of spectroscopic overlap with the VIMOS Public Extragalactic Redshift Survey (VIPERS; Guzzo et al. 2014). VIPERS selects $i < 22.5$ galaxies with an additional colour selection to target galaxies in the redshift range $0.5 < z < 1.5$ and is thus highly incomplete at $z < 0.5$. We use redshift quality flags 3 and 4 (>95 per cent secure) for all of the spectroscopic surveys and create a matched catalogue in order to directly compare photometric redshifts with spectroscopic redshifts in Sections 3.3 and 4.2.

3.3 The dependence of photometric redshift errors on galaxy type

There is one main advantage of the methodology we propose in this paper over the reconstruction methods, discussed in Section 1. The unknown galaxy bias is fully accounted for in our modelling if the

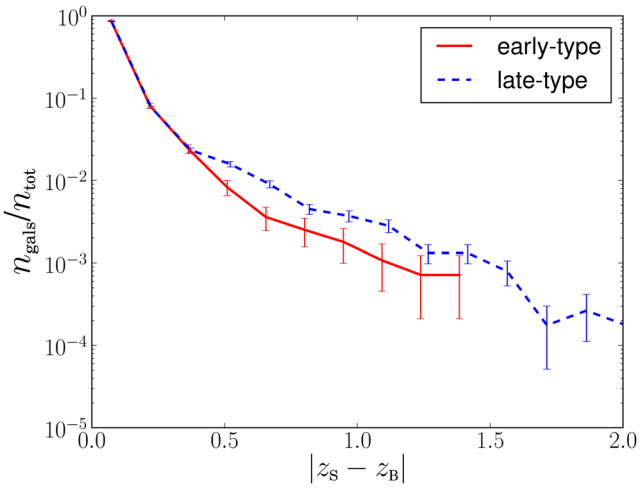

Figure 2. The distribution of the absolute magnitude of the differential between spectroscopic redshift and photometric redshifts as a function of galaxy type as measured by BPZ. There are 2783 ET galaxies and 11,341 LT galaxies.

average galaxy bias of the population of outliers at a true redshift z is not significantly different from the average galaxy bias of the main population at that same redshift. If that is the case, we do not need to incorporate nuisance parameters that model scale-dependent and redshift-dependent galaxy bias in our analysis even though the typical photometric sample will have a different mean galaxy bias than the spectroscopic sample with which it is cross-correlated (e.g. the highly biased BOSS galaxies). All these averaged galaxy bias properties are encompassed by the autocorrelation signal $w_{ii}^{\rm SP,O}$ (equation 12). Outliers in the photometric redshift measurements come from two sources. The first is from random photometric errors as the majority of the faint galaxies used in cosmological analyses are detected below 10σ. We would not expect this form of outlier to be galaxy-type dependent. The second source, however, is template or training set degeneracies, where a low-redshift red galaxy has the same colour as a high-redshift intrinsically blue galaxy. These outliers are clearly galaxy type-dependent.

To test the dependence of photometric redshift outliers on galaxy type, Fig. 2 shows the absolute difference between spectroscopic and photometric redshifts, $|z_{\rm s} - z_{\rm B}|$ for the CFHTLenS–VVDS-matched catalogue described in Section 3.2. Early-type galaxies are selected with the BPZ template type $T_{\rm B} < 1.5$ (shown solid) and late-type galaxies are selected with $2 < T_{\rm B} < 4$ (shown dashed), where these ranges in $T_{\rm B}$ are shown by Velander et al. (2014) to separate red and blue galaxies well. These distributions have been normalized by the total number of galaxies for the given type and are shown on a log-scale to enhance the differences. We see that the numbers for early-type and late-type galaxies for $|z_{\rm s} - z_{\rm B}| > 0.4$ are different, with the majority of extreme outliers ($\Delta z > 0.2$) being late-type. For the purposes of our analysis, the different galaxy bias properties of these extreme outliers would only impact upon the conclusions we drew from the most separated redshift bins where we find little signal to constrain the redshift distributions anyway.

4 RESULTS

We measure the cross-correlations between three combinations of the photometric and spectroscopic surveys described above – CFHTLenS–BOSS, RCSLenS–BOSS, and RCSLenS–WiggleZ. For the eight redshift bins shown in Fig. 3, we measure the cross-correlation $w_{ij}(\theta)$ using equation (3) between spectroscopic redshift bin i and photometric redshift bin j for seven logarithmically spaced angular bins in the range 1 arcmin $< \theta <$ 35 arcmin. Errors are jack-knife-resampled such that each 1 deg^2 pointing is a jack-knife sub-sample. These sub-samples maximize the regularity in the shapes of the jack-knife regions, as advocated by Cabré et al. (2007) and Norberg et al. (2009). This internal method of error estimation is approximate, as Norberg et al. (2009) have found jack-knife-resampled covariance matrices to be somewhat biased on small angular scales. All of the covariance matrices pass tests for positive definiteness and have eigenvalues spanning a reasonable range.

4.1 Photometric redshift accuracy in CFHTLenS

In Fig. 3, we present our results for the measured cross-correlations between the CFHTLenS and BOSS galaxies as filled circles. Each panel represents a different cross-correlation between a spectroscopic and a photometric bin with the spectroscopic redshifts increasing from top to bottom and the photometric redshifts increasing from left to right. The red dashed line shows the predicted clustering for the off-diagonal panels using equation (12). The orange solid band shows the predicted magnification from equation (8) given HOD parameters taken from Parejko et al. (2013) for the LOWZ sample and Miyatake et al. (2013) for the CMASS sample. The α values used are [0.52, 0.52, 0.46, 0.47, 0.45, 0.68, 0.89, 0.89] for CFHTLenS, where each array value corresponds to the photometric redshift bins in ascending order. The errors obtained by bootstrap resampling the α values are all of the order of 1×10^{-3}, so we do not explicitly quote them here. The band of values is set by the minimum and maximum values possible from sampling the HOD parameters within their 1σ uncertainties. The purple and cyan bands show the best-fitting models when we allow for modification of the $\Phi_j(z)$ as described in Section 2.5. We discuss these best-fitting models below in Section 4.1.1.

Focusing first on the data measurements, the diagonal cross-correlations are depicted in red and have the highest amplitude, as is expected given the $\Phi_j(z)$ plotted in Fig. 1. That is, the $\Phi_j(z)$ selected by z_B peak are in the appropriate range. However, the significant overlap between adjacent bins also indicates that we would expect relatively high amplitudes in the off-diagonal panels closest to the diagonal panels, and this trend is confirmed in Fig. 3. The panels that are furthest away from the diagonal reflect cross-correlations with smaller amplitude. Intrinsic clustering can only cause positive cross-correlations, and thus, it is clear that the most widely separated bins have little intrinsic clustering and a negligible amount of contamination between low-z photometric redshift bins and high-z spectroscopic redshift bins and vice-versa.

Fig. 3 provides a wealth of information. To illustrate this, concentrate on a particular sample, the $0.7 < z_B < 0.9$ bin. Here, we find the leakage to the $0.57 < z_{\rm s} < 0.7$ bin as revealed by the non-zero cross-correlation measured with the spectroscopic sample in this range, is well-modelled by the best-fitting $\Phi_j(z)$. The anti-correlation seen with lower spectroscopic redshift samples agrees with the predictions from lensing magnification. Interestingly, some widely separated bins contain a negative signal (of the order of -1 per cent at 1 arcmin) in 1 or 2σ tension with the magnification predictions. These anticorrelations could potentially be explained by systematic effects such as object detection and deblending problems that are not fully characterized and accounted for in the angular clustering measurements (Morrison & Hildebrandt 2015; Simet &

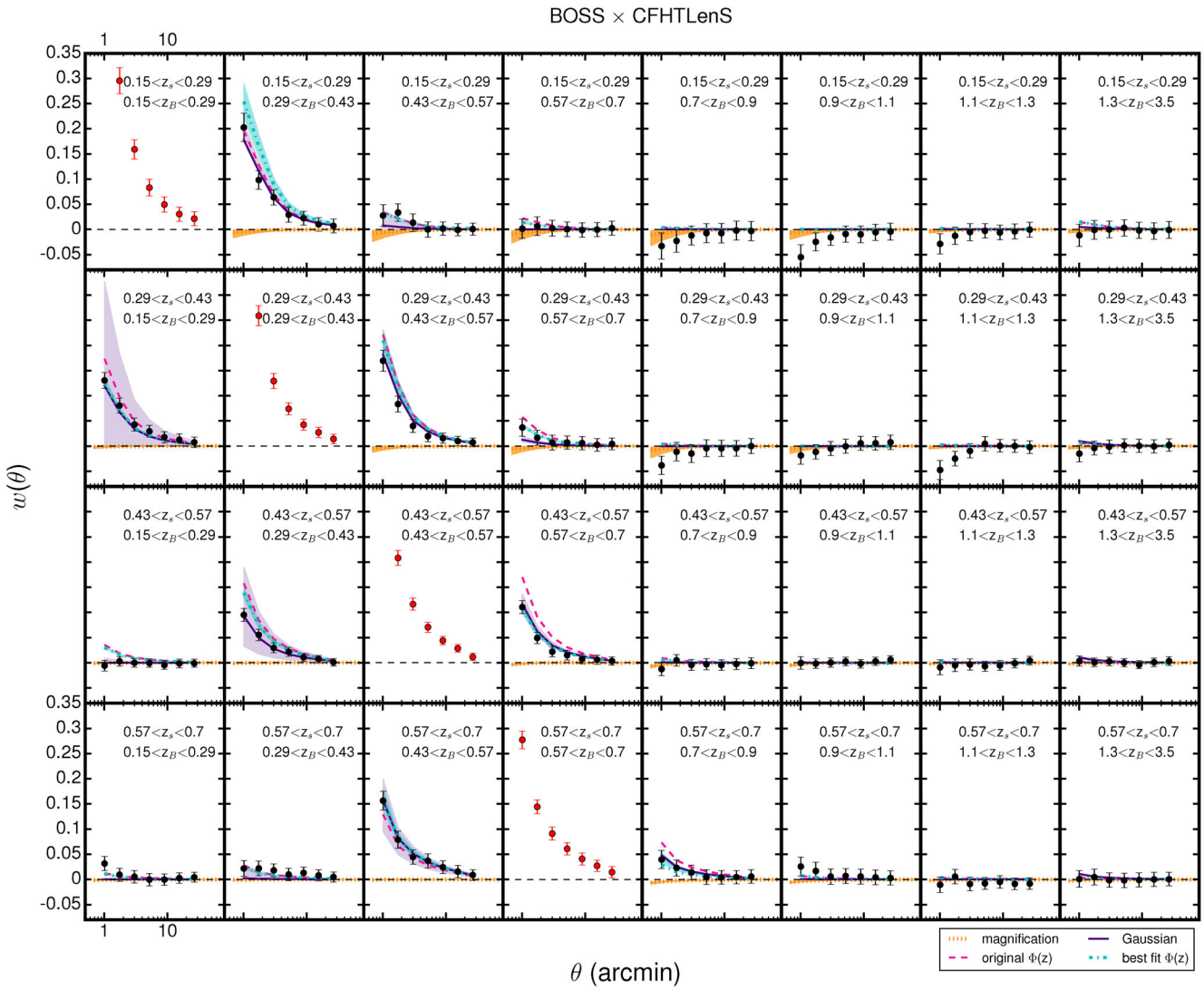

Figure 3. The measured cross-correlations between the CFHTLenS (photometric redshift) bin j and BOSS (spectroscopic redshift) bin i galaxies. The data points are weighted means over the two CFHTLenS fields (W1 and W4), with the weights given by the number of pairs $(D_iD_j)_\theta$ (see equation 3). The red dashed line shows the predicted ij clustering where $i \neq j$ based on equation (12). The orange band shows a predicted magnification contribution based on halo model fits to the BOSS galaxies from the literature. The width of the orange band quantifies how the predicted magnification signal changes for the uncertainties in the quoted halo model parameters from the literature. The cyan band shows the predicted ij clustering after the $\Phi_j(z)$ have been shifted by the best-fitting quantities given in Table 1. The purple band shows the predicted ij clustering given by best-fitting Gaussian $p(z)$.

Mandelbaum 2015; Suchyta et al. 2016). In Appendix C, we investigate these systematic effects and find that while they likely play a role in the negative signals, we are unable to constrain their contribution without the aid of sophisticated image simulations.

Briefly returning to the question of galaxy bias discussed in Section 3.3, we repeat our analysis for the CFHTLenS-BOSS cross-correlation for the case of BPZ template type $T_B < 1.5$ and compare it to the case of $2 < T_B < 4$. Here we find that $w_{ii}^{SP,O}$ does differ in terms of amplitude and angular dependence for red and blue galaxies. However, the angular dependence (i.e. shape) is similar over different photometric redshift bins, which supports the assertion that the evolution of the scale-dependent photometric galaxy bias is captured by the diagonal clustering measurements in each bin. From this test and the analysis presented in Section 3.3, we conclude that varying galaxy bias does not impact upon the conclusions drawn in this paper. However, the different galaxy biases in the outlier population will be important to model for reconstruction methods and future higher fidelity implementations of the methodology presented in this analysis.

4.1.1 *Improved CFHTLenS photometric redshift distributions*

We determine best-fitting $\Phi_j(z)$ for each photometric redshift bin using the procedure outlined in Section 2.5, which consists of two approaches. The first method fits a shift to the original BPZ $\Phi_j(z)$ along the z-dimension, and Fig. 4 shows an example of the MCMC results for CFHTLenS cross-correlated with BOSS. Fig. 4 and subsequent corner plots were made with TRIANGLE.PY[7] (Foreman-Mackey et al. 2014). We impose the prior $-0.15 < \Delta z_j < 0.15$ to avoid sampling the parameter space where the $\Phi_j(z)$ swap positions. The best-fitting values for the redshift bias in each redshift bin are given in Table 1.

[7] https://github.com/dfm/corner.py

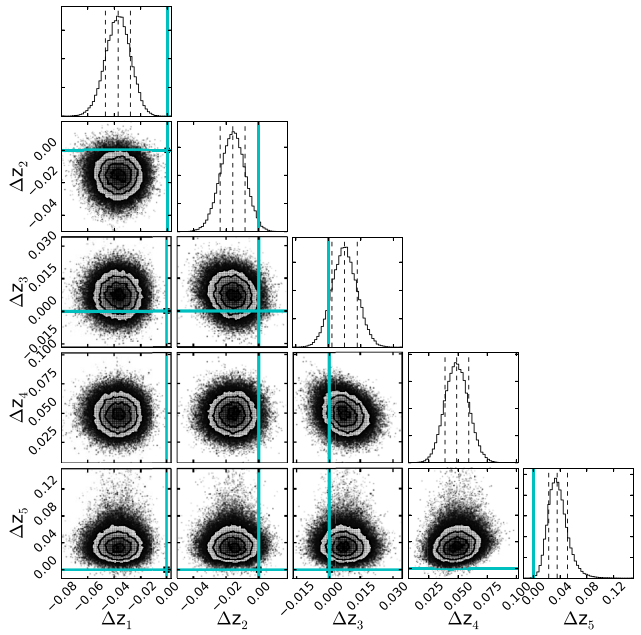

Figure 4. Joint fit shifts for photometric bin j for CFHTLenS cross-correlated with BOSS. The dashed black lines mark the 16th, 50th and 84th percentiles of the samples in the marginalized distributions, and the solid cyan lines mark $\Delta z_j = 0$.

Table 1. A summary of the best jointly fit Δz_j shifts and best-fitting μ_{zj} and σ_{zj} for the three spectroscopic–photometric redshift survey combinations investigated in this work. The effective unmasked overlap area is provided for each photometric/spectroscopic survey combination. The reduced χ^2 is also provided, with the sub-script 'orig' for the BPZ $\Phi_j(z)$, 'shift' for the BPZ $\Phi_j(z)$ after best-fitting shifts have been applied, and 'Gauss' for the best-fitting Gaussians.

	CFHTLenS/BOSS	RCSLenS/BOSS	RCSLenS/WiggleZ
$A_{\rm eff}$	66.3	180.2	175.1
Δz_1	$-0.037^{+0.009}_{-0.010}$	–	–
Δz_2	$-0.016^{+0.008}_{-0.008}$	$-0.095^{+0.007}_{-0.007}$	$0.070^{+0.019}_{-0.019}$
Δz_3	$0.007^{+0.006}_{-0.006}$	$-0.084^{+0.005}_{-0.005}$	$-0.001^{+0.011}_{-0.011}$
Δz_4	$0.049^{+0.010}_{-0.010}$	$0.236^{+0.026}_{-0.019}$	$0.024^{+0.018}_{-0.022}$
Δz_5	$0.036^{+0.016}_{-0.013}$	$0.143^{+0.024}_{-0.016}$	$-0.015^{+0.018}_{-0.022}$
Δz_6	–	–	$-0.040^{+0.029}_{-0.028}$
$\chi^2_{\rm red,orig}$	1.31	1.70	1.09
$\chi^2_{\rm red,shift}$	1.17	1.29	1.07
μ_{z1}	$0.233^{+0.019}_{-0.020}$	$0.152^{+0.003}_{-0.001}$	$0.281^{+0.007}_{-0.013}$
μ_{z2}	$0.327^{+0.013}_{-0.017}$	$0.291^{+0.002}_{-0.001}$	$0.327^{+0.025}_{-0.022}$
μ_{z3}	$0.505^{+0.010}_{-0.012}$	$0.437^{+0.005}_{-0.004}$	$0.455^{+0.015}_{-0.013}$
μ_{z4}	$0.678^{+0.013}_{-0.014}$	$0.698^{+0.002}_{-0.003}$	$0.583^{+0.019}_{-0.010}$
μ_{z5}	$0.823^{+0.048}_{-0.056}$	$0.832^{+0.048}_{-0.083}$	$0.822^{+0.046}_{-0.056}$
μ_{z6}	–	–	$1.063^{+0.026}_{-0.041}$
σ_{z1}	$0.052^{+0.023}_{-0.022}$	$0.169^{+0.009}_{-0.008}$	$0.206^{+0.003}_{-0.006}$
σ_{z2}	$0.091^{+0.034}_{-0.017}$	$0.148^{+0.007}_{-0.007}$	$0.209^{+0.008}_{-0.015}$
σ_{z3}	$0.098^{+0.026}_{-0.018}$	$0.105^{+0.004}_{-0.004}$	$0.142^{+0.013}_{-0.013}$
σ_{z4}	$0.114^{+0.011}_{-0.016}$	$0.227^{+0.002}_{-0.004}$	$0.201^{+0.018}_{-0.023}$
σ_{z5}	$0.0846^{+0.031}_{-0.038}$	$0.073^{+0.028}_{-0.047}$	$0.269^{+0.027}_{-0.034}$
σ_{z6}	–	–	$0.188^{+0.056}_{-0.055}$
$\chi^2_{\rm red,Gauss}$	0.95	1.75	0.95

For the main results, we have chosen to only assign free parameters to the photometric redshift bins that have at least one adjacent spectroscopic redshift bin. For CFHTLenS-BOSS, this limits us to the five photometric redshift bins in the range $0.15 < z_B < 0.9$. In Fig. D1 of Appendix D, we show an example of the full parameter sampling for all eight photometric redshift bins for CFHTLenS-BOSS, where it is clear that there are degeneracies for the photometric redshift bins that do not have an adjacent spectroscopic bin. For CFHTLenS-BOSS, we fit five free parameters for shifting the $\Phi_j(z)$, obtaining values of Δz_j ranging from $-0.037^{+0.009}_{-0.010}$ for $j = 1$ to $0.049^{+0.010}_{-0.010}$ for $j = 4$. This best fit is represented by the cyan band in Fig. 3, where the width quantifies how the model changes for the uncertainties in the best-fitting model parameters.

The second method described in Section 2.5 is to fit Gaussian $\Phi_j(z)$, each with a mean, μ_{zj}, and a standard deviation, σ_{zj}. For CFHTLenS-BOSS, we fit 10 free parameters for Gaussian $\Phi_j(z)$. We use the minimum and maximum redshifts of each bin j defined in Section 3.1, $z_{{\rm min},j}$ and $z_{{\rm max},j}$, to impose the priors $z_{{\rm min},j} < \mu_{zj} < z_{{\rm max},j}$ and $0.001 < \sigma_{zj} < z_{{\rm max},j} - z_{{\rm min},j}$. These priors are chosen to avoid sampling the parameter space where the $\Phi_j(z)$ are extremely flat. The values are provided in Table 1. We do not fit the model to any data points corresponding to $z_B > 0.9$ because the outliers cannot be well modelled by a single Gaussian. In Fig. D2 of Appendix D, we show the MCMC sampling of the 10 free Gaussian $\Phi_j(z)$ parameters. Fig. D2 shows much stronger degeneracies than seen for the MCMC sampling of the shift method in Fig. 4. The best-fitting Gaussian model is represented by the purple band in Fig. 3. The large uncertainties are reflected in the width of the purple band.

We assess a goodness of fit using the reduced χ^2, which is the χ^2 given by equation (14) divided by the degrees of freedom (DOF). The naïve estimate of DOF is 387 = 392 data points – 5 fit parameters for the shifted $\Phi_j(z)$ and 214 = 224 data points – 10 fit parameters for the Gaussian $\Phi_j(z)$. However, see Andrae, Schulze-Hartung & Melchior (2010) for an argument that these DOF estimates should be considered as upper limits, and a standard reduced χ^2 is not an appropriate measure of the goodness of fit. As previously mentioned, we also issue the caveat that our covariance matrices are obtained via internal jack-knife resampling and may be somewhat biased. None the less, we quote the χ^2/DOF to illustrate the performance of the fitting: $\chi^2_{\rm no-shift} = 508.06/387 = 1.31$ for the original BPZ $\Phi_j(z)$, $\chi^2_{\rm shift} = 454.31/387 = 1.17$ for the BPZ $\Phi_j(z)$ after application of the best-fitting shifts, and $\chi^2_{\rm Gauss} = 202.32/214 = 0.95$ for the best-fitting Gaussian $\Phi_j(z)$. These values are also summarized in Table 1.

4.2 Comparison of photometric and spectroscopic redshifts for a bright CFHTLenS galaxy sample

In this section, we compare the original and improved CFHTLenS redshift distributions from our method with those obtained directly by examining galaxies that have both spectroscopic and photometric redshifts. For this comparison, we limit our photometric sample to bright galaxies with $i < 23$ to match the completeness of the comparison spectroscopic sample described in Section 3.2. We split the sample in four photometric redshift bins spanning $0.15 < z_B < 0.7$. The total number of spectroscopic redshifts with matched photometric redshifts in the range $0.15 < z_B < 0.7$ are 3925 (VVDS), 3031 (EGS), and 11 108 (VIPERS).

For each photometric redshift bin, we define a measure of bias $z_{\rm bias}$ to be the difference between the median redshift ($\tilde{z}$) determined from the spectroscopic redshifts and from the $\Phi_j(z)$. We choose the median as it is less sensitive to incompleteness in the spectroscopic

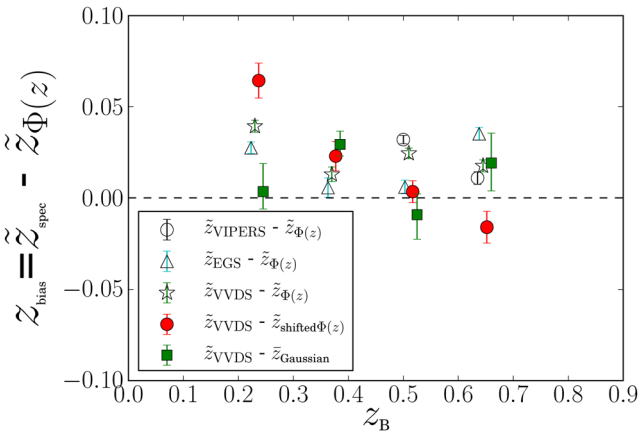

Figure 5. Direct comparison of the medians of the spectroscopic and photometric distributions for galaxies with $i < 23$. The open symbols show the difference between the medians of VIPERS (circles), DEEP2 EGS (triangles) or VVDS (stars) and the medians of the original BPZ $\Phi_j(z)$. The closed symbols show the difference between the median of VVDS and the $\Phi_j(z)$ after application of the best-fitting shifts (circles) or the best-fitting Gaussian $\Phi_j(z)$ (squares). The data points are slightly horizontally offset for clarity. For the cross-correlation analysis, the CFHTLenS galaxies were cut to $i < 23$ to be more consistent with the magnitude completeness limits of the comparison spectroscopic data sets.

sample at $z_{spec} > 1.3$. Error bars are determined by bootstrap resampling of the spectroscopic redshifts. Fig. 5 shows z_{bias} determined from each of the three spectroscopic surveys; VVDS (open star), VIPERS (open circle) and EGS (open triangle). Whilst these surveys are relatively complete, we have not investigated whether the sample is fully representative of the photometric sample, both in terms of colour–space coverage (see for example, Masters et al. 2015) and redshift coverage. Even with the applied bright magnitude limits, there is a noticeable amount of scatter even between the z_{bias} obtained from the two more complete surveys VVDS and DEEP2 EGS, suggesting some sample variance not taken into account by the bootstrap resampled error bars. Residual differences between the sampling of the galaxy populations can potentially be accounted for in future work using the re-weighting method of Bonnett et al. (2016). For all bins, however, we conclude that the photometric redshifts underestimate the true median redshift of the galaxy sample.

We next measure the CFHTLenS-BOSS cross-correlations for bright CFHTLenS galaxies with $i < 23$ and determine redshift offsets and best-fitting Gaussian distributions for the four redshift bins using our MCMC analysis. The resulting bias that we measure between the spectroscopic redshift distribution and our improved photometric redshift distribution is shown in Fig. 5 where we now use the median of the $\Phi_j(z)$ with best-fitting shifts applied and the median of the best-fitting Gaussian, respectively. We find that applying the best-fitting shift to the photometric redshift distributions leads to an even stronger underestimate of the true median redshift of the galaxy sample (closed circles) for the lower two redshift bins. However, the Gaussian model (closed squares) results in a z_{bias} that is consistent with zero. The total errors consist of the bootstrap resampled errors from the spectroscopic redshifts added in quadrature with the errors from the best-fitting values. The latter could potentially be slightly underestimated, as we ignore the covariance matrices between the spectroscopic bins.

From this analysis, we can conclude that modelling errors under the assumptions that the shape of the redshift distribution is accurate, and that the bias can be represented by a linear shift in the distribution is insufficient to capture the true underlying distribution when dealing with real data. Our linear-shift model, also advocated by The Dark Energy Survey Collaboration et al. (2015) and tested in Appendix A, was based on the well-modelled catastrophic outliers within CFHTLenS. It does, however, overlook the potential for the width of the main peak in the distribution to vary. For example, if the width were underestimated in the $\Phi_j(z)$, the model would underestimate the amplitude of the cross-correlation with both adjacent redshift bins. A shift, in contrast, appears as an underestimate of the signal in one of the adjacent bins, with an overestimate in the other adjacent bin. We see some evidence of this behaviour in the $i < 23$ sample.

Our Gaussian $\Phi_j(z)$ model allows us to test the effects of being able to change the width of the redshift distribution as well as the mean, with the caveat that our single-moded Gaussian $\Phi_j(z)$ are too simple to characterize catastrophic outliers. In this bright-galaxy measurement of bias in the main peak of the redshift distribution, the Gaussian model provides the most robust result.

4.3 Photometric redshift accuracy in RCSLenS

In this section, we repeat the cross-correlation analysis for the RCSLenS survey, but now with two spectroscopic surveys, BOSS (shown in Fig. 6) and WiggleZ (shown in Fig. 7). Comparing the measurements between RCSLenS and CFHTLenS (in Fig. 3), we find the amplitude and shape of the diagonal panels for the CFHTLenS and RCSLenS cross-correlations with BOSS are similar; however, these are qualitatively different from the RCSLenS cross-correlations with WiggleZ. This difference reflects the contrasting nature of the BOSS galaxy population with galaxy biases given by $b_{LOWZ} = 1.6$ (Chuang et al. 2016) and $b_{CMASS} = 1.9$ (Sánchez et al. 2014) compared with the WiggleZ galaxy population with $b_{WiggleZ} = 1.0$ (Blake et al. 2010). There are higher amplitudes of correlation between the BOSS galaxies (with their higher mean galaxy bias) and both CFHTLenS and RCSLenS galaxies compared to the lower amplitudes of correlation between the WiggleZ galaxies (with their lower mean galaxy bias) and RCSLenS galaxies. The typical signal to noise for RCSLenS–WiggleZ measurements is markedly less than for either the CFHTLenS–BOSS or the RCSLenS–BOSS cross-correlations.

In terms of the nature of the off-diagonal panels which reveal the strength of the photometric redshift errors, there is a difference between the measurements for CFHTLenS–BOSS and RCSLenS–BOSS highlighting the poorer quality of the four-band RCSLenS photometric redshifts. The RCSLenS–BOSS off-diagonal measurements show positive signals, particularly in the cross-correlations between the lowest spectroscopic redshift range and the highest photometric redshift ranges (top-right corner) and between the lowest photometric ranges and all spectroscopic ranges (left-most column). This signal indicates the significant presence of catastrophic outliers in the photometric redshifts for RCSLenS, which is consistent with the characterization of the photometric redshifts when compared directly against spectroscopic redshifts in RCSLenS as compared to CFHTLenS (Hildebrandt et al. 2016).

The orange band in Fig. 6 is the magnification signal computed in the same way as for Fig. 3 with α values of [0.61, 0.42, 0.68, 0.94, 0.85, 1.29, 1.53, 1.13]. In contrast to CFHTLenS, the α values are greater than one for the three highest redshift bins, leading to a positive magnification signal as seen in the top-right corner of Fig. 6. Focusing on $0.15 < z_s < 0.29$ cross-correlated with $1.1 < z_B < 1.3$, the orange band shows agreement with the measured

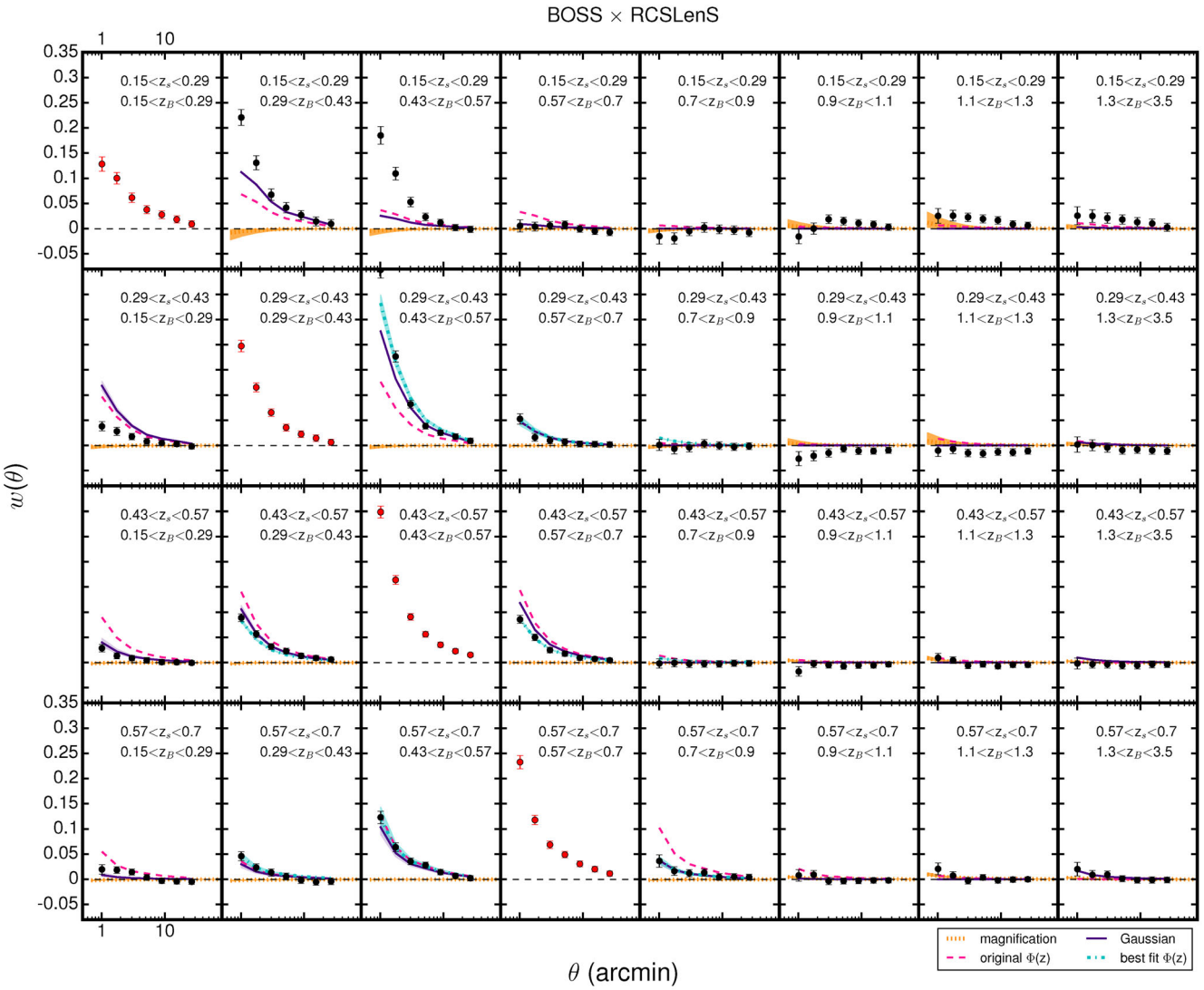

Figure 6. As in Fig. 3, cross-correlating RCSLenS (photometric redshift) bin j and BOSS (spectroscopic redshift) bin i galaxies. The data points are weighted means over the six RCSLenS fields that overlap BOSS, with the weights given by the number of pairs $(D_iD_j)_\theta$ (see equation 3).

cross-correlations for the first two θ bins, but the overall shape is qualitatively different. This is likely caused by an interplay between magnification and clustering of catastrophic outliers; the latter has not been accounted for in the $\Phi_j(z)$ and thus is not reflected in our model. There are no orange magnification prediction bands for Fig. 7, as the WiggleZ galaxies are typically less massive than the BOSS galaxies, and the amplitudes of their magnification signals are likely correspondingly lower in the one-halo regime. In the two-halo regime, only a couple of the cross-correlation measurements have a S/N over 2 (the highest is 2.7), thereby obviating the need for magnification modelling.

4.3.1 Improved RCSLenS photometric redshift distributions

As in Section 4.1.1, we apply the methods of Section 2.5 to RCSLenS and summarize the best-fitting values for each survey combination in Table 1. We again assign free parameters to the photometric redshift bins that have at least one adjacent spectroscopic bin. We impose a similar prior as in Section 4.1.1, except we extend the maximum possible shift to 0.3 for the $j = 4$ and $j = 5$ bins. For the shifting method, we additionally exclude the first redshift bin (the first row and column of Fig. 6) from our analysis due to the high catastrophic outlier rates indicated by the measurements. The BPZ $\Phi_j(z)$ (and hence our model) do not account for these catastrophic outlier rates and no amount of shifting will aid the cause. For RCSLenS-BOSS, we fit four free parameters for shifting the BPZ $\Phi_j(z)$ in the photometric redshift bins in the range $0.29 < z_B < 0.9$, obtaining values of Δz_j ranging from $-0.095^{+0.007}_{-0.007}$ for $j = 2$ to $0.236^{+0.026}_{-0.019}$ for $j = 4$. WiggleZ extends to slightly higher redshifts, thereby allowing us to fit five free parameters for the photometric redshift bins from 0.29 to 1.1 for the case of RCSLenS–WiggleZ, obtaining values of Δz_j ranging from $-0.040^{+0.029}_{-0.028}$ for $j = 6$ to $0.070^{+0.019}_{-0.019}$ for $j = 2$. The best-fitting shift model is again represented by the cyan band for RCSLenS–BOSS in Fig. 6 and for RCSLenS–WiggleZ in Fig. 7.

We also fit a Gaussian $\Phi_j(z)$, each with a mean, μ_{zj}, and a standard deviation, σ_{zj}. As for CFHTLenS, we do not fit the model to any data points corresponding to $z_B > 0.9$ for RCSLenS–BOSS and $z_B > 1.1$ for RCSLenS–WiggleZ because the outliers cannot

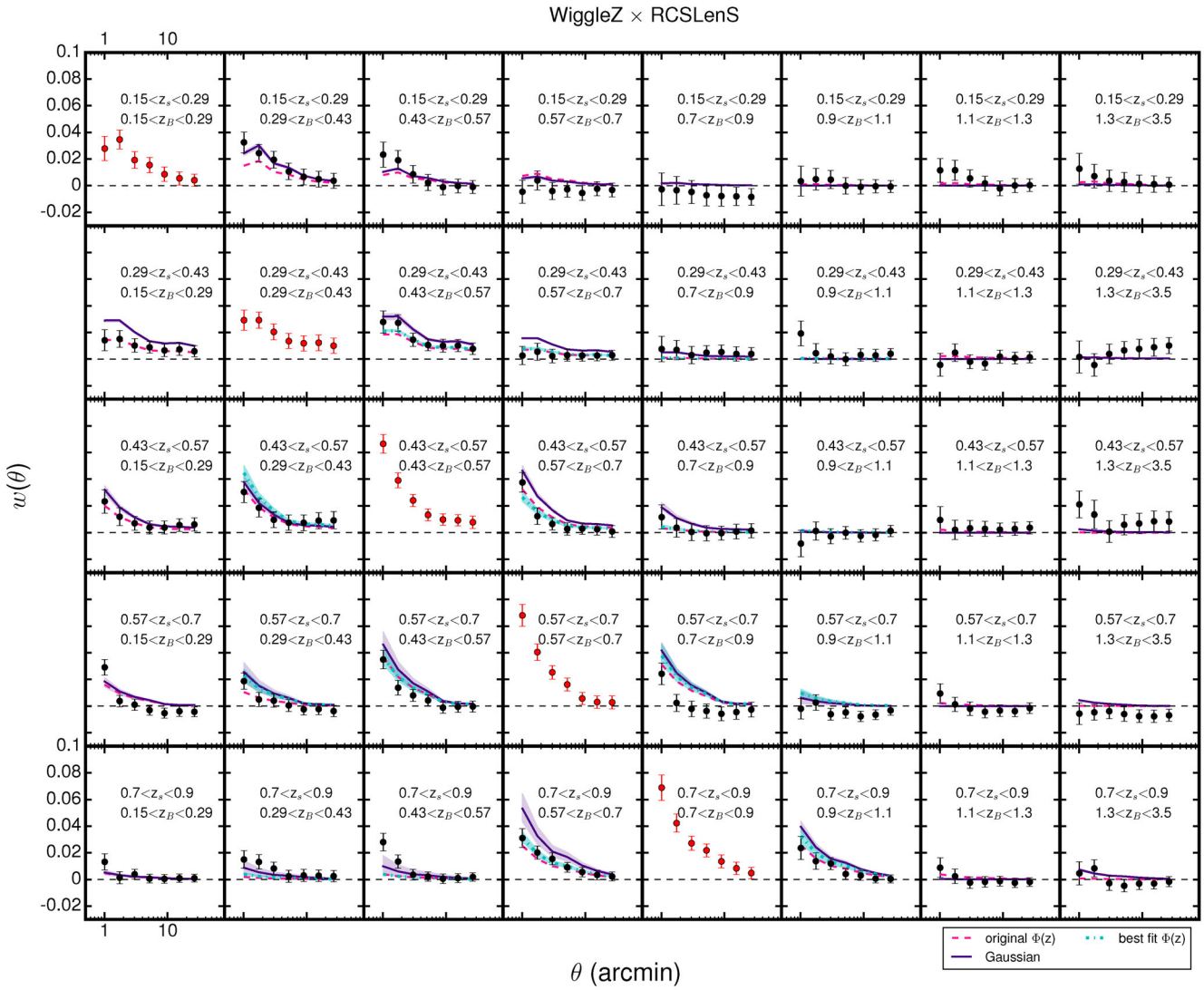

Figure 7. As in Fig. 3, cross-correlating RCSLenS (photometric redshift) bin j and WiggleZ (spectroscopic redshift) bin i galaxies. The data points are weighted means over the five RCSLenS fields that overlap WiggleZ, with the weights given by the number of pairs $(D_iD_j)_\theta$ (see equation 3). Note that there is no magnification prediction plotted here.

be well modelled by a single Gaussian. We impose similar priors to those described in Section 4.1.1, but we extend the maximum possible σ_{zj} by adding a value to account for photometric redshift scatter. Specifically, we apply the prior $0.001 < \sigma_{zj} < (z_{\mathrm{max},j} - z_{\mathrm{min},j}) + 0.08(1 + z_{\mathrm{mid},j})$ with $z_{\mathrm{mid},j}$ being the midpoint between $z_{\mathrm{min},j}$ and $z_{\mathrm{max},j}$. For RCSLenS–BOSS, we fit 10 free parameters. For RCSLenS–WiggleZ, we fit 12 free parameters. Both sets of best-fitting parameters are summarized in Table 1. The best-fitting Gaussian model is represented by the purple band for RCSLenS–BOSS in Fig. 6 and for RCSLenS–WiggleZ in Fig. 7. The best-fitting Gaussian model also fails to reproduce the large signals seen in the cross-correlation between $0.15 < z_s < 0.29$ and the three photometric redshift bins in the range $0.15 < z_B < 0.57$ in Fig. 6 (and to a lesser extent in Fig. 7).

For the sake of comparison, we provide the χ^2/DOF. The DOF for RCSLenS–BOSS is $752 = 756$ data points – 4 fit parameters for the shifted $\Phi_j(z)$ and $662 = 672$ data points – 10 fit parameters for the Gaussian $\Phi_j(z)$. The DOF for RCSLenS–WiggleZ is $835 = 840$ data points – 5 fit parameters for the shifted $\Phi_j(z)$ and $863 = 875$ data points – 12 fit parameters for the Gaussian $\Phi_j(z)$. The χ^2/DOF for RCSLenS-BOSS are: $\chi^2_{\mathrm{no-shift}} = 1280.44/752 = 1.70$ for the original BPZ $\Phi_j(z)$, $\chi^2_{\mathrm{shift}} = 967.93/752 = 1.29$ for the BPZ $\Phi_j(z)$ after application of the best-fitting shifts, and $\chi^2_{\mathrm{Gauss}} = 1155.4/662 = 1.75$ for the best-fitting Gaussian $\Phi_j(z)$. The χ^2/DOF for RCSLenS-WiggleZ are: $\chi^2_{\mathrm{no-shift}} = 907.84/835 = 1.09$ for the original BPZ $\Phi_j(z)$, $\chi^2_{\mathrm{shift}} = 891.59/835 = 1.07$ for the BPZ $\Phi_j(z)$ after application of the best-fitting shifts, and $\chi^2_{\mathrm{Gauss}} = 823.94/863 = 0.95$ for the best-fitting Gaussian $\Phi_j(z)$.

Fig. 8 shows the comparison of the best-fitting $\Phi_j(z)$ shifts and best-fitting Gaussian μ_{zj} and σ_{zj}. Ideally, the best-fitting parameters from the BOSS cross-correlation with RCSLenS would agree with those from the WiggleZ cross-correlation with RCSLenS. The best-fitting $\Phi_j(z)$ shifts are significantly discrepant, whilst there is better agreement for the best-fitting Gaussian parameters. We posit that the main source of this disagreement stems from our estimates of $\Phi_j(z)$ (from BPZ or Gaussians) being a poor characterization of the underlying redshift distribution. Future work might benefit from a more sophisticated model that retains the ability to model both the catastrophic outliers and the width of the redshift distribution.

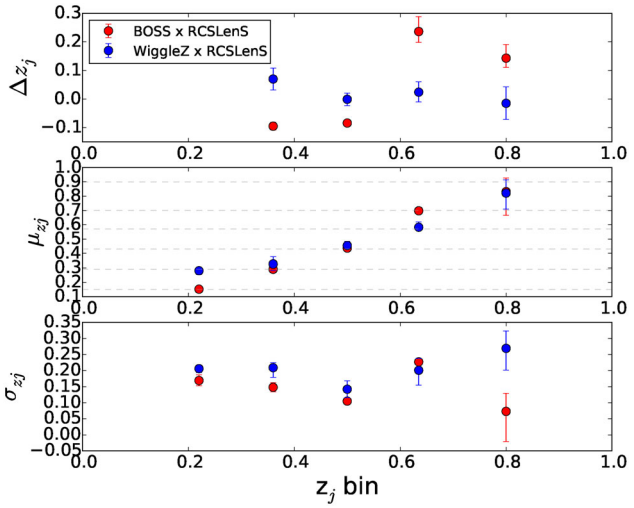

Figure 8. Comparison of best-fitting parameters for RCSLenS from the cross-correlation with BOSS and WiggleZ. The top panel shows the best-fitting shifts when the input models are the BPZ $\Phi_j(z)$. The lower two panels show the best-fitting means and standard deviations when the input models are Gaussians. The middle panel contains dashed grey lines to indicate the boundaries of the redshift bins. The values and errors are summarized in Table 1. Note that the plotted error bars correspond to 2σ, and there are no constraints for Δz_1, as explained in Section 4.3.1.

5 CONCLUSIONS

In this paper, we have extended the formalism of Benjamin et al. (2010) with the initial aim to verify galaxy redshift distributions of a sample of galaxies, as determined from the sum of their photometric redshift probability distributions, $\Phi_j(z)$. By cross-correlating the galaxy positions of different photometric redshift bins, Benjamin et al. (2013) showed that the measured galaxy clustering between photometric redshift bins was consistent with the level of clustering expected when the redshift distributions are estimated from BPZ in this way. For the cosmological analyses of CFHTLenS that then followed (Benjamin et al. 2013; Heymans et al. 2013; Kilbinger et al. 2013; Simpson et al. 2013; Fu et al. 2014; Kitching et al. 2014), the photometric error distribution was therefore assumed to be known with zero uncertainty. In this analysis, we have increased the fidelity of this test by using overlapping spectroscopy from BOSS and, in the case of the RCSLenS, also from WiggleZ. Cross-correlating photometrically selected galaxy samples with galaxies binned by spectroscopic redshift significantly enhances the signal to noise in the measured clustering between different redshift bins, thus allowing for a more stringent test of the photometrically derived redshift distributions. For CFHTLenS, we can draw the same conclusions as Benjamin et al. (2013), that the catastrophic outlier rate is well predicted by the $\Phi_j(z)$. This can be seen by comparing the measured clustering signal with the model prediction, for widely separated bins ($\Delta z > 0.2$), in Fig. 3. Where the model is seen to fail, however, is around the peak of the redshift distribution, where significant deviations between the signal and model are seen for the first and fourth CFHTLenS redshift bin with $0.15 < z_B < 0.29$ and $0.57 < z_B < 0.7$. A direct comparison between spectroscopic and photometric redshifts for a bright sample in Fig. 5 also indicates that the $\Phi_j(z)$ model also fails in terms of the width or scatter in the redshift distribution. This analysis implies that the cosmological analyses of CFHTLenS should have included systematic error terms in their analysis to account for bias and scatter in their redshift distributions that were not accurately modelled by the $\Phi_j(z)$.

For RCSLenS, the conclusion that we can draw from this type of analysis about the catastrophic outlier rate is complicated by the fact that the lowest photometric redshift bin ($0.15 < z_B < 0.29$) exhibits a strong catastrophic outlier rate when galaxy photometric redshifts are individually compared to their directly measured spectroscopic redshifts (Hildebrandt et al. 2016). This outlier rate is not predicted by the $\Phi_j(z)$ in this redshift range. Model predictions for the cross-correlation of this low-redshift bin with higher photometric redshift bins will therefore be incorrect, as they depend on the measured autocorrelation signal in this low-redshift bin (see equation 13). A disagreement between model and signal for cross-correlations with this bin will therefore exist, even if the catastrophic outlier rate is accurately represented in the higher redshift bins, as suggested by the direct comparison in Hildebrandt et al. (2016). Excluding this low-redshift bin from our analysis, we find that the catastrophic errors scattering galaxies from high redshift down to $z > 0.29$ are well represented by the $\Phi_j(z)$. Around the peak of each redshift distribution, however, significant deviations are again found, this time for the four bins spanning $0.29 < z_B < 0.9$.

As discussed in Section 2.5, we can use our derived formalism to determine joint offsets in the peaks of the tomographic redshift distributions which are close in redshift to a spectroscopic sample. To undertake this analysis, we need to assume that the overall shape of each distribution is sufficiently accurate (i.e. that the catastrophic outliers are well represented). The offsets for each redshift bin that satisfy these constraints are given in Table 1, showing significant biases up to $\Delta z \sim 0.236$. We found that whilst the catastrophic outliers were well modelled by the $\Phi_j(z)$, the scatter was not, leading to inconsistencies when directly comparing our results with deep spectroscopic surveys in Section 4.2 and when comparing results between the BOSS and WiggleZ surveys. We therefore also determine the best-fitting Gaussian $\Phi_j(z)$ for each bin which provides a more accurate estimate of the bias and scatter in each photometric redshift bin. By definition however, this single-moded Gaussian is unable to model catastrophic outliers.

We have investigated the influence of the astrophysical features of galaxy bias and magnification, as well as the systematic effects of object detection and deblending. We found that these features do not impact upon the conclusions presented here but will need to be investigated in more detail for future studies seeking to draw tighter constraints on redshift distributions. With the expected signal to noise of upcoming deeper surveys, magnification may play a more significant role in the angular cross-correlation signal between bins widely separated in redshift. Complete simulation pipelines including a full picture of the underlying physics (clustering and lensing) and the observing, object detection and cataloguing process will be necessary to fully understand and disentangle the physical and systematic effects.

5.1 Impact on cosmological parameter estimation with CFHTLensS

Our methodology has been shown to provide a robust tool to verify redshift distributions for photometric surveys where overlapping but incomplete spectroscopy exists. The recent existence of this 'same-sky' survey data has allowed us to test the photometric redshift distributions used in the CFHTLenS weak lensing analyses with much higher fidelity than was previously possible. We can use the results of our analysis and scaling-relations from Jain & Seljak (1997) to estimate the impact on cosmological parameter estimation

from using inaccurate redshift distributions in previous CFHTLenS analyses, for example, Heymans et al. (2013).

Weak lensing is most sensitive to a combination of the clustering amplitude σ_8 and the matter density parameter Ω_m. Defining $S_8 = \sigma_8 \Omega_m^{0.47}$, the two-point shear correlation function ξ_+ for a flat Λcold dark matter cosmology is related to S_8 as

$$\xi_+ \propto z_s^{1.52} S_8^{2.58}, \quad (15)$$

for a single-lensed source redshift slice at z_s (Jain & Seljak 1997). We can use this as a 'toy-model' to indicate how errors in the source redshift propagates into biases on cosmological parameters. Considering the largest correction from Table 1, a bias of 0.049 in the photometric redshift bin spanning $0.57 < z_B < 0.7$ would correspond to an overestimate in the recovered S_8 parameter from this tomographic bin by 4 per cent. The second largest correction from Table 1 is for the lowest redshift bin which was already excluded from all CFHTLenS analyses as a result of concerns over the photometric redshift accuracy in this bin (Hildebrandt et al. 2012). We refer the reader to Joudaki et al. (2016) where the CFHTLenS tomographic cosmological analysis is revisited, taking into account the photometric redshift errors uncovered in this work.

ACKNOWLEDGEMENTS

We are grateful to the RCS2 team for planning the survey, applying for observing time, and conducting the observations. We acknowledge use of the Canadian Astronomy Data Centre operated by the Dominion Astrophysical Observatory for the National Research Council of Canada's Herzberg Institute of Astrophysics. We thank Marcello Cacciato and Maciek Bilicki for helpful comments on the manuscript and Chris Morrison, Sarah Bridle and Marika Asgari for useful discussions. AC and CH acknowledge support from the European Research Council under FP7 grant number 240185. Part of this work was developed while at the Aspen Center for Physics (NSF grant 1066293), and AC thanks the Center for their hospitality. CB acknowledges the support of the Australian Research Council through the award of a Future Fellowship. HH is supported by an Emmy Noether grant (No. Hi 1495/2-1) of the Deutsche Forschungsgemeinschaft. RN acknowledges support from the German Federal Ministry for Economic Affairs and Energy (BMWi) provided via DLR under project no.50QE1103. MV acknowledges support from the European Research Council under FP7 grant number 279396 and the Netherlands Organization for Scientific Research (NWO) through grants 614.001.103.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is http://www.sdss3.org/.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

Author Contributions: All authors contributed to the development and writing of this paper. The authorship list reflects the lead authors of this paper (AC, CH) followed by two alphabetical groups. The first alphabetical group includes key contributors to the science analysis and interpretation in this paper. CB made the angular clustering measurements, and HH led the RCSLenS collaboration. The second group consists of the founding core team and those whose significant efforts produced the final RCSLenS data product or data products used in this analysis.

REFERENCES

Abrahamse A., Knox L., Schmidt S., Thorman P., Tyson J. A., Zhan H., 2011, ApJ, 734, 36
Ahn C. P. et al., 2014, ApJS, 211, 17
Anderson L. et al., 2014, MNRAS, 441, 24
Andrae R., Schulze-Hartung T., Melchior P., 2010, preprint (arXiv:e-prints)
Applegate D. E. et al., 2014, MNRAS, 439, 48
Benítez N., 2000, ApJ, 536, 571
Benjamin J., van Waerbeke L., Ménard B., Kilbinger M., 2010, MNRAS, 408, 1168
Benjamin J. et al., 2013, MNRAS, 431, 1547
Bernstein G., Huterer D., 2010, MNRAS, 401, 1399
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Blake C. et al., 2010, MNRAS, 406, 803
Blake C. et al., 2016, MNRAS, 456, 2806
Bonnett C. et al., 2016, Phys. Rev. D, 94, 042005
Bordoloi R. et al., 2012, MNRAS, 421, 1671
Cabré A., Fosalba P., Gaztañaga E., Manera M., 2007, MNRAS, 381, 1347
Chuang C.-H. et al., 2016, MNRAS, 461, 3781
Cunha C. E., Lima M., Oyaizu H., Frieman J., Lin H., 2009, MNRAS, 396, 2379
Cunha C. E., Huterer D., Lin H., Busha M. T., Wechsler R. H., 2014, MNRAS, 444, 129
Davis M. et al., 2007, ApJ, 660, L1
de Putter R., Doré O., Das S., 2014, ApJ, 780, 185
Drinkwater M. J. et al., 2010, MNRAS, 401, 1429
Duncan C. A. J., Joachimi B., Heavens A. F., Heymans C., Hildebrandt H., 2014, MNRAS, 437, 2471
Eisenstein D. J. et al., 2011, AJ, 142, 72
Erben T. et al., 2009, A&A, 493, 1197
Erben T. et al., 2013, MNRAS, 433, 2545
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Foreman-Mackey D., Price-Whelan A., Ryan G., Emily Smith M., Barbary K., Hogg D. W., Brewer B. J., 2014, triangle.py v0.1.1
Fu L. et al., 2014, MNRAS, 441, 2725
Gilbank D. G., Gladders M. D., Yee H. K. C., Hsieh B. C., 2011, AJ, 141, 94
Goodman J., Weare J., 2010, Commun. Appl. Math. Comp. Sci., 5, 65
Guzzo L. et al., 2014, A&A, 566, A108
Harnois-Déraps J., Vafaei S., Van Waerbeke L., 2012, MNRAS, 426, 1262
Heavens A. F., Joachimi B., 2011, MNRAS, 415, 1681
Heymans C. et al., 2012, MNRAS, 427, 146
Heymans C. et al., 2013, MNRAS, 432, 2433
Hildebrandt H., van Waerbeke L., Erben T., 2009, A&A, 507, 683
Hildebrandt H. et al., 2012, MNRAS, 421, 2355
Hildebrandt H. et al., 2016, MNRAS
Jain B., Seljak U., 1997, ApJ, 484, 560
Joudaki S. et al., 2016, preprint (arXiv:e-prints)
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Kitching T. D. et al., 2014, MNRAS, 442, 1326
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Le Fèvre O. et al., 2005, A&A, 439, 845

Mandelbaum R. et al., 2008, MNRAS, 386, 781
Masters D. et al., 2015, preprint (arXiv: e-prints)
Matthews D. J., Newman J. A., 2010, ApJ, 721, 456
McQuinn M., White M., 2013, MNRAS, 433, 2857
Melchior P. et al., 2015, MNRAS, 449, 2219
Ménard B., Scranton R., Fukugita M., Richards G., 2010, MNRAS, 405, 1025
Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint (arXiv:e-prints)
Miller L. et al., 2013, MNRAS, 429, 2858
Miyatake H. et al., 2013, preprint (arXiv:e-prints)
Moessner R., Jain B., 1998, MNRAS, 294, L18
More S., Miyatake H., Mandelbaum R., Takada M., Spergel D. N., Brownstein J. R., Schneider D. P., 2015, ApJ, 806, 2
Morrison C. B., Hildebrandt H., 2015, MNRAS, 454, 3121
Morrison C. B., Scranton R., Ménard B., Schmidt S. J., Tyson J. A., Ryan R., Choi A., Wittman D. M., 2012, MNRAS, 426, 2489
Nakajima R., Mandelbaum R., Seljak U., Cohn J. D., Reyes R., Cool R., 2012, MNRAS, 420, 3240
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Newman J. A., 2008, ApJ, 684, 88
Newman J. A. et al., 2013, ApJS, 208, 5
Newman J. A. et al., 2015, Astropart. Phys., 63, 81
Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, MNRAS, 396, 19
Parejko J. K. et al., 2013, MNRAS, 429, 98
Peebles P. J. E., 1973, ApJ, 185, 413
Planck Collaboration et al., 2014, A&A, 571, A16
Rahman M., Ménard B., Scranton R., Schmidt S. J., Morrison C. B., 2015, MNRAS, 447, 3500
Sánchez A. G. et al., 2014, MNRAS, 440, 2692
Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, MNRAS, 431, 3307
Schneider M., Knox L., Zhan H., Connolly A., 2006, ApJ, 651, 14
Schulz A. E., 2010, ApJ, 724, 1305
Scranton R. et al., 2005, ApJ, 633, 589
Seljak U., 2000, MNRAS, 318, 203
Sheldon E. S., Cunha C. E., Mandelbaum R., Brinkmann J., Weaver B. A., 2012, ApJS, 201, 32
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Simet M., Mandelbaum R., 2015, MNRAS, 449, 1259
Simpson F. et al., 2013, MNRAS, 429, 2249
Suchyta E. et al., 2016, MNRAS, 457, 786
The Dark Energy Survey Collaboration et al., 2015, preprint (arXiv:e-prints)
Velander M. et al., 2014, MNRAS, 437, 2111
Wittman D., 2009, ApJ, 700, L174
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

APPENDIX A: VALIDATION TESTS ON MOCK GALAXIES

We test the methodology described in Sections 2.2–2.5 on an idealized case consisting of lognormal distributions of 1000 000 sources of known clustering in a lightcone of area 25 deg^2, with an underlying Gaussian $\Phi_j(z)$ with mean of 0.4 and a standard deviation of 0.3, cut in the range $0.1 < z < 0.7$. 10 000 sources are randomly picked as spectroscopic 'BOSS' sources. 200 000 sources are randomly picked as photometric redshift 'CFHTLenS' sources. The photometric redshift sources are assigned Gaussian scatters with mean of 0 and a standard deviation of 0.05. To simulate BPZ-like redshift probability distributions, each source is assigned a $p(z)$, peaking at the scattered photometric redshift value. The spectroscopic and photometric sources are divided into six redshift bins of width 0.1 in the range $0.1 < z < 0.7$, and cross-correlations are measured. We followed the steps outlined in Section 2.3 and show the measured cross-correlations compared with the predicted cross-correlations in Fig. A1. Fig. A2 shows the results of MCMC sampling shifts of the redshift distributions for each of the bins (described in Section 2.5. The best-fitting shifts are consistent with 0, as was input into the mocks, thus validating our method in this idealized scenario.

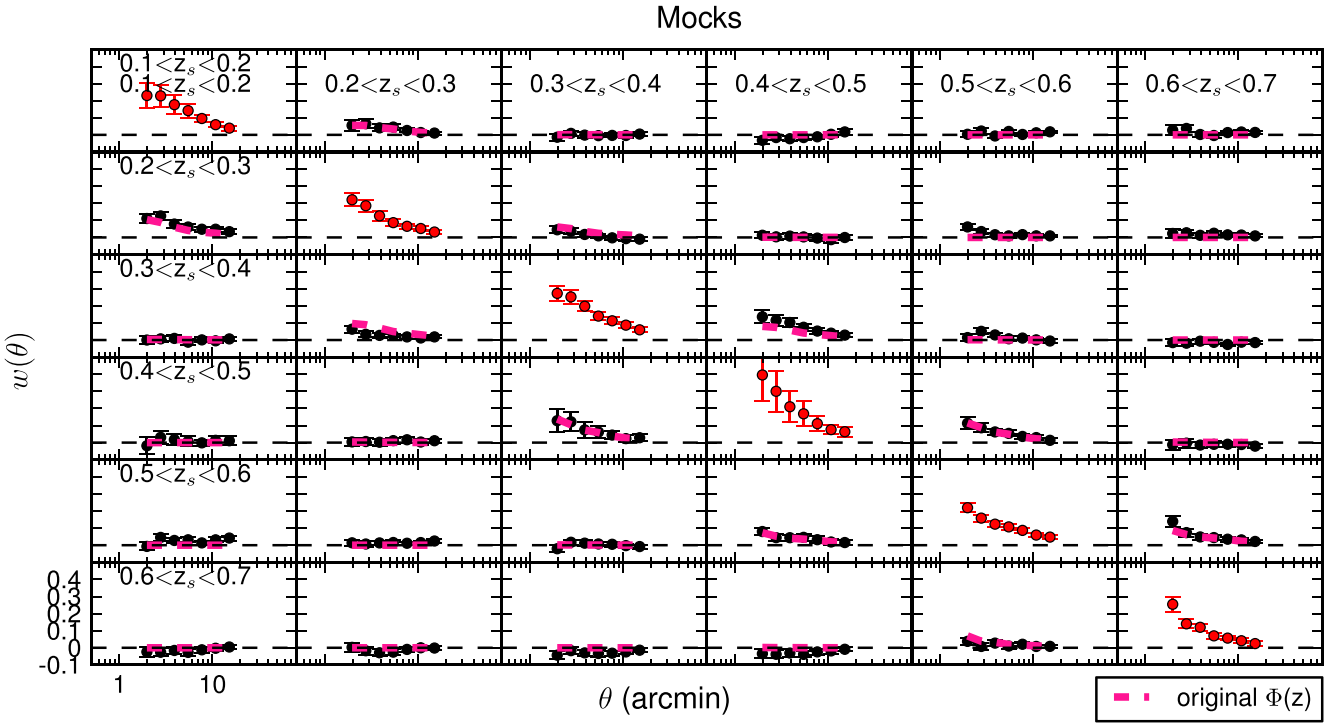

Figure A1. Mock analysis: the measured cross-correlations between the mock photometric galaxies and the mock spectroscopic galaxies.

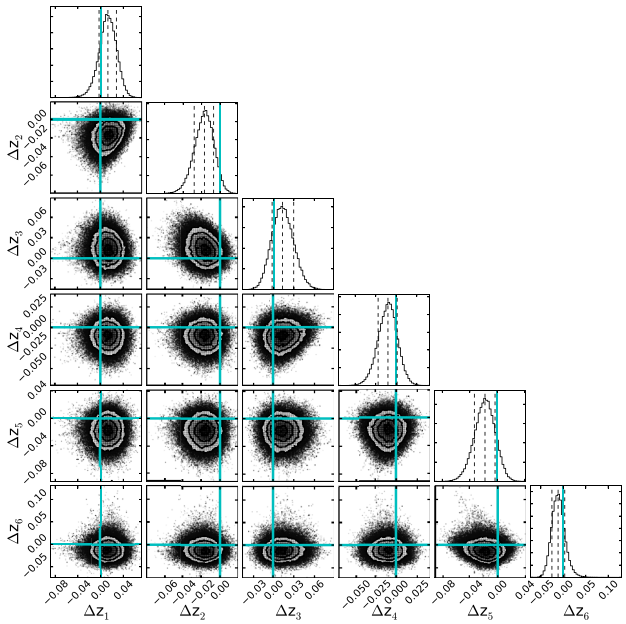

Figure A2. Mock-analysis: best-fitting shifts in the redshift distributions for six photometric redshift bins using joint MCMC sampling.

APPENDIX B: HALO MODEL

The halo model provides an analytic framework for describing how galaxies occupy dark matter haloes, and the CHOMP software used in this work follows the formalism of Seljak (2000). We assume cosmological parameters from Planck Collaboration et al. (2014), with $\Omega_{\rm m} = 0.315$, $\Omega_\Lambda = 0.685$, $\sigma_8 = 0.829$, $n_{\rm s} = 0.9603$ and $\Omega_{\rm b}h^2 = 0.022\,05$. Using the mass function from Sheth & Tormen (2002), the density profile from Navarro, Frenk & White (1996), and a halo bias model, we can derive correlation functions. To describe the numbers of central and satellite galaxies as a function of halo mass, we assume the functional forms given in Zheng, Coil & Zehavi (2007), using the best-fitting parameters from Parejko et al. (2013) for the LOWZ sample and v1 of Miyatake et al. (2013) for the CMASS sample into these functional forms. The number of central galaxies is given by

$$N_c(M) = \frac{1}{2}\left[1 + {\rm erf}\left(\frac{\log({\rm M}) - \log({\rm M}_{\rm min})}{\sigma_{\log {\rm M}}}\right)\right], \quad \text{(B1)}$$

where $M_{\rm min}$ is the minimum mass for a halo to host a central galaxy, and $\sigma_{\log M}$ is the width of the cutoff. The satellite term follows

$$N_{\rm s}(M) = N_{\rm c}(M)\left(\frac{M - M_0}{M_1^{'}}\right)^{\alpha}, \quad \text{(B2)}$$

where M_0 is the minimum mass for a halo to host satellite galaxies, and $M_1^{'}$ is the mass differential at which a halo is expected to have one satellite galaxy. Adding the central and satellite terms together gives the average number of galaxies occupying a halo of mass M.

The LOWZ parameters, taken from table 3 of Parejko et al. (2013), are $\log 10(M_{\rm min}/{\rm M}_\odot) = 13.25 \pm 0.26$, $\log 10(M_1^{'}/{\rm M}_\odot) = 14.18 \pm 0.39$, $\sigma_{\log M} = 0.98 \pm 0.57$, $\kappa = 1.04 \pm 0.71$, and $\alpha = 0.94 \pm 0.49$. Their κ corresponds to our $M_0/M_{\rm min}$. The CMASS parameters were originally taken from table 2 of v1 of Miyatake et al. (2013) as $\log 10(M_{\rm min}/{\rm M}_\odot) = 13.21^{+0.13}_{-0.11}$, $\log 10(M_1^{'}/{\rm M}_\odot) = 14.15^{+0.09}_{-0.08}$, $\sigma_{\log M} = 0.56^{+0.11}_{-0.09}$, $\kappa < 0.58$, and $\alpha = 1.06 \pm 0.49^{+0.11}_{-0.13}$. The final published CMASS parameters appear in table 1 of More et al. (2015) for three stellar mass sub-samples. The parameter values we use in this work fall within the range of values spanned by the three stellar mass sub-samples in More et al. (2015).

APPENDIX C: EFFECTS OF OBJECT DETECTION AND DEBLENDING ON CLUSTERING MEASUREMENTS

Clustering signals can contain not only contributions from physical phenomena like spatial correlation and magnification but also from systematic effects from the object detection and selection process itself. In this section, we first confirm that not accounting for small-scale (<9 arcsec) selection features in the random catalogues can affect the measured clustering on a range of scales; we next investigate how object detection and deblending might affect the number of faint photometric galaxies in the vicinity of bright spectroscopic galaxies; finally, we perform a brute-force check by re-running our analysis on the data with photometric masks placed over the locations of known spectroscopic galaxies.

First, we set up a toy experiment to investigate how small-scale (<9.3 arcsec) selection features in the random catalogues can affect the measured clustering on a range of scales by creating three mock galaxy catalogues based on the public N-body simulations described in Harnois-Déraps, Vafaei & Van Waerbeke (2012):[8] a 'BOSS-like' clustered galaxy catalogue with $b_{\rm g} = 2$ and a number density of 1 arcmin^{-2}, a 'photometric-like' clustered galaxy catalogue with $b_{\rm g} = 1$ and a number density of 10 arcmin^{-2}, and a 'no-clustering' galaxy catalogue with Poisson distribution and a number density of 10 arcmin^{-2}. For the photometric-like (PHOT) and no-clustering (NC) galaxy catalogues, we have full versions and masked versions where all of the PHOT/NC galaxies within a 9.3 arcsec radius of a galaxy from the BOSS-like catalogue are cut out. The 9.3 arcsec radius is convenient because it corresponds to 2 pixels in the original N-body simulation. All galaxies are at a single redshift $z = 0.525$. We measure the $w(\theta)$ using equation (3) for six different combinations quoted as [data sample 2, random sample 2]. Data sample 1 is always the BOSS-like catalogue, and random sample 1 is always a full (unmasked) NC sub-sample. Where there are multiple sub-samples from the NC galaxy catalogue used in the $w(\theta)$ measurement (e.g. if both data sample 2 and random sample 2 are both drawn from the NC galaxy catalogue), we ensure that the sub-samples are mutually exclusive and do not share any of the same galaxies. The combinations are: (1) [full PHOT, full NC], (2) [masked PHOT, full NC], (3) [masked PHOT, masked NC], (4) [full NC, full NC], (5) [masked NC, full NC], and (6) [masked NC, masked NC]. The results are shown in Fig. C1 where the combinations go from top left to top right and continue from bottom left to bottom right. The left-most and right-most columns show the true cross-correlations when the random catalogues properly account for the properties of the data catalogues. The top-middle panel shows a dip at small θ in the cross-correlation between a BOSS-like sample and a masked PHOT sample with a corresponding random sample that is not masked. Similarly, the lower-middle panel shows an anti-correlation for the $w(\theta)$ between a BOSS-like sample and a masked NC sample with a corresponding random sample that is not masked.

[8] http://www.cfhtlens.org

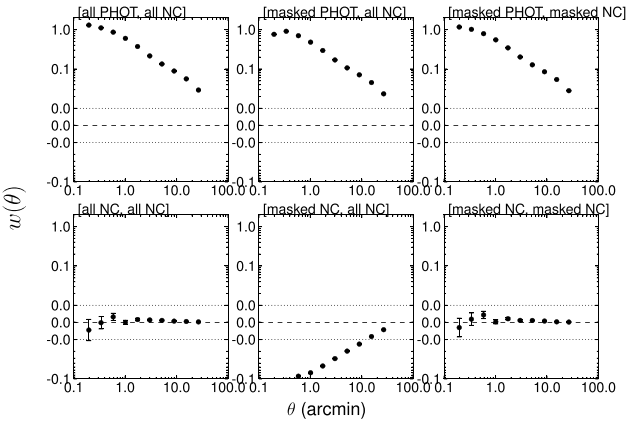

Figure C1. Clustering measurements for different combinations of data and random mocks. All combinations are measured with equation (3) and involve a BOSS-like catalogue (data sample 1) and a corresponding random catalogue (random sample 1) that reflects the geometry of the BOSS-like catalogue. The different combinations are labelled as [data sample 2, random sample 2], where the samples are drawn from either 'masked' or 'full' photometric-like (PHOT) or a no-clustering (NC) catalogues as described in the text. The left-hand and right-hand columns of panels illustrate the case where the random sample properly accounts for the masking of the data sample. However, the middle column of panels shows that the signal is damped if masked galaxies in the data sample are not correspondingly masked in the random sample.

The small-scale selection effects of object detection and deblending have been studied in the context of cluster galaxies which (typically bright, large in size, and residing in crowded fields) obscure nearby fainter or smaller galaxies (Applegate et al. 2014; Melchior et al. 2015; Simet & Mandelbaum 2015). The obscuration could, in principle, lead to a dearth of objects detected close to the brighter galaxies and masquerade as a spurious magnification signal. Because the impact is uncertain, these selection effects are not included in the CFHTLenS/RCSLenS masks. We investigate possible systematic contributions to the clustering signal around BOSS galaxies from the cataloguing process by running simulations with BALROG.[9] This public software allows us to add simulated galaxies to CFHTLenS images around known BOSS galaxies and run SExtractor with the same object detection and deblending parameters used for the actual catalogues. We can repeat the process many times and measure the recovered fraction of objects as a function of the angular separation from the BOSS galaxies. The results are shown in Fig. C2, and a horizontal line marks the average level at which objects at any location in an image can be recovered (due to noise). Fig. C2 indicates that there is a lower fraction of objects recovered at angular scales smaller than 20 arcsec.

The conclusion we can draw from this analysis is that if the random catalogue does not include small-scale selection effects, the measured $w(\theta)$ will be diluted out to very large scales. We re-run the analysis with photometric catalogues that have 1 arcmin and 30 arcsec masks centred on the BOSS galaxy positions in both the data and random catalogues. For both mask radii, the S/N renders the masked $w(\theta)$ consistent with the $w(\theta)$ measured before masking. Therefore, we conclude that this is not a dominant systematic in our analysis but flag it as an important systematic in the future. We leave further tests to future work. Much more sophisticated image simulations and mock galaxy catalogues will be required to fully disentangle physical effects (magnification) from systematic effects introduced in the observation and measurement process.

[9] https://github.com/emhuff/Balrog

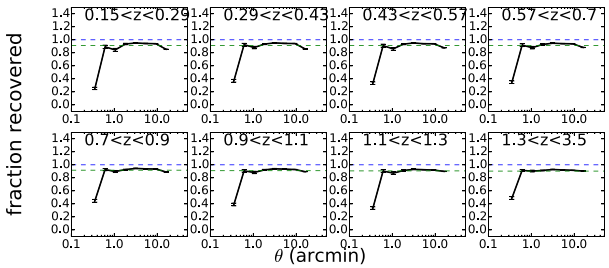

Figure C2. Fraction of recovered objects from simulations.

APPENDIX D: FULL MCMC FOR CFHTLENS–BOSS

In Fig. D1, we show the results of MCMC sampling the redshift bias parameters corresponding to all eight redshift bins ([0.15, 0.29], [0.29, 0.43], [0.43, 0.57], [0.57, 0.7], [0.7, 0.9], [0.9, 1.1], [1.1, 1.3], [1.3, 3.5]). This example corresponds to the overlap between CFHTLenS and BOSS. As the spectroscopic coverage of BOSS only extends to the first four redshift bins, there are clear degeneracies in the lower half of Fig. D1 which reflect the fact that the spectroscopic redshifts up to $z\sim0.7$ can offer only limited information to constrain the higher redshift photometric galaxies.

In Fig. D2, we show the results of MCMC sampling Gaussian $\Phi_j(z)$ described by their means μ_{zj} and standard deviations σ_{zj} for the same data set.

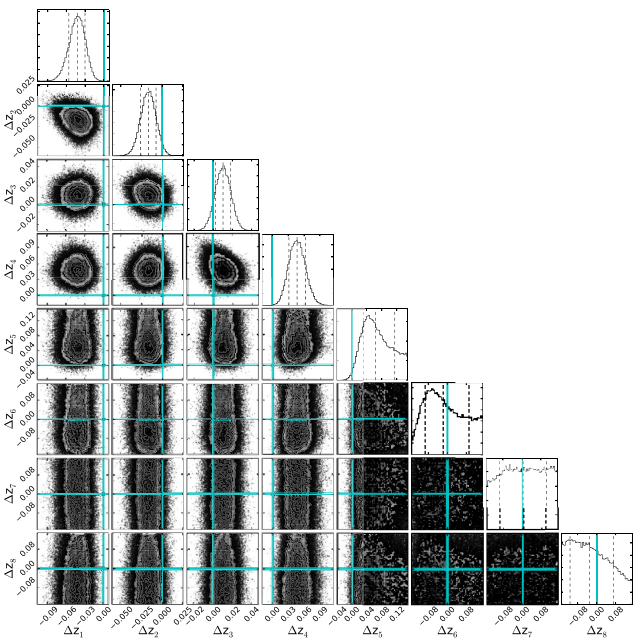

Figure D1. MCMC sampling of the redshift bias parameters for all eight redshift bins for CFHTLenS-BOSS.

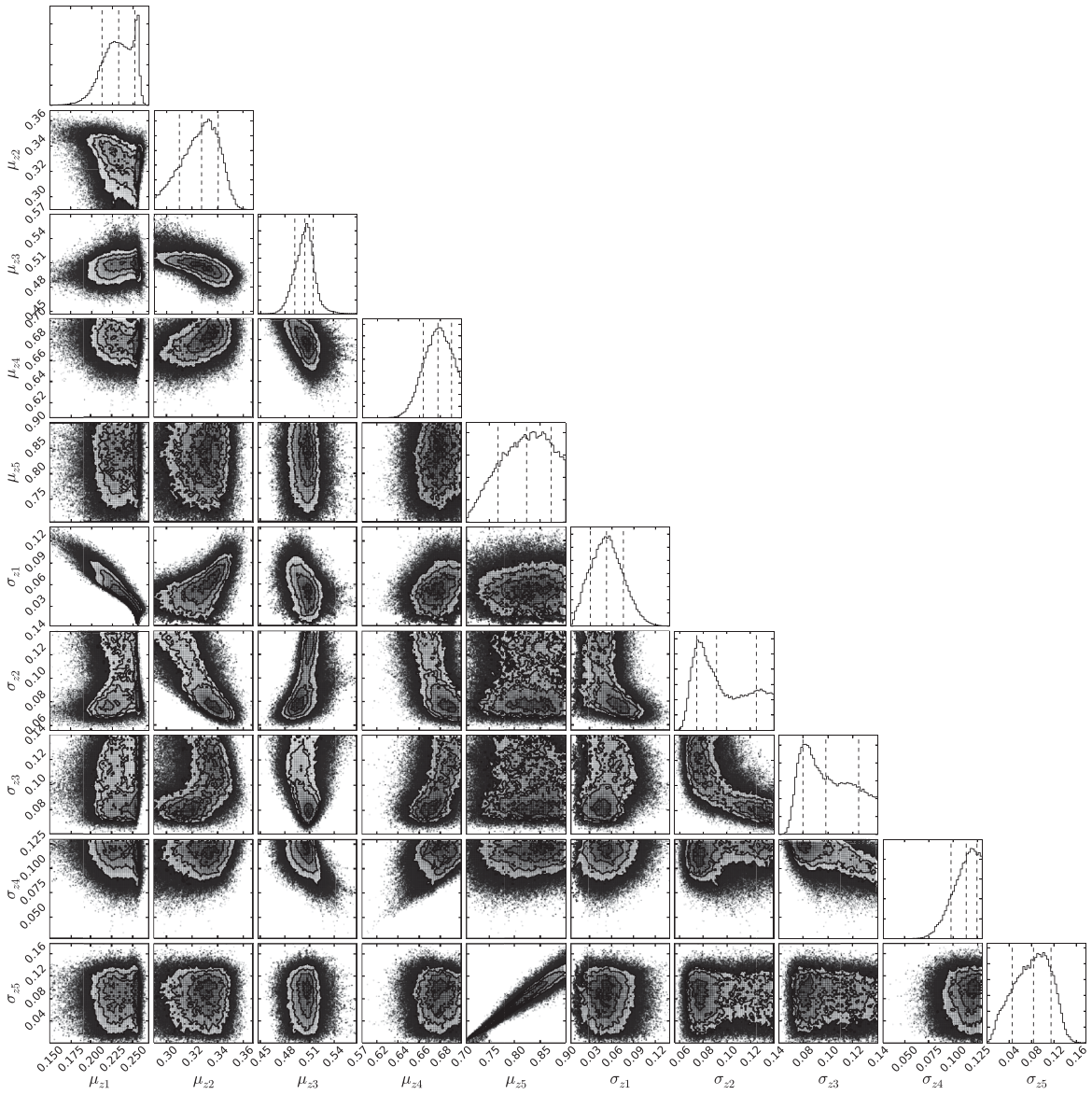

Figure D2. MCMC sampling of the best-fitting Gaussian $\Phi_j(z)$ for five redshift bins for CFHTLenS–BOSS. The redshift bin limits are given by [0.15, 0.29], [0.29, 0.43], [0.43, 0.57], [0.57, 0.7] and [0.7, 0.9].

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