

# Dark Energy Survey Year 1 results: validation of weak lensing cluster member contamination estimates from $P(z)$ decomposition

T. N. Varga,<sup>1,2★</sup> J. DeRose,<sup>3,4,5</sup> D. Gruen,<sup>1b,3,4,5</sup> T. McClintock,<sup>1b,6</sup> S. Seitz,<sup>1,2</sup> E. Rozo,<sup>7</sup> M. Costanzi,<sup>2</sup> B. Hoyle,<sup>1,2</sup> N. MacCrann,<sup>8,9</sup> A. A. Plazas,<sup>10</sup> E. S. Rykoff,<sup>4,5</sup> M. Simet,<sup>11,12</sup> A. von der Linden,<sup>13</sup> R. H. Wechsler,<sup>3,4,5</sup> J. Annis,<sup>14</sup> S. Avila,<sup>15</sup> E. Bertin,<sup>16,17</sup> D. Brooks,<sup>18</sup> E. Buckley-Geer,<sup>14</sup> D. L. Burke,<sup>4,5</sup> A. Carnero Rosell,<sup>19,20</sup> M. Carrasco Kind,<sup>21,22</sup> J. Carretero,<sup>23</sup> C. E. Cunha,<sup>4</sup> C. B. D'Andrea,<sup>24</sup> L. N. da Costa,<sup>20,25</sup> J. De Vicente,<sup>19</sup> S. Desai,<sup>26</sup> H. T. Diehl,<sup>14</sup> J. P. Dietrich,<sup>27,28</sup> P. Doel,<sup>18</sup> A. E. Evrard,<sup>29,30</sup> B. Flaugher,<sup>14</sup> P. Fosalba,<sup>31,32</sup> J. Frieman,<sup>14,33</sup> J. García-Bellido,<sup>34</sup> E. Gaztanaga,<sup>31,32</sup> D. W. Gerdes,<sup>29,30</sup> R. A. Gruendl,<sup>21,22</sup> J. Gschwend,<sup>20,25</sup> G. Gutierrez,<sup>14</sup> W. G. Hartley,<sup>18,35</sup> D. L. Hollowood,<sup>36</sup> K. Honscheid,<sup>8,9</sup> D. J. James,<sup>37</sup> T. Jeltema,<sup>36</sup> K. Kuehn,<sup>38</sup> N. Kuropatkin,<sup>14</sup> M. Lima,<sup>20,39</sup> M. A. G. Maia,<sup>20,25</sup> M. March,<sup>24</sup> J. L. Marshall,<sup>40</sup> P. Melchior,<sup>41</sup> F. Menanteau,<sup>21,22</sup> C. J. Miller,<sup>29,30</sup> R. Miquel,<sup>23,42</sup> R. L. C. Ogando,<sup>20,25</sup> A. K. Romer,<sup>43</sup> E. Sanchez,<sup>19</sup> V. Scarpine,<sup>14</sup> M. Schubnell,<sup>30</sup> S. Serrano,<sup>31,32</sup> I. Sevilla-Noarbe,<sup>19</sup> M. Smith,<sup>44</sup> F. Sobreira,<sup>20,45</sup> E. Suchyta,<sup>46</sup> M. E. C. Swanson,<sup>22</sup> G. Tarle,<sup>30</sup> D. Thomas,<sup>15</sup> D. L. Tucker,<sup>14</sup> and Y. Zhang<sup>14</sup> (DES Collaboration)

*Affiliations are listed at the end of the paper*

Accepted 2019 August 1. Received 2019 June 16; in original form 2018 December 13

## ABSTRACT

Weak lensing source galaxy catalogues used in estimating the masses of galaxy clusters can be heavily contaminated by cluster members, prohibiting accurate mass calibration. In this study, we test the performance of an estimator for the extent of cluster member contamination based on decomposing the photometric redshift  $P(z)$  of source galaxies into contaminating and background components. We perform a full scale mock analysis on a simulated sky survey approximately mirroring the observational properties of the Dark Energy Survey Year One observations (DES Y1), and find excellent agreement between the true number profile of contaminating cluster member galaxies in the simulation and the estimated one. We further apply the method to estimate the cluster member contamination for the DES Y1 redMaPPer cluster mass calibration analysis, and compare the results to an alternative approach based on the angular correlation of weak lensing source galaxies. We find indications that the correlation based estimates are biased by the selection of the weak lensing sources in the cluster vicinity, which does not strongly impact the  $P(z)$  decomposition method. Collectively, these benchmarks demonstrate the strength of the  $P(z)$  decomposition method in alleviating membership contamination and enabling highly accurate cluster weak lensing studies without broad exclusion of source galaxies, thereby improving the total constraining power of cluster mass calibration via weak lensing.

**Key words:** gravitational lensing: weak – galaxies: clusters: general – cosmology: observations.

★ E-mail: [t.varga@physik.lmu.de](mailto:t.varga@physik.lmu.de)

## 1 INTRODUCTION

Galaxy clusters trace the highest peaks of the cosmic density field and their abundance and distribution constitutes a powerful cosmological probe (Allen, Evrard & Mantz 2011; Dodelson et al. 2016). This mode of inference poses two major tasks: detecting galaxy clusters from observational data, and defining a mass–observable relation (MOR) to compare the observed cluster abundances with the predicted ones. The efficient pathways to identify galaxy clusters differ between the available wavelengths and targeted redshifts ranges: In optical wavelengths and low-redshifts ( $z < 1$ ) clusters can be detected as an overdensity of quenched, red, early-type galaxies (Koester et al. 2007; Rykoff et al. 2014), giving rise to the cluster mass proxy known as *optical richness* ( $\lambda$ ). In other wavelengths, clusters can be identified through the X-ray emission (Truemper 1993; Mantz et al. 2010) or through the Sunyaev–Zeldovich effect (Sunyaev & Zeldovich 1970, 1972; Bleem et al. 2015) induced by the presence of hot intracluster gas. While these methods are suitable to detect clusters, they do not provide a direct measure of their masses. The MOR must be calibrated using additional information.

The best method for calibrating cluster masses today is via weak gravitational lensing, as it is directly sensitive to the gravitational potential. For this reason weak lensing cluster mass calibration studies (von der Linden et al. 2014a, b; Applegate et al. 2014; Hoekstra et al. 2015; Mantz et al. 2015; Okabe & Smith 2016; Battaglia et al. 2016; Melchior et al. 2017; Simet et al. 2017; Murata et al. 2017; Dietrich et al. 2017; McClintock et al. 2018) have become a necessary component of cluster cosmology analyses. Weak lensing mass estimates carry their own set of uncertainties, both systematic and statistical. It is expected that to fully utilize the statistical power of ongoing sky surveys, the amplitude of the MOR must be calibrated with at most a few percent total uncertainty (Weinberg et al. 2013). With the growing depth, area, and statistical power of various sky surveys the proper characterization of systematic uncertainties is becoming the highest priority. Indeed, current analyses are systematics dominated (e.g. table 6 of McClintock et al. 2018), meaning that to improve on the overall cosmological constraining power we have to improve our understanding of systematic errors in the mass calibration.

One important systematic impacting weak lensing analyses is the contamination of the source galaxy catalogue by galaxies associated with the cluster. This contamination is a result of the uncertainty in photometric redshift estimates, as few-band surveys do not provide enough information to select a pure and close to complete background sample of galaxies. Contaminating galaxies dilute the measurement, requiring one to boost the raw signal to recover the true signal. Hence the effect is traditionally referred to as the *boost factor* (Sheldon et al. 2004; Applegate et al. 2014; Hoekstra et al. 2015; Gruen et al. 2014; Simet et al. 2017; Melchior et al. 2017; Medezinski et al. 2018b; Leauthaud et al. 2017). Conversely, when many-band photometric information is available, the contamination can be strongly reduced, but with increased observational cost (Applegate et al. 2014).

Previous studies made use of multiple approaches in characterizing cluster member contamination: Sheldon et al. (2004) and Simet et al. (2017) estimated the boost factor profiles from the transverse correlation of source galaxies around cluster centres, while Applegate et al. (2014) and Medezinski et al. (2018b, a) utilized the colour information in a ‘colour-cut’ method. Gruen et al. (2014) and Dietrich et al. (2017) estimated the correction factor from decomposing the source population into a cluster and background component. This latter method was expanded by Mel-

chior et al. (2017) who estimated the contamination rate based on a decomposition of the photometric redshift probability distribution function (p.d.f.,  $P(z)$ ) estimates of source galaxies, which was also employed by Chang et al. (2018) and Stern et al. (2018).

In this study, we aim to validate the cluster member contamination estimates obtained through  $P(z)$  decomposition, and provide a detailed description for the case of the Dark Energy Survey Year One observations (DES Y1) cluster weak lensing analysis of McClintock et al. (2018). The structure of this paper is as follows: in Section 2, we outline the framework and formalism of the  $P(z)$  decomposition method, in Section 3, we perform tests on simulated DES-like observations as well as actual DES data, and finally in Section 4, we present the boost factor results used in the DES Y1 redMaPPer weak lensing cluster mass calibration (McClintock et al. 2018).

For the DES Y1 data analysis, we assume a flat  $\Lambda$  cold dark matter ( $\Lambda$ CDM) cosmology with  $\Omega_m = 0.3$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , with distances defined in physical coordinates, rather than comoving. The DES-like mock observations assume a flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.286$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_b = 0.047$ ,  $n_s = 0.96$ , and  $\sigma_8 = 0.82$ .

## 2 $P(Z)$ DECOMPOSITION FORMALISM

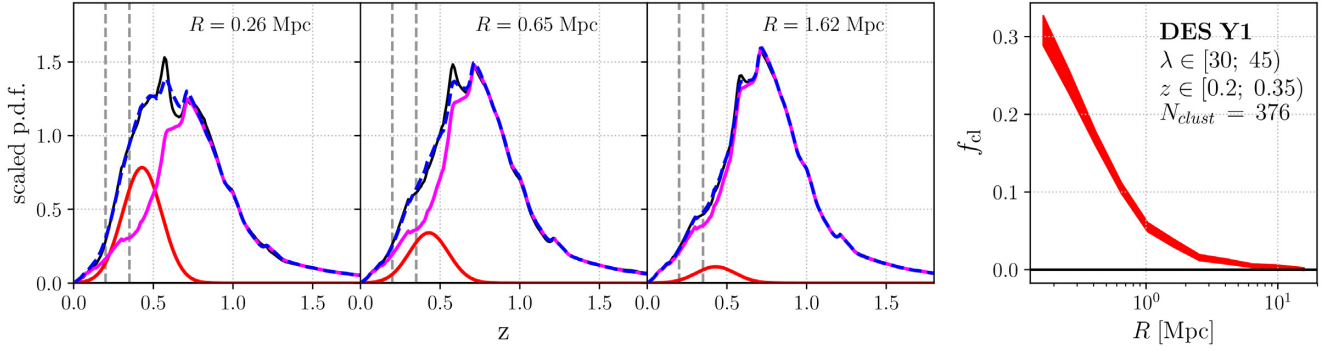
Our aim is to estimate the cluster member contamination affecting weak lensing measurements. With an estimate of the contamination rate that has sufficiently low systematic and statistical uncertainty, we can correct for the bias in the raw weak lensing signal.

The present approach infers the fraction of contaminating cluster member galaxies  $f_{cl}$  from the photometric redshift  $P(z)$  p.d.f. of the appropriately selected and weighted source galaxies. By comparing the  $P(z)$  of sources near clusters with the  $P(z)$  of galaxies in field lines of sight we identify a feature indicative of the presence of cluster galaxies. The schematic of this approach is illustrated on Fig. 1 for a selection of galaxy clusters. The three left-hand panels show the average  $P(z)$  of galaxies at three different transverse separations from the clusters, along with the reference  $P(z)$  of field galaxies. The feature characteristic for the cluster member contamination is shown as a red Gaussian curve. We note that the position of the Gaussian is expected to be at a higher redshift than the redshift range of the targeted clusters, as only those cluster member galaxies contribute to the contamination whose redshift estimates are significantly biased high. The relative strength of this feature at different radii is taken as a tracer of the cluster member contamination rate profile  $f_{cl}(R)$ , shown on the right-hand panel of Fig. 1.

### 2.1 Weak lensing formalism

Weak lensing analyses of galaxy clusters rely on a large sample of background *source* galaxies. The images of these background source galaxies are distorted due to the gravitational potential of the lens (the galaxy cluster), and thus can be used to trace the underlying matter distribution of the lens (Bartelmann & Schneider 2001).

In most scenarios, a catalogue of background source galaxies is constructed from optical imaging data, and thus their exact distances are not known. To remedy this, photometric redshift algorithms are employed to provide an estimate of their redshifts. Such methods involve large uncertainty for individual galaxies, and potential bias for the ensemble, due to the limited information available (Hoyle et al. 2017; Hildebrandt et al. 2017). The uncertainty of photometric redshifts mean that the background source catalogue can only be defined in approximate terms, such that it may also include foreground galaxies, and galaxies which are at the lens redshift. In



**Figure 1.** Left-hand panels:  $P(z)$  decomposition at three different radial ranges for the cluster sample with richness  $\lambda \in [30; 45]$  and redshift  $z \in [0.2; 0.35]$  in DES Y1 data. Black lines: average-weighted  $P(z|R)$  of source galaxies. Red lines:  $P(z)$  of the Gaussian contamination component scaled by the estimated cluster member contamination rate  $f_{cl}$ . Magenta lines: average  $P(z|\text{field})$  scaled by  $1 - f_{cl}$ . Blue dashed: model  $P(z)$  calculated from the sum of the magenta and red lines. The vertical dashed lines indicate the redshift range of galaxy clusters in the cluster selection. Right-hand panel: the cluster member contamination rate  $f_{cl}$  profile calculated from the decomposition presented on the left-hand panels: the red shaded range corresponds to the amplitudes of the Gaussian components at each radial range. The width of the shaded area indicates the  $1\sigma$  uncertainty region.

contrast, the redshifts of galaxy clusters are typically known with very high precision either from spectroscopic follow-up, or from the ensemble photometric redshift estimates of their red cluster member population (Rykoff et al. 2016; McClintock et al. 2018). In this analysis we neglect any uncertainty in the cluster redshift  $z_{lens}$ , and consider that the uncertainty in the redshift of source galaxies  $z_{src}$  is captured in their  $P(z_{src})$  p.d.f.

The photometric redshift  $P(z)$  assigned to a galaxy is determined by its observed properties, most commonly its colours, and also by our prior knowledge about its likely redshifts. To reach good photo- $z$  performance (e.g. low bias), the prior should be strongly dependent on the selection function. Defining this in practice requires a reference sample of galaxies for which the mapping between redshift and observed properties is known, and is representative of the target galaxy selection. Hence when redshift estimates calibrated with one selection are used together with a significantly different selection during the science analysis, they are no longer guaranteed to retain their fiducial performance. (Bonnett et al. 2015; Hoyle et al. 2017).

Background galaxies ( $z_{src} > z_{lens}$ ) at different redshifts (distances) contribute to the lensing signal with different amplitudes. This is characterized by the inverse of the *critical surface density*:

$$\Sigma_{crit}(z_{lens}, z_{src}) = \frac{c^2}{4\pi G} \frac{D_s(z_{src})}{D_l(z_{lens}) D_{ls}(z_{lens}, z_{src})}, \quad (1)$$

where  $D_s$ ,  $D_l$ , and  $D_{ls}$  denote *angular diameter* distances to the source galaxy, the lens, and between the lens and the source, respectively.

In a cluster weak lensing scenario, the quantity of interest is the average tangential component of the reduced gravitational shear  $g_T = \gamma_T / (1 - \kappa)$  where  $\gamma$  is the weak lensing shear and  $\kappa$  is the convergence. This is estimated from the shapes and alignments of the source galaxies through the ellipticity measure  $e$ , where we assume that  $\langle e \rangle \approx \langle g \rangle$ . The shear signal is related to the *excess surface mass density*  $\Delta\Sigma$ , expressible from the physical mass distribution of the lens system via:

$$\gamma_T(R) = \frac{\bar{\Sigma}(<R) - \bar{\Sigma}(R)}{\Sigma_{crit}} = \frac{\Delta\Sigma(R)}{\Sigma_{crit}}, \quad (2)$$

and the convergence is defined as

$$\kappa(R) = \bar{\Sigma}(R) / \Sigma_{crit}, \quad (3)$$

where  $R$  is the projected separation from the lens,  $\bar{\Sigma}(<R)$  and  $\bar{\Sigma}(R)$  refer to the average surface mass density within radius  $R$ , and at radius  $R$ , respectively.

Following Sheldon et al. (2004), we define the maximum-likelihood estimator for the stacked excess surface mass density of multiple clusters:

$$\widetilde{\Delta\Sigma} = \frac{\sum_i \sum_j^{src \ lens} w_{i,j} e_{i,j}^T / \langle \Sigma_{crit}^{-1} \rangle_{i,j}}{\sum_i \sum_j^{src \ lens} w_{i,j}} = \frac{\sum_i \sum_j^{src \ lens} w_{i,j} \Delta\Sigma_{i,j}}{\sum_i \sum_j^{src \ lens} w_{i,j}}, \quad (4)$$

with weights

$$w_{i,j} = \langle \Sigma_{crit}^{-1} \rangle_{i,j}^2 / \sigma_{e,i}^2, \quad (5)$$

Here,  $e_{i,j}^T$  corresponds to the tangential component of the *estimated ellipticity*  $e$  of the  $i$ th source galaxy relative to the  $j$ th lens,  $\sigma_{e,i}^2$  is the variance of the estimated shape for galaxy  $i$ , and

$$\langle \Sigma_{crit}^{-1} \rangle_{i,j} = \int dz_{src} P_i(z_{src}) \Sigma_{crit,i,j}^{-1}(z_{lens}, z_{src}) \quad (6)$$

is defined as the *effective* inverse critical surface density for source–lens pair  $i, j$ .<sup>1</sup> We note that  $\Delta\Sigma$  relates to  $\gamma$ , however the distortion of source galaxies is determined by  $g$ , thus the effect of magnification must be accounted for during the modelling of the measured lensing signal.

The  $\widetilde{\Delta\Sigma}$  estimator defined in equation (4) is unbiased if the redshift p.d.f.  $P(z_{src})$  is estimated correctly. In general, intrinsic bias in the photometric redshift estimates would also bias the weak lensing measurement,<sup>2</sup> while uncertainty alone can be propagated self-consistently (Applegate et al. 2014). For this reason, weak lensing surveys spend great effort calibrating photometric redshifts for their weak lensing source galaxy catalogues (Kelly et al. 2014; Hoyle et al. 2017; Hildebrandt et al. 2017; Tanaka et al. 2018). In case the photometric redshifts are biased, it requires additional work and loss of constraining power to ensure that the photo- $z$  bias

<sup>1</sup>We estimate  $\langle \Sigma_{crit}^{-1} \rangle_{i,j}$ , as  $\langle \Sigma_{crit} \rangle_{i,j}$  is less numerically stable.

<sup>2</sup>Depending on the form of the lensing estimator, it is possible for the redshift estimate to be biased in a way such that the lensing estimator is still unbiased. This however does not hold for arbitrary photo- $z$  bias.

is appropriately propagated into the systematic error budget of the final scientific results (DES Collaboration 2017).

## 2.2 Boost factor formalism

Galaxy clusters represent a large overdensity of physically associated galaxies, consisting both of actual cluster member galaxies, and also the galaxies inhabiting correlated structures. Thus, cluster lines of sight are different from the average line of sight: the galaxy overdensity is concentrated at a tight peak in redshift, much narrower than the resolution of photometric redshift estimates of individual galaxies. In the transverse direction, the number density of the cluster-related galaxy population decreases as one moves away from the cluster centre, where different galaxy populations (e.g. red and blue cluster galaxies) follow different radial profiles (Navarro, Frenk & White 1996; Rykoff et al. 2014).

Observationally, the weak lensing source galaxy catalogue is built from galaxies selected according to morphological, photometric and spatial selection criteria (Zuntz et al. 2017; Hildebrandt et al. 2017; Mandelbaum et al. 2018), and galaxies associated with the targeted clusters also enter the catalogue if they satisfy those criteria. Ideally cluster galaxies would be excluded (e.g. as in Schrabback et al. 2018) since they are at the lens redshift and carry no lensing signal. However in a wide field survey, priors used to estimate photometric redshifts do not account for the presence of the targeted clusters. Consequently, the redshift estimates can be greatly biased, and allow for cluster galaxies leaking into the source catalogue with non-zero weights. Hence, in the above picture, the cluster member contamination can be viewed as a form of selection or representativeness bias impacting the redshift estimates. The redshift bias may further depend on galaxy type, resulting in different rates of contamination by different populations of cluster galaxies. Thus, defining a high purity background sample of sufficient volume may not be possible. We note that depending on the radial separations, the contamination can originate from both the targeted galaxy clusters and also from galaxies in the correlated matter structures. For reasons of brevity, we refer to both of these sources as cluster member contamination, as they can not be disentangled based purely on available redshift information.

To quantify the required boost factor correction we need to consider the impact of contamination on the  $\Delta\Sigma$  estimator defined in equation (4). Following the method developed in Gruen et al. (2014) and extended in Melchior et al. (2017), we assume a model for the *true* line-of-sight distribution of source galaxies selected during the measurement as a combination of two terms: a cluster galaxy component which is effectively a Dirac- $\delta$  function located at  $z_{\text{clust}}$ , and a non-cluster or background component taken to be the lensing weighted distribution of source galaxies in field lines of sight.

Via the above line-of-sight model, we can expand equation (4) into the sum of contributions from the cluster (cl) and background (bg) terms:

$$\begin{aligned} \widetilde{\Delta\Sigma}_{\text{est}} &= \frac{\sum_{j,i \in \text{cl}} w_{i,j} \Delta\Sigma_{i,j} + \sum_{j,i \in \text{bg}} w_{i,j} \Delta\Sigma_{i,j}}{\sum w_{i,j}} \\ &= \left( \frac{\sum_{j,i \in \text{cl}} w_{i,j}}{\sum w_{i,j}} \right) \underbrace{\langle \Delta\Sigma_{i,j} \rangle_{\text{cl}}}_0 + \left( \frac{\sum_{j,i \in \text{bg}} w_{i,j}}{\sum w_{i,j}} \right) \underbrace{\langle \Delta\Sigma_{i,j} \rangle_{\text{bg}}}_{\langle \Delta\Sigma_{i,j} \rangle_{\text{true}}} \end{aligned} \quad (7)$$

of which  $\langle \Delta\Sigma_{i,j} \rangle_{\text{cl}}$  carries no signal, while  $\langle \Delta\Sigma_{i,j} \rangle_{\text{bg}}$  is the ‘true’ signal we would estimate if there was no contamination.  $\sum_{j,i \in \text{cl}}$

and  $\sum_{j,i \in \text{bg}}$  denotes a sum over source–lens pairs with cluster members and background galaxies respectively. We furthermore define the effective contamination rate of cluster galaxies  $f_{\text{cl}}$  via

$$f_{\text{cl}} = \frac{\sum_{j,i \in \text{cl}} w_{i,j}}{\sum_{j,i} w_{i,j}}, \quad (8)$$

which we can use to express the boost correction needed to recover the true signal

$$\widetilde{\Delta\Sigma}_{\text{corr}}(R) = \frac{\widetilde{\Delta\Sigma}(R)}{1 - f_{\text{cl}}(R)}. \quad (9)$$

Here,  $\widetilde{\Delta\Sigma}$  denotes the raw measured lensing signal obtained from equation (4),  $\Delta\Sigma_{\text{corr}}(R)$  denotes the lensing signal corrected for contamination, and  $\mathcal{B} \equiv (1 - f_{\text{cl}})^{-1}$  is referred to as the *boost factor*. Hence, in the above framework, the cluster member contamination correction for a given measurement scenario is completely characterized by the  $f_{\text{cl}}(R)$  profile.

## 2.3 Estimating the contamination using $P(z)$ decomposition

We estimate the contamination rate from the available colour–magnitude information of source galaxies, where, due to the overdensity of the cluster we expect that the contaminating cluster galaxies will appear as a sub-population. We follow Melchior et al. (2017), and make use of the *observed* lensing-weighted average redshift probability distribution of the sources

$$P(z | R) = \frac{\sum_{i,j} w_{i,j} P_i(z, R)}{\sum_{i,j} w_{i,j}}, \quad (10)$$

which contain information compressed from colour–magnitude space into a probability distribution function. We measure this at different projected radii  $R$  around the cluster. The weights  $w_{i,j}$  for each source are identical to the ones introduced in equation (4). In this framework, the estimated redshifts represent information compression from the colour–magnitude space into a single  $P(z)$  estimate per sample. Contaminating cluster members contribute to the average photometric redshift  $P(z)$ s differently in different radial ranges. Thus by tracking the changes in the  $P(z)$  as a function of radius, we can recover an estimate of the underlying cluster member contamination.

We model the observed redshift distribution  $P(z | R)$  as a combination of two terms, reflecting the cluster and background populations defined in Section 2.2:

$$P(z | R) = f_{\text{cl}}(R) \cdot P_{\text{memb}}(z) + (1 - f_{\text{cl}}(R)) \cdot P_{\text{bg}}(z), \quad (11)$$

where  $P_{\text{memb}}(z)$  is the redshift distribution of *contaminating* cluster member galaxies and  $P_{\text{bg}}(z)$  is the distribution of background galaxies (Gruen et al. 2014; Melchior et al. 2017). We approximate the second term by the appropriately weighted redshift distribution of the average survey *field* line of sight:  $P_{\text{bg}}(z) \approx P_{\text{field}}(z)$ . As an ansatz, we consider  $P_{\text{memb}}(z)$  to be a Gaussian distribution. The validity of this assumption is tested in Section 3.2.5. The free parameters of the decomposition are the mean and width of the Gaussian  $P_{\text{memb}}(z)$ , and the contamination rate  $f_{\text{cl}}(R)$ .

An example for this  $P(z)$  decomposition method is shown on Fig. 1 for the case of DES Y1 data. There, a qualitatively similar behaviour is visible for the different radial bins, and the contamination increases with decreasing radius.



### 3 METHOD VALIDATION

We perform a validation benchmark to test the robustness and performance of the  $P(z)$  decomposition boost estimates. First, we outline the primary assumptions of the decomposition method in Section 3.1, validate the method in a mock analysis scenario in Section 3.2, and perform consistency tests on DES Y1 data in Section 3.3.

#### 3.1 Model assumptions

The  $P(z)$  decomposition method relies on several assumptions about both the contaminating and the background galaxies which impact the efficacy of the method. Some of these we explore below, i.e. the potential intrinsic alignment of cluster galaxies (Section 3.1.1), the impact of weak lensing magnification (Section 3.1.2), and the influence of blinding and intracluster light on photometry used in the decomposition estimates (Section 3.1.3). Other assumptions are tested in later sections, i.e. the Gaussian ansatz for  $P_{\text{memb}}(z)$  (Section 3.2.5), and the influence of the chosen background model (Section 3.3.3).

##### 3.1.1 Intrinsic alignments

Contaminating galaxies physically connected to the lens system possess an *intrinsic alignment* due to the tidal forces acting between them. When cluster members are included in a lensing measurement, intrinsic alignments could appear as negative tangential shear around clusters due to the preferential radial orientation of galaxies. This effect is difficult to decouple from the physical lensing signal. Recent spectroscopic follow-up studies of Hao et al. (2011) and Sifón et al. (2015) found no significant signal for preferential alignments of cluster member galaxies with respect to the cluster centres. Huang et al. (2018) found significant detection only when considering a high luminosity subset of galaxies, but no detection when considering their complete galaxy sample. For this reason we assume that the dominant effect of contaminating cluster members is the dilution of the lensing signal.

##### 3.1.2 Magnification

Weak lensing magnification caused by the potential of the cluster changes the observed number density and luminosity function of background galaxies. This translates into biased photometric redshift estimates for background galaxies, which would result in the  $P(z)$  of background galaxies near clusters to be different from the  $P(z)$  of the similarly weighted galaxies in field lines of sight. Gruen & Brimiouille (2017, their appendix C), however, finds, that under realistic survey assumptions biases in photometric redshift estimates due to the increased flux of magnified sources, and due to the different surface density of magnified galaxies are sub-per-cent effects that partially cancel one another.

##### 3.1.3 Impact of blinding and intracluster light

The potential bias due to blinding and source obscuration in the estimated  $P(z)$ s is difficult to estimate, as it would require detailed understanding of detection and shape measurement selection probabilities, as well as the photometric transfer function of representative source galaxy samples in cluster fields (see e.g. Chang et al. 2015; Suchyta et al. 2016). To a first approximation, we expect blinding and source obscuration to uniformly impact

all source galaxies, leading to an amplitude shift in the  $P(z)$  of the selected source galaxies. However given the excellent match of the cluster background population  $P(z)$  and the field background population  $P(z)$  at large  $z \gg z_l$  (visible in Fig. 1), we assume that the impact of these effects is strongly subdominant. In Section 3.3.3, we nevertheless perform a simple consistency test for differences in the background  $P(z)$ .

The presence of *intracluster light* biases the photometric redshift estimates, influencing the recovered  $P(z)$ s in a manner similar to blinding. However Gruen et al. (2019, their appendix A) estimated the impact of this effect to be negligible for the radial scales considered in this study.

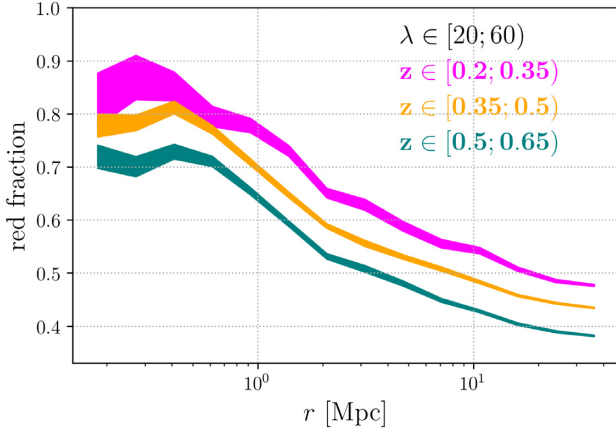
#### 3.2 Benchmark on the Buzzard mock observations

We test the  $P(z)$  decomposition method against the true cluster member contamination in a simulated environment, mirroring the measurement setup of McClintock et al. (2018). In Sections 3.2.1 and 3.2.2, we introduce the simulated observations and the mock galaxy clusters. In Section 3.2.3, we perform the  $P(z)$ -decomposition on the simulated *Buzzard* data. In Section 3.2.4, we determine the true contamination of our photo- $z$  selected source sample. In Section 3.2.5, we test the validity of the Gaussian ansatz for  $P_{\text{memb}}(z)$ . Finally, in Section 3.2.6, we discuss the agreement between the true and estimated contamination rates.

##### 3.2.1 Buzzard simulated light-cones

The *Buzzard*-suite of cosmological simulations (DeRose et al. 2019) consists of mock DES Y1 catalogues generated by combining three  $N$ -body light-cones created using L-GADGET2, a version of GADGET2 (Springel 2005) optimized for memory efficiency. The initial conditions were set-up via second-order Lagrangian perturbation theory using 2LPTIC (Crocce, Pueblas & Scoccimarro 2006). The light-cones were produced on the fly using simulation boxes with volumes  $1050^3$ ,  $2600^3$ , and  $4000^3 (h^{-1}\text{Mpc})^3$ ; the corresponding particle masses are  $3.3 \times 10^{10}$ ,  $1.6 \times 10^{11}$ , and  $5.9 \times 10^{11} h^{-1}\text{M}_\odot$ . The resulting light-cones were joined at redshifts 0.34 and 0.9, arranged such that the highest resolution simulations are used at lower redshifts. These simulation boxes assume a flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.286$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_b = 0.047$ ,  $n_s = 0.96$ , and  $\sigma_8 = 0.82$ . The galaxy catalogues were created by assigning galaxies to dark matter particles via the ADDGALS algorithm (Wechsler, DeRose & Busha, in preparation). ADDGALS calibrates the relation between the large-scale density and the  $r$ -band absolute magnitudes of galaxies as measured using subhalo abundance matching (Conroy, Wechsler & Kravtsov 2006; Reddick et al. 2013; Lehmann et al. 2017) in a high-resolution  $N$ -body simulation. For each simulated galaxy, spectral energy distributions (SEDs) are assigned from the SDSS DR7 VAGC (Cooper et al. 2011) by finding the galaxy in the data with the closest match in  $M_r - \Sigma_5 - z$  space, where  $M_r$  is the galaxy's rest frame  $r$ -band absolute magnitude and  $\Sigma_5$  is the distance to the fifth-nearest galaxy in projection (see appendix E2 of DeRose et al. 2019). Photometric noise is added in accordance with the DES Y1 depth map of Drlica-Wagner et al. (2017), and  $g$ ,  $r$ ,  $i$ ,  $z$  fluxes in the DES filters are generated from the previously assigned SEDs.

In this study, we use version 1.3 of the *Buzzard* mock catalogues. Only the main 'SPT' area of DES Y1 is simulated, and the footprint is restricted to  $\text{RA} < 0$  to exclude areas where the DES coverage is more inhomogeneous. This yields a simulated sky survey with a



**Figure 2.** Fraction of red galaxies as a function of radius around redMaPPer clusters within the *Buzzard* mock observation across different redshift ranges but fixed richness. Shaded areas indicate  $1\sigma$  statistical uncertainties estimated from JK resampling.

total area of  $1120 \text{ deg}^2$ . The resulting galaxy fluxes include the effect of weak lensing magnification based on ray tracing along their lines of sight, with the highest redshift galaxy being located at  $z = 2.35$  (DeRose et al. 2019). For the purposes of the current measurement, we selected sources in a way that is meant to approximate the source selection in the DES Y1 analysis (Zuntz et al. 2017) by applying signal-to-noise cuts following MacCrann et al. (2018). This sample is defined purely to mirror the properties of actual DES source galaxies, and does not contain shear or photometry systematics. We then run the BPZ template based photometric redshift algorithm (Benítez 2000; Coe et al. 2006) on this mock catalogue to obtain a  $P(z)$  estimate for each source galaxy with equivalent settings as used by Hoyle et al. (2017) for the DES Y1 data. Given plausible galaxy colours and identical measurement setup, we expect BPZ to possess similar performance in Buzzard as in the DES Y1 data.

### 3.2.2 Simulated galaxy clusters

In order to obtain a simulated cluster sample similar to the one presented by McClintock et al. (2018), we run the redMaPPer algorithm (Rykoff et al. 2014) on the mock galaxy catalogues with the same configuration as the real DES Y1 data.

RedMaPPer is a red-sequence-based optical matched filter cluster finder which produces an estimate on the position, the optical richness  $\lambda$ , and redshift of the detected clusters. This yields a cluster catalogue with comparable distribution in angular position and redshift to the catalogue in the DES Y1 data set. A catalogue of reference random points is also generated, which are defined as positions and redshifts where the survey conditions (e.g. mask and observational depth) allow clusters of given richness to be detected.

While the redMaPPer algorithm is sensitive only to the over-density of red-sequence galaxies, we also test the blue galaxy content of clusters as they are expected to significantly contribute to the contamination. We calculate the fraction of red galaxies as a function of radii using a rest-frame magnitude limit of  $m_r > -19$ . For this, we take a galaxy as red if it belongs to the red sequence defined in the rest-frame colour–magnitude space, which in practice corresponds to a cut of  $(g - r) > 0.2 \cdot r - 0.028$ . Fig. 2 shows this red fraction across different redshift ranges, where we

find good qualitative agreement with previous observational studies (e.g. Butcher & Oemler 1978; Hansen et al. 2009, their fig. 12). The red fractions are different across different redshift bins with a larger blue cluster member population at higher redshifts, which is expected from the time dependence of the galaxy quenching process.

A difference between the real and mock cluster catalogues is that clusters in the simulation appear to have a stronger redshift evolution in richness at a given halo mass relative to expectations from existing scaling relations. This fact along with the reduced simulated footprint results in a lower number of clusters in richness bins at low redshift compared to McClintock et al. (2018). In addition, the DES Y1 data are deeper than the reference data set used by the ADDGALS algorithm to populate SEDs, and for this reason, the mock galaxy populations and their relative abundances at faint magnitudes or high redshifts might differ from reality. Because of this, as well as because of differences in source galaxy selection and between our real and synthetic datasets, we do not expect the cluster member contamination rates in the mock observations to be equal to our DES Y1 measurements. Nevertheless, the mocks include many qualitative aspects of the real observations, and for this reason we make use of them as a controlled environment to benchmark and validate the performance of the  $P(z)$  decomposition under somewhat simplified circumstances.

### 3.2.3 Decomposition results in simulated catalogues

We estimate boost factors for the redMaPPer clusters in the mocks using an identical measurement setup as McClintock et al. (2018). Hence, the  $\Delta\Sigma$  estimator takes the form of:

$$\widehat{\Delta\Sigma} = \frac{\sum \omega_{i,j} e_{T,i}}{\sum \omega_{i,j} \Sigma_{\text{crit},i}^{-1}} \quad (12)$$

with

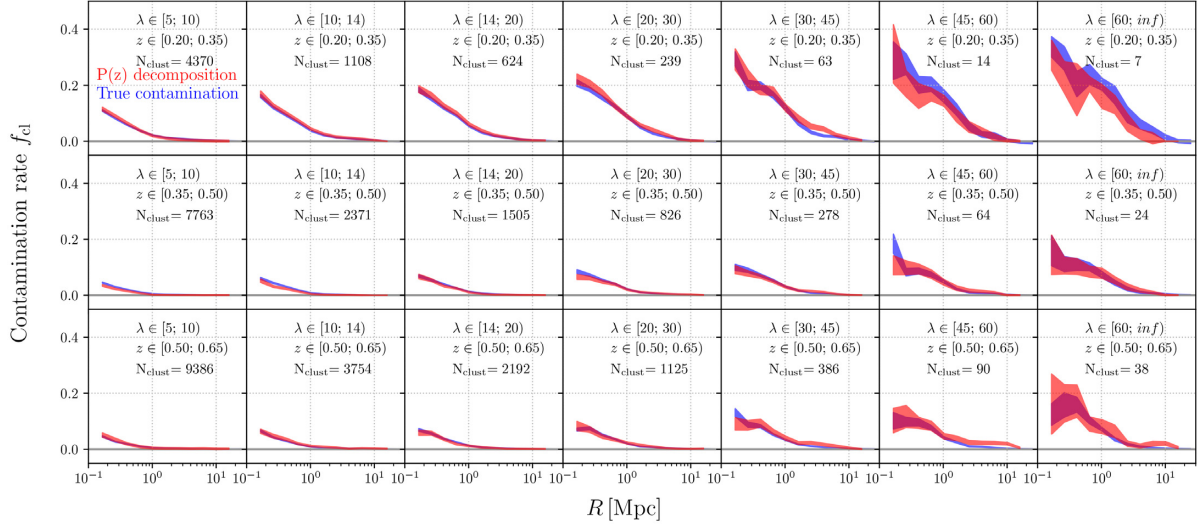
$$\omega_{i,j} \equiv \Sigma_{\text{crit},i}^{-1}(z_{l_j}, \langle z_{s_i} \rangle) \text{ if } \langle z_{s_i} \rangle > z_{l_j} + 0.1. \quad (13)$$

Where  $\Sigma_{\text{crit},i}^{-1}$  is calculated at a source redshift randomly drawn from the corresponding  $P(z)$ , while  $\Sigma_{\text{crit}}^{-1}(z_{l_j}, \langle z_{s_i} \rangle)$  represents the value at the mean redshift of the source. The mock galaxy catalogue does not include shear biases thus we set the shear and selection responses to unity (see equation 12 of McClintock et al. 2018). Following equations (7) and (8), the contamination fraction is given by:

$$f_{\text{cl}} = \frac{\sum_{\text{cl}} \omega_{i,j} \Sigma_{\text{crit},i,j}^{-1}}{\sum \omega_{i,j} \Sigma_{\text{crit},i,j}^{-1}}. \quad (14)$$

We divide the clusters into bins of redshift  $z \in [0.2; 0.4]$ ,  $[0.4; 0.5]$ , and  $[0.5; 0.65]$ , and richness:  $\lambda \in [5; 10]$ ,  $[10; 14]$ ,  $[14; 20]$ ,  $[20; 30]$ ,  $[30; 45]$ ,  $[45; 60]$ , and  $[60; \infty)$ .  $\widehat{\Delta\Sigma}$  is calculated in 11 logarithmically spaced radial bins ranging from 0.2 to 30 Mpc. For each cluster sample and radial range we save  $\min(N_{\text{pairs}}; 2 \times 10^4)$  representative source–lens pairs selected in a uniform random way, and calculate the mean  $P(z)$  of that source population weighted by  $\omega_{\Sigma_{\text{crit}}^{-1}}$ . The estimate on  $f_{\text{cl}}(R)$  is then found by the  $P(z)$  decomposition method outlined in Section 2.3. For the field component, we take the  $P(z)$  of sources in the outermost radial bin, which we find to be identical to the weighted  $P(z)$  of sources selected around random points in a series of Kolmogorov–Smirnov tests.

The decomposition is calculated by considering all radial scales simultaneously where we require the redshift positions and widths of the cluster components to be identical at different radial ranges. The mixing amplitudes  $f_{\text{cl}}(R)$  between the cluster and reference



**Figure 3.** Cluster member contamination measured in the *Buzzard* mock observations for the various bins of clusters in richness  $\lambda$  and redshift  $z$  defined in Section 3.2.3. Red: contamination profiles calculated via  $P(z)$  decomposition using BPZ photometric redshift estimates with realistic photometric noise added to the mock galaxies. Blue: true contamination profiles calculated from the truth catalogues of the simulations via equation (16).

$P(z)$  are left free across radial bins. Hence, the inner radial scales where the contamination is stronger provide constraints about the cluster component for the outer radial ranges. The  $f_{cl}$  profile model for a cluster sample has  $N_{rbin} + 2$  free parameters, and the decomposition is performed via a least-squares Levenberg–Marquardt algorithm, where the optimized quantity is the mean-squared deviation between the measured  $P(z | R)$  and the model prediction defined in equation (11). This boost factor calculation algorithm is implemented in the XPIPE PYTHON package,<sup>3</sup> which was also used by Chang et al. (2018) and McClintock et al. (2018), and contains an identical setup to Melchior et al. (2017).

To estimate the uncertainty on the recovered  $f_{cl}(R)$ , we use jackknife (JK) resampling following Efron (1982):

$$\mathbf{C}_{f_{cl}}^{JK} = \frac{K-1}{K} \sum_k \left( \widetilde{f}_{cl(k)} - \widetilde{f}_{cl(\cdot)} \right)^T \cdot \left( \widetilde{f}_{cl(k)} - \widetilde{f}_{cl(\cdot)} \right), \quad (15)$$

where  $\widetilde{f}_{cl(\cdot)} = \frac{1}{K} \sum_k \widetilde{f}_{cl(k)}$  and  $\widetilde{f}_{cl(k)}$  denotes the contamination rate estimated via equation (14). We make use of  $K = \min\{100; N_{clust}\}$  simply connected spatial regions  $\mathcal{R}_k$  for each cluster sample, defined via a spherical  $k$ -means algorithm,<sup>4</sup> and  $\widetilde{f}_{cl(k)}$  is calculated from all clusters except those in region  $\mathcal{R}_k$ . With this method, we estimate the covariance between all radial ranges within each richness and redshift bin.

The recovered contamination profiles are shown in Fig. 3, and are qualitatively similar to those observed in the real data. The overall behaviour is consistent with theoretical expectations: For all cluster bins, the contamination rate decreases with increasing redshift, and a clear trend is apparent where richer clusters at a given redshift range produce greater contamination rates.

### 3.2.4 True contamination in simulated catalogues

We calculate the true contamination as the *excess* rate for galaxies to be located within the immediate  $\mathcal{S}_j \equiv [z_j - \Delta z; z_j + \Delta z]$  vicinity

of the clusters, defined via

$$f_{cl}^{\text{true}} \equiv \frac{\sum_j \sum_{z_{s,i} \in \mathcal{S}_j} \omega_{i,j} \Sigma_{crit,i,j}^{-1}}{\sum_j \sum_i \omega_{i,j} \Sigma_{crit,i,j}^{-1}} - \frac{\sum_l \sum_{z_{s,i} \in \mathcal{S}_l} \omega_{i,j} \Sigma_{crit,i,l}^{-1}}{\sum_l \sum_i \omega_{i,j} \Sigma_{crit,i,l}^{-1}}, \quad (16)$$

where  $N_c$  refers to the number of clusters,  $N_r$  to the number of random points, and  $N_s$  to the number of source galaxies, while  $\omega_{i,j}$  is the lensing weight associated with the source–lens pairs defined in equation (13).  $\mathcal{S}_j$  and  $\mathcal{S}_l$  refer to the immediate true redshift vicinities of clusters and random points, respectively. That is,  $f_{cl}^{\text{true}}$  is the probability of finding a galaxy within the redshift range  $\mathcal{S}$  around the clusters, minus the same probability for random lines of sight, where the second term we obtain by saving source–lens pairs around redMaPPer random points. This is equivalent to a cylindrical selection of contaminating galaxies, which is motivated by the fact that the contamination originates not only from physically bound galaxies, but also from galaxies in the extended correlated structures.

Fig. 4 illustrates the above approach. It is clear that a large fraction of source–lens pairs near cluster centres (blue) actually lie at the cluster redshift. Comparing this with the distribution of galaxies around random points (black), the contamination rate is taken as the excess area under the curve within the  $\pm \Delta z$  (dashed) region. Based on Fig. 4, we adopt  $\Delta z = 0.05$  as our fiducial redshift width for the purposes of computing the true contamination rate. The resulting  $f_{cl}^{\text{true}}$  profiles are shown on Fig. 3 as the blue shaded regions, where the  $1\sigma$  uncertainties are estimated from JK resampling using the same approach as in Section 3.2.3.

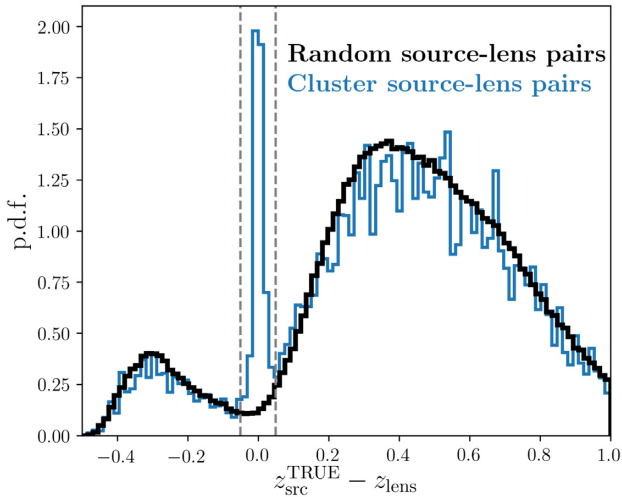
### 3.2.5 Validity of Gaussian cluster model

With the formalism introduced in Section 3.2.4, we can compute the average, weighted, photometric redshift  $P(z)$  for likely cluster galaxies which are located within the  $\pm \Delta z^{\text{TRUE}}$  vicinities of clusters. As shown on Fig. 5, these  $P(z)$ s have a prominent peak located slightly above the cluster redshift range. Cluster galaxy  $P(z)$ s also possess a long tail extending up to high redshifts. This is an intrinsic

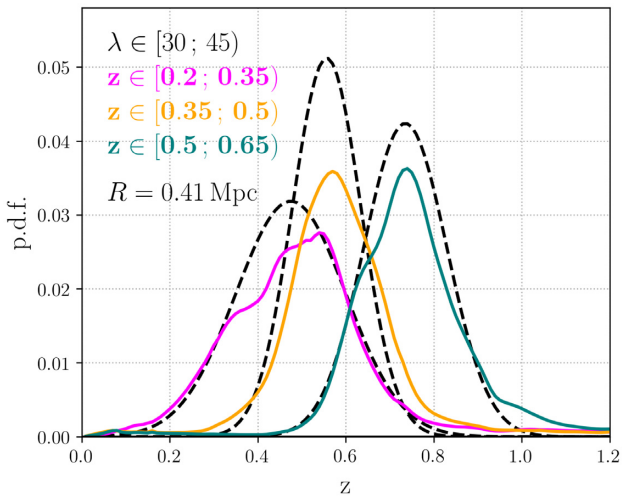
<sup>3</sup><https://github.com/vargatn/xpipe>

<sup>4</sup>[https://github.com/esheldon/kmeans\\_radec](https://github.com/esheldon/kmeans_radec)





**Figure 4.** Schematic for estimating the *true* cluster member contamination fraction in the *Buzzard* mock observations. The figure shows histograms of the weighted, true redshift separation of source–lens pairs at different radial distances from galaxy clusters with  $z \in [0.35; 0.5]$  and  $\lambda \in [30; 45]$ . Blue: source–lens pairs at low radial scales around clusters ( $R < 0.78$  Mpc). Black: source–lens pairs around redMaPPer selected random points in the same radial range. Grey dashed:  $\Delta z = \pm 0.05$  vicinity of the cluster redshift.



**Figure 5.** Comparison of the Gaussian  $P(z)$  component model with the actual  $P(z)$ s of likely cluster galaxies in the *Buzzard* mock observations. Coloured curves:  $P(z)$  of source galaxies within  $\pm \Delta z^{\text{TRUE}}$  of the clusters, in different cluster redshift bins, but at the same richness and radial bin. Dashed curves: best-fitting curves of the Gaussian cluster component model  $P_{\text{memb}}(z)$ . Note that the shown  $P(z)$ s are normalized for  $z \in [0; 3.5]$ .

feature of photometric redshift estimation, as some cluster galaxies have spectral types that exhibit these types of degeneracies.

Fig. 5 also shows the best-fitting Gaussian cluster component models  $P_{\text{memb}}(z)$ . These are obtained from the decomposition method in Section 3.2.3 and are not informed of the true cluster member  $P(z)$ s. Due to the chosen analytic form of  $P_{\text{memb}}(z)$ , the long high-redshift tail of the actual cluster galaxy  $P(z)$ s can not be recovered, which results in the apparent offset of the (normalized) p.d.f.s on Fig. 5. For the shown samples, the best-fitting Gaussian contains within  $1\sigma$  57 per cent, 47 per cent and 56 per cent of

probability of the actual  $P(z)$  of cluster members. Nevertheless, the Gaussians recover the approximate position and width of the peaks, and possesses fewer degrees of freedom than alternative high-skewness models.

### 3.2.6 Discussion of simulation benchmarks

Fig. 3 compares the boost factor profiles estimated from  $P(z)$  decomposition from simulations as described in Section 3.2.3, with the actual cluster member contamination rate calculated in Section 3.2.4. We find that our estimated boost factors are in excellent agreement with the true member contamination rates in the *Buzzard* mock simulations. The uncertainties shown in Fig. 3 are estimated via JK resampling, and do not incorporate *systematic* uncertainties. Hence, we estimate this systematic uncertainty by requiring consistency between the true and estimated  $f_{\text{cl}}$  profiles across all parameter bins with  $N_{\text{clust}} > 50$ . Via this approach we find a global relative systematic uncertainty of  $\delta_{\text{sys}} < 1$  per cent across different richness – redshift selections, where the total covariance is given by  $C_{i,j} = C_{i,j}^{\text{JK}} + \delta_{i,j} \cdot f_{\text{cl},i}^2 \cdot \delta_{\text{sys}}^2$ .

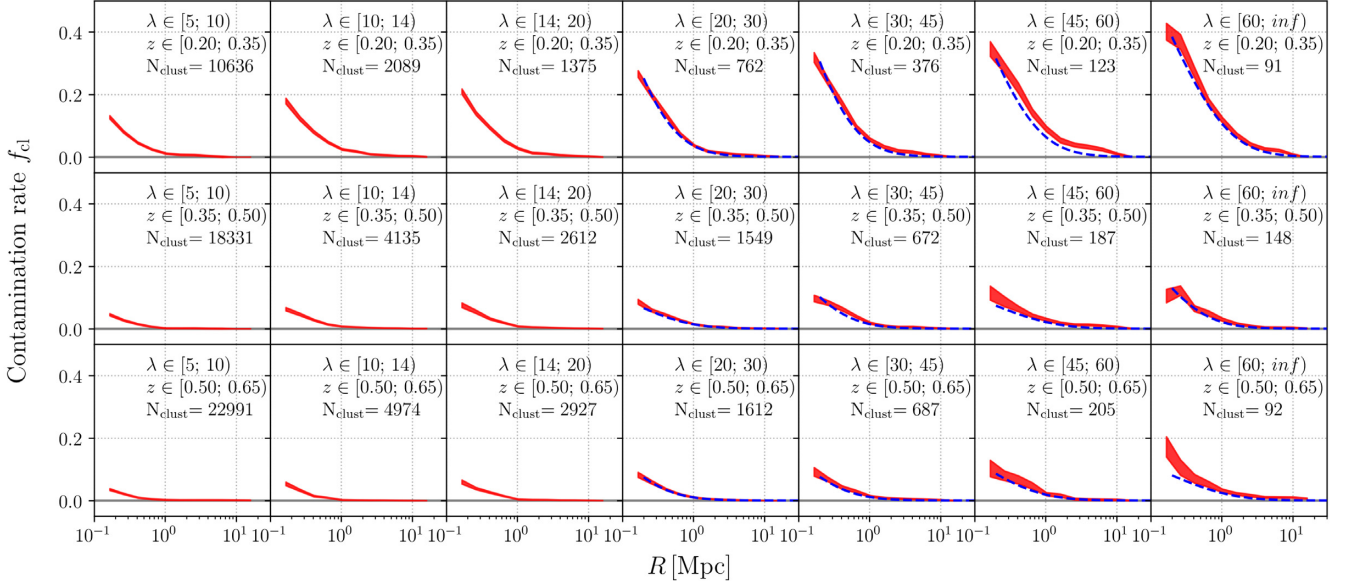
We note that the simulated galaxy catalogues include the effects of magnification with the typical angular resolution of 0.6 arcmin, corresponding to approximately 0.15, 0.2 and 0.24 Mpc in the different redshift bins. While this low resolution allows for only weak constraints, the good agreement between the estimated and true  $f_{\text{cl}}$  profiles indicate that magnification does not play a significant role in the resolved radial ranges.

The purpose of this simulation benchmark is to test how well the  $P(z)$  decomposition predictions match the contamination within the simulation, not to extrapolate for the real DES data. Thus, we do not require full realism from the simulated environment. Nevertheless, the simulated colour distribution of galaxies has been studied by DeRose et al. (2019) in a setup nearly identical to version 1.3 of *Buzzard* used in the present study. They found the simulated galaxy properties to be broadly consistent with reality except for a slight systematic shift on the colour of the blue cloud. We do not anticipate that this manifests in a qualitative difference on the performance of the boost factor estimator compared to real data. We however note that the presented mock scenario is constructed to resemble the DES analysis setup, including the choice of photometric redshift estimation algorithm. For this reason, the observed performance is not necessarily indicative for a significantly different survey and analysis scenario. While the abundances, radial profiles, and colour properties of cluster galaxies in the simulation may be slightly different from reality, we expect *Buzzard* to be qualitatively similar to the real DES Y1 data. Hence we take the excellent performance of the  $P(z)$  decomposition in this setting as a strong motivation for its applicability for real observations.

### 3.3 Analysis on DES Y1 data

In this section, we apply the  $P(z)$  decomposition method to DES Y1 data, following the exact measurement setup presented in McClintock et al. (2018). The structure of this section is the following: in Section 3.3.1, we present the relevant parts of the DES Y1 data set relating to the galaxy cluster catalogue and the weak lensing source galaxies, in Section 3.3.2, we derive the form of the necessary boost factor correction, while in Section 3.3.3 present a simple test on the robustness of our contamination model, and finally in Section 3.3.4, we compare with the alternative method of correlation-based boost factor estimate.





**Figure 6.** Cluster member contamination measured in the DES Y1 data for the various bins of clusters in richness  $\lambda$  and redshift  $z$  defined in Section 3.3.2. Red: contamination profiles calculated via  $P(z)$  decomposition using BPZ redshift estimates. Blue: boost factor model from Section 4.2. The curves correspond to the best-fitting parameters found by McClintock et al. (2018) from a likelihood optimization performed jointly with the cluster mass reconstruction.

### 3.3.1 The DES Y1 data set

The DES Y1 observations cover approximately 1800 deg<sup>2</sup> of the southern sky in  $g,r,i,z$  bands. These observations are processed via a variety of photometric data reduction steps into the Y1 GOLD catalogue (Drlica-Wagner et al. 2017) which is the main science quality catalogue of DES. Using the fiducial multi-epoch, multi-object fitting algorithm (MOF) DES finds the 10 $\sigma$  limiting magnitudes of this data set for 2 arcsec apertures to be  $g \approx 23.7$ ,  $r \approx 23.5$ ,  $i \approx 22.9$ , and  $z \approx 22.2$ . Based on these observations McClintock et al. (2018) defined a locally volume-limited catalogue of galaxy clusters identified via the redMaPPer algorithm. In the Y1 footprint the average MOF limiting magnitude is deep enough to detect a  $0.2 L^*$  galaxy up to  $z \approx 0.7$ , thereby setting the maximum depth of the volume-limited cluster sample.

Approximately 1500 deg<sup>2</sup> of this catalogue is further processed by the METACALIBRATION algorithm (Huff & Mandelbaum 2017; Sheldon & Huff 2017) to define a source galaxy sample (Zuntz et al. 2017). This source galaxy catalogue consists of an ellipticity estimate  $e_i$  for each galaxy, along with ancillary quantities used to perform the bias calibration via the *response*  $R = R_\gamma + R_{\text{sel}}$  of the ellipticity estimates to shear and the source galaxy selection function respectively.

Photometric redshift  $P(z)$ s are calculated via the BPZ template-based algorithm (Hoyle et al. 2017). Two separate redshift estimates are derived: one based on the MOF-based galaxy colours listed in the GOLD catalogue, and a second based on the photometric model obtained from METACALIBRATION. This second METACALIBRATION based redshift estimate is required to properly account for the selection response correction, however is found to have greater scatter compared to the MOF-based redshift estimates. For this reason, McClintock et al. (2018) opted to use the METACALIBRATION estimates only in selecting and weighting source–lens pairs. Hoyle et al. (2017) found these redshift estimates to be mildly biased in the mean redshift. Since the  $P(z)$  decomposition method is only sensitive to the relative shape of the  $P(z)$ s, we do not expect the impact of this bias to be significant.

### 3.3.2 Contamination estimator for DES Y1-like data

In this section, we derive the required boost factor correction for the  $\Delta\Sigma$  estimator employed for the DES Y1 analysis. For this, equations (12) and (13) are replaced with:

$$\widetilde{\Delta\Sigma} = \frac{\sum \hat{\omega}_{i,j} e_{T,i}}{\sum \hat{\omega}_{i,j} R_{\gamma,i,j}^T \Sigma_{\text{crit},i,j}^{-1}}. \quad (17)$$

where

$$\hat{\omega}_{i,j} \equiv \Sigma_{\text{crit}}^{-1}(z_{l_j}, \langle z_{s_i}^{\text{MCAL}} \rangle) \text{ if } \langle z_{s_i}^{\text{MCAL}} \rangle > z_{l_j} + 0.1, \quad (18)$$

which is the general form of the estimator. Here, we neglected the selection response term, which McClintock et al. (2018) found to be subdominant compared to the shear response  $R_\gamma^T$ , where the superscript refers to the response matrix rotated into the tangential frame. In the above estimator, the weighting and selection is performed based on the mean METACALIBRATION based redshift estimates  $\langle z_{s_i}^{\text{MCAL}} \rangle$ , while  $\Sigma_{\text{crit},i,j}^{-1}$  is calculated using a random draw from the MOF-based redshift  $P(z)$ .

Following equations (7) and (8), we find the contamination rate to be:

$$f_{\text{cl}} = \frac{\sum_{\text{cl}} \hat{\omega}_{i,j} R_{\gamma,i,j}^T \Sigma_{\text{crit},i,j}^{-1}}{\sum \hat{\omega}_{i,j} R_{\gamma,i,j}^T \Sigma_{\text{crit},i,j}^{-1}}. \quad (19)$$

We perform the  $P(z)$  decomposition in a setup identical to Section 3.2.3, but using weights according to equation (19), and make use of a randomly selected, representative subsample of the source–lens pairs from McClintock et al. (2018). The detailed description of our results is presented in Section 4, while the boost profiles themselves are shown on Fig. 6.

### 3.3.3 Sensitivity to background component choice

The performance of the  $P(z)$  decomposition method is dependent on how well the ansatz for the background component resemble the p.d.f. of actual background galaxies. Furthermore, the average

$P(z)$  estimated for a galaxy sample may contain minor features (e.g. wiggles and peaks) which depend on the internal setup of the photometric redshift algorithm (e.g. distribution of templates within the BPZ algorithm), and do not themselves relate to the physical distribution of galaxies (Bonnert et al. 2015; Melchior et al. 2017). Differences originating from these non-physical reasons may also impact the robustness of the contamination estimates.

We test the self-consistency of the decomposition method and its sensitivity to the minor features in the estimated  $P(z)$ s by extending the fiducial  $P(z)$  decomposition analysis with a second step. In this second step, the reference  $P(z)$  model is updated from the ‘field’  $P(z)$  to the observed  $P(z)$  at  $R \approx 1$  Mpc minus the Gaussian cluster model found in the previous step. The  $f_{cl}$  fit is then repeated with this new reference  $P(z)$  component, while keeping the position and width of the Gaussian cluster model component unchanged. The resulting boost factor profiles are shown on Fig. 7, overlaid with the fiducial boost factor profiles. The two iterations agree very well, and following the approach used in Section 3.2.6, we estimate a relative systematic uncertainty of  $<1$  per cent, motivating that the choice for the background component propagates to only a negligible difference in the final contamination profiles.

### 3.3.4 Comparison with correlation-based boost factors

An alternative way for estimating boost factors is via the angular clustering of source galaxies around clusters, as only the contaminating galaxies are correlated with the cluster (Sheldon et al. 2004; Applegate et al. 2014; Hoekstra et al. 2015; Simet et al. 2017; Leauthaud et al. 2017). We calculate this correlation function via the estimator:

$$\text{Corr} = \frac{N_R}{N_D} \cdot \frac{DD}{RR} - 1, \quad (20)$$

where DD and RR are defined as  $\sum \omega \mathbf{R}_y^T \Sigma_{crit}^{-1}$  around redMaPPer clusters and random points respectively, while  $N_D$  refers to the number of clusters, and  $N_R$  to the number of random points (Landy & Szalay 1993). The results of this measurement are shown on Fig. 7. The correlation function estimates are, for many cluster samples, preferentially lower than the  $P(z)$  decomposition estimates, especially at the two lower redshift selections.

This can be understood as clusters impacting the spatial distribution of source galaxies in ways other than contamination by cluster galaxies: e.g. the density and blending of cluster members may lead to a bias against selecting sources near clusters (Simet et al. 2015; Leauthaud et al. 2017; Zuntz et al. 2017), which can explain the preferential lower estimates. Such effects cannot be captured by random points, as they relate to the presence of the cluster in the line of sight, and are not well characterized for the DES Y1 METACALIBRATION shear catalogues. By contrast, the  $P(z)$  decomposition method is insensitive to colour-agnostic fluctuations in the source selection, and to the number density profile of source galaxies.

## 4 RESULTS FOR DES Y1 DATA

### 4.1 Boost factor estimates

We present the contamination rate estimates from applying our method to the DES Y1 data in Fig. 6. In the present calculation, we consider all cluster selections, but note that in McClintock et al. (2018) only the  $\lambda > 20$  clusters enter the determination of the MOR. The qualitative behaviour of the contamination rate

profiles agrees well with the theoretical expectation of decreasing contamination with increasing radius. As expected, the amplitude of the contamination increases with cluster richness. Furthermore, the contamination rates are higher for low-redshift clusters, as for those fainter cluster member galaxies can also be detected, whose photometric redshifts are less accurate.

We find that the ‘peak’ in the  $P(z)$  due to contaminating galaxies is very prominent at low radii for all cluster bins, and the presence of this feature is critical for the applicability of the decomposition method. The best-fitting parameters of the Gaussian  $P_{\text{memb}}(z)$  model are presented in Fig. 8, along with the used prior ranges. The means of these best-fitting Gaussian  $P_{\text{memb}}(z)$  distributions differ from the redshift ranges of the clusters. However, this is expected from the way source galaxies are selected in the DES analysis: only those cluster member galaxies enter the source selection whose estimated mean redshift scatters towards higher redshifts.

The contamination rate profiles  $f_{cl}(R)$  shown in Fig. 6 can be directly translated into a multiplicative correction factor  $\mathcal{B} \equiv (1 - f_{cl})^{-1}$  necessary for recovering an unbiased estimate on  $\Delta\Sigma$  via equation (9).

### 4.2 Analytic boost factor model

We model the boost factor profile using a Navarro–Frenk–White (NFW) profile (Navarro et al. 1996):

$$\mathcal{B}_{\text{model}}(R) = 1 + B_0 \frac{1 - F(x)}{x^2 - 1}, \quad (21)$$

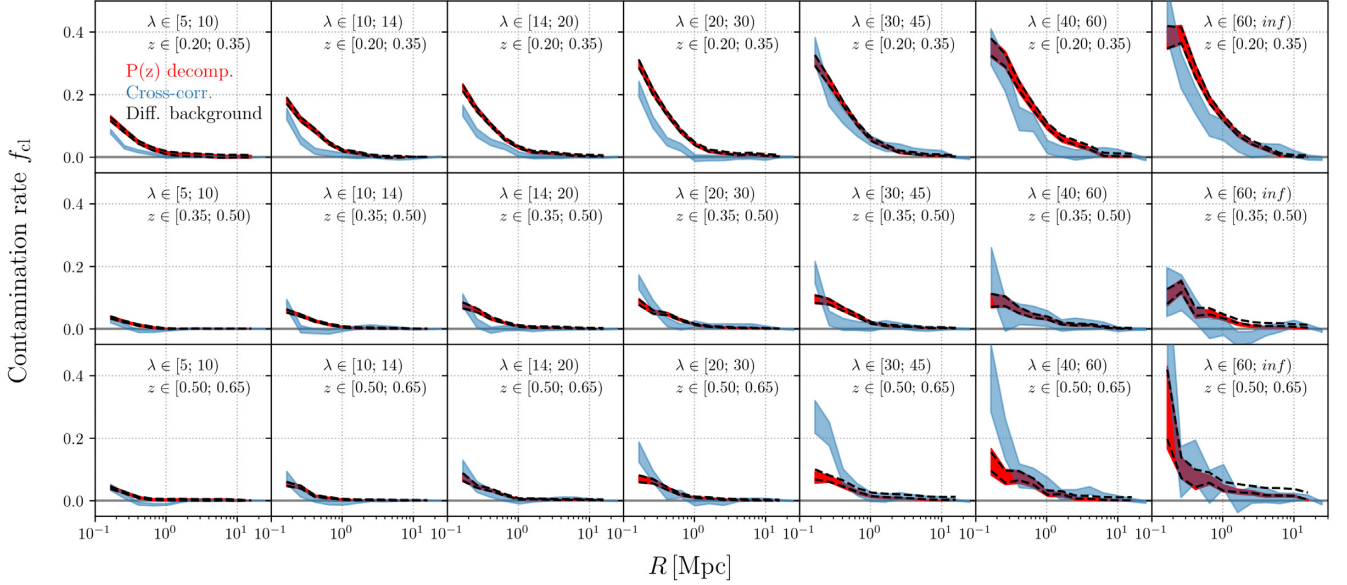
where  $x = R/R_s$ , and

$$F(x) = \begin{cases} \frac{\tan^{-1} \sqrt{x^2 - 1}}{\sqrt{x^2 - 1}} & : x > 1 \\ 1 & : x = 1 \\ \frac{\tanh^{-1} \sqrt{1 - x^2}}{\sqrt{1 - x^2}} & : x < 1 \end{cases}. \quad (22)$$

This model has two free parameters per cluster bin:  $B_0$  and  $R_s$  which characterize the amplitude and scale radius of the correction profile, respectively.

The best-fitting boost model profiles are overlaid on Fig. 6 to the raw contamination rate estimates. Following the approach used in Section 3.2.6 we estimate a relative systematic uncertainty of  $<1$  per cent. In McClintock et al. (2018), these fits are performed in a joint likelihood analysis together with the mass profile model and systematic corrections. This way the estimated statistical uncertainty of the boost factors is propagated self-consistently into their final mass constraints. The model parameters are not tied to the mass parameters of clusters to allow for freedom in describing the boost factors. Due to the excellent performance of the  $P(z)$  decomposition method in our tests, and as the systematic uncertainty found in a simulated environment was subdominant compared to the JK error estimate, McClintock et al. (2018) did not assume any additional systematic uncertainty to this source of systematic error.

We note that in the Monte Carlo chains run by McClintock et al. (2018), the  $R_s$  and  $B_0$  parameters were found to be degenerate, hence the increase in actual contamination does not translate into an obvious increase in  $B_0$ . However this was found to propagate into only a mild change in the boost factor profile over the studied radial range. The recovered cluster masses were robust against degeneracy in the boost factor model parameters, and are not significantly impacted. Nevertheless, we find that the scale radius of the contamination component is typically at least twice as large as the scale radius of the NFW halo. This is consistent with the expectation that the galaxy



**Figure 7.** Comparison of the  $P(z)$  decomposition based contamination estimates (red) with the cross-correlation-based estimates (blue) across the different richness and redshift bins. Black: results from the background sensitivity test. The *black* background sensitivity results are found to agree very well with the fiducial red curves, while the blue correlation-based estimates appear to be globally biased low.

distribution of clusters can be described as an NFW distribution with lower (approximately half) concentration than the underlying dark matter halo (Budzynski et al. 2012).

## 5 SUMMARY AND CONCLUSIONS

In this study, we carried out a detailed method validation on the  $P(z)$  decomposition cluster member contamination estimation algorithm proposed by Gruen et al. (2014) and Melchior et al. (2017). This approach relies on the decomposition of the average redshift  $P(z)$ s of source galaxies around galaxy clusters into a cluster member and background component, to obtain an estimate on the relative number of contaminating galaxies which are mistakenly included in the source galaxy catalogue. Since its inception this method has been used by studies ranging from the DES Science Verification cluster mass calibration (Melchior et al. 2017), to cluster weak lensing studies focusing on the detection of the *splashback*-feature (Chang et al. 2018), and to the mass calibration of SPT-selected clusters (Stern et al. 2018). It also serves as an important constituent of the weak lensing mass calibration on DES Y1 (McClintock et al. 2018), which will be used in deriving cosmological constraints based on the number counts of optically identified galaxy clusters (DES Collaboration).

In order to demonstrate the applicability of  $P(z)$  decomposition-based boost factors, we performed a series of tests benchmarking various aspects of the approach. We find the following:

- (i) The method performed well in a mock survey simulation (Section 3.2), yielding excellent agreement between the estimated contamination rates and the actual true number of contaminants extracted from the truth catalogues of the simulation.
- (ii) Within the mock analysis, we investigated the validity of the Gaussian ansatz for the cluster  $P(z)$  component (Section 3.2.4). We found that it recovers the approximate redshift and width of the peak within the  $P(z)$  of the contaminating galaxies. Furthermore, the Gaussian ansatz did not appreciably bias the estimated contamination.

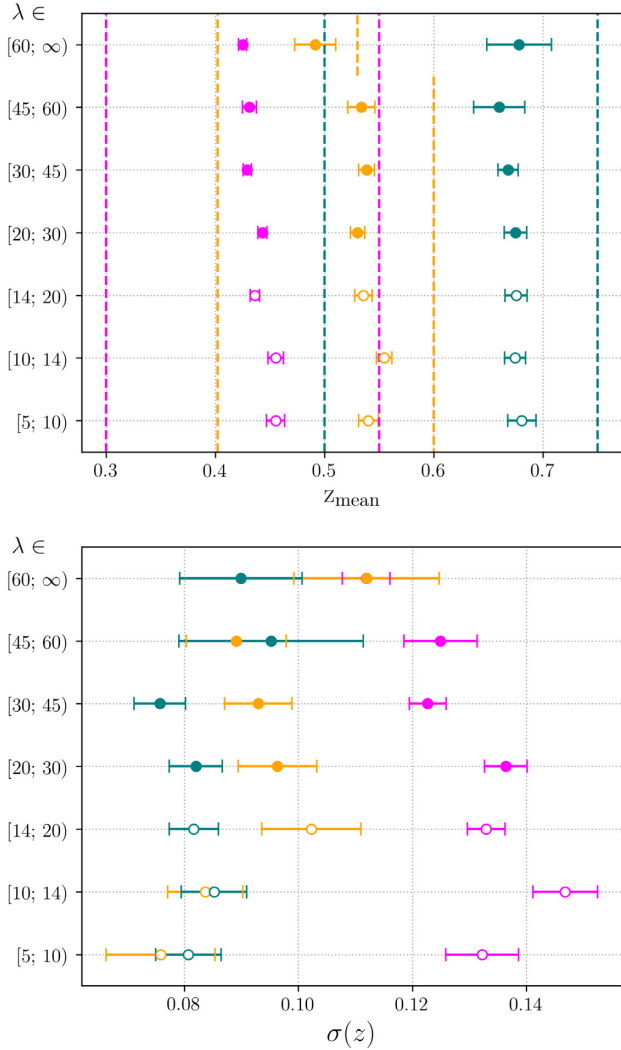
- (iii) We tested the sensitivity of the contamination estimates to the choice of the background  $P(z)$  component on DES Y1 data (Section 3.3.3), and found an excellent agreement between the boost factors derived via the fiducial and alternative background components.

- (iv) We compared the method with an alternative, transverse correlation-based contamination estimate in Section 3.3.4. We found indications that the alternative method is preferentially underestimating the contaminations, which is likely an imprint of the radial source galaxy selection function. This is expected to impact the decomposition predictions to a lesser extent as it does not make use of the number profile of sources.

Excluding galaxies from the source catalogue can also reduce the cluster member contamination, however it may also reduce the statistical power of the measurement if the exclusion criteria are too broad. Hence, it presents a trade-off in the total error budget between the systematic uncertainty originating from boost factors and statistical uncertainty such as shape noise. However based on the consistency tests presented in this paper, and on the fact that boost factors played a strongly subdominant role in the total error budget of McClintock et al. (2018), it appears that the  $P(z)$  decomposition method is sufficient to provide boost factor estimates for current cluster weak lensing analyses. We note that this determination is dependent on the characteristics of the sky survey e.g. depth, area, and number of filters. Notably cluster weak lensing studies such as Medezinski et al. (2018b, a) and Miyatake et al. (2018) in the ongoing Hyper Suprime-Cam Survey (Aihara et al. 2018) favoured the approach of trying to excluding cluster member galaxies via colour-colour or  $P(z)$  cuts.

During the DES Y1 analysis we propagated uncertainties by making use of a simple analytic model – an NFW profile – to describe the boost factor correction. Previously Melchior et al. (2017) also used an analytic model, while others such as Chang et al. (2018) and Stern et al. (2018) chose to directly use the recovered boost factor profiles in correcting their  $\Delta\Sigma$  measurements. While the NFW model was found to be a sufficient description for the





**Figure 8.** Best-fitting parameters for the Gaussian  $P_{\text{memb}}$  component found for the  $P(z)$  decomposition in the DES Y1 data. The colours indicate the various cluster redshift bins:  $z \in [0.2, 0.35)$ ,  $z \in [0.35, 0.5)$ , and  $z \in [0.5, 0.65)$  is denoted by magenta, orange, and green, respectively. Top panel: mean redshift of the cluster member components. The dashed coloured lines indicate the prior range for the mean of the cluster component. Bottom panel: standard deviation of the cluster member components.

analysis of McClintock et al. (2018), it is likely that with the increasing precision of future studies more complex boost factor models might become necessary.

We quantified several possible sources of systematic uncertainty impacting the  $P(z)$  decomposition method, finding  $<1$  percent relative systematic uncertainty based on benchmarks on mock observations,  $<1$  per cent relative systematic uncertainty originating from the choice of the background  $P(z)$  component, and  $<1$  per cent relative systematic uncertainty from requiring good global agreement between the numerical boost factor estimates and the analytic model. From these contributions we estimate that the decomposition method under optimal circumstances can provide boost factor estimates with approximately 2 per cent relative global systematic uncertainty. However, we note that specific circumstances such as the performance of the photometric redshift algorithm, or the source galaxy selection function will impact the accuracy and precision of the  $P(z)$  decomposition method.

## ACKNOWLEDGEMENTS

This paper has gone through internal review by the DES collaboration. TNV and SS are supported by the SFB-Transregio 33 ‘The Dark Universe’ by the Deutsche Forschungsgemeinschaft (DFG) and the DFG cluster of excellence ‘Origin and Structure of the Universe’. JD, DG, and RHW received support from the U.S. Department of Energy under contract number DE-AC02-76SF00515. DG was also supported by *Chandra* award number GO8-19101A, issued by the Chandra X-ray Observatory Center. ER was supported by the DOE grant DE-SC0015975, by the Sloan Foundation, grant FG-2016-6443, and the Cottrell Scholar program of the Research Corporation for Science Advancement. AvdL is supported by the U.S. Department of Energy under award number DE-SC0018053.

Part of our computations have been carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP). A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This research made use of computational resources at SLAC National Accelerator Laboratory, and the authors thank the SLAC computational team for support. Some of the computing for this project was performed on the Sherlock cluster. We would like to thank Stanford University and the Stanford Research Computing Center for providing computational resources and support that contributed to these research results. This research also used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under contract number DE-AC02-05CH11231.

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

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l’Espai (IEEC/CSIC), the Institut de Física d’Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the

University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under grant numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015-71825, ESP2015-66861, FPA2015-68048, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAAS-TRO), through project number CE110001020, and the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under contract number DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

## REFERENCES

- Aihara H. et al., 2018, *PASJ*, 70, S4
- Allen S. W., Evrard A. E., Mantz A. B., 2011, *ARA&A*, 49, 409
- Applegate D. E. et al., 2014, *MNRAS*, 439, 48
- Bartelmann M., Schneider P., 2001, *Phys. Rep.*, 340, 291
- Battaglia N. et al., 2016, *J. Cosmol. Astropart. Phys.*, 8, 013
- Benítez N., 2000, *ApJ*, 536, 571
- Bleem L. E. et al., 2015, *ApJS*, 216, 27
- Bonnett C. et al., 2016, *Phys. Rev. D*, 94, 042005
- Budzynski J. M., Kposov S. E., McCarthy I. G., McGee S. L., Belokurov V., 2012, *MNRAS*, 423, 104
- Butcher H., Oemler A., Jr., 1978, *ApJ*, 219, 18
- Chang C. et al., 2015, *ApJ*, 801, 73
- Chang C. et al., 2018, *ApJ*, 864, 83
- Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, *AJ*, 132, 926
- Conroy C., Wechsler R. H., Kravtsov A. V., 2006, *ApJ*, 647, 201
- Cooper M. C. et al., 2011, *ApJS*, 193, 14
- Crocce M., Puelbas S., Scoccimarro R., 2006, *MNRAS*, 373, 369
- DeRose J. et al., 2019, preprint ([arXiv:1901.02401](https://arxiv.org/abs/1901.02401))
- DES Collaboration, 2018, *Phys. Rev. D*, 98, 043526
- Dietrich J. P. et al., 2019, *MNRAS*, 483, 2871
- Dodelson S., Heitmann K., Hirata C., Honscheid K., Roodman A., Seljak U., Slosar A., Trodden M., 2016, preprint ([arXiv:1604.07626](https://arxiv.org/abs/1604.07626))
- Drlica-Wagner A. et al., 2017, *ApJS*, 235, 33
- Efron B., 1982, The Jackknife, the Bootstrap and other Resampling Plans, CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM). SIAM, Philadelphia
- Gruen D., Brimiouille F., 2017, *MNRAS*, 468, 769
- Gruen D. et al., 2019, *MNRAS*, 488, 4389
- Gruen D. et al., 2014, *MNRAS*, 442, 1507
- Hansen S. M., Sheldon E. S., Wechsler R. H., Koester B. P., 2009, *ApJ*, 699, 1333
- Hao J., Kubo J. M., Feldmann R., Annis J., Johnston D. E., Lin H., McKay T. A., 2011, *ApJ*, 740, 39
- Hildebrandt H. et al., 2017, *MNRAS*, 465, 1454
- Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685
- Hoyle B. et al., 2018, *MNRAS*, 478, 592
- Huang H.-J., Mandelbaum R., Freeman P. E., Chen Y.-C., Roze E., Rykoff E., 2018, *MNRAS*, 474, 4772
- Huff E., Mandelbaum R., 2017, preprint ([arXiv:1702.02600](https://arxiv.org/abs/1702.02600))
- Kelly P. L. et al., 2014, *MNRAS*, 439, 28
- Koester B. P. et al., 2007, *ApJ*, 660, 221
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Leauthaud A. et al., 2017, *MNRAS*, 467, 3024
- Lehmann B. V., Mao Y.-Y., Becker M. R., Skillman S. W., Wechsler R. H., 2017, *ApJ*, 834, 37
- MacCrann N. et al., 2018, *MNRAS*, 480, 4614
- Mandelbaum R. et al., 2018, *PASJ*, 70, S25
- Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, *MNRAS*, 406, 1759
- Mantz A. B. et al., 2015, *MNRAS*, 446, 2205
- McClintock T. et al., 2019, *MNRAS*, 482, 1352
- Medezinski E. et al., 2018a, *PASJ*, 70, S28
- Medezinski E. et al., 2018b, *PASJ*, 70, 30
- Melchior P. et al., 2017, *MNRAS*, 469, 4899
- Miyatake H. et al., 2019, *ApJ*, 875, 17
- Murata R., Nishimichi T., Takada M., Miyatake H., Shirasaki M., More S., Takahashi R., Osato K., 2017, *ApJ*, 854, 120
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Okabe N., Smith G. P., 2016, *MNRAS*, 461, 3794
- Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, *ApJ*, 771, 30
- Rykoff E. S. et al., 2014, *ApJ*, 785, 104
- Rykoff E. S. et al., 2016, *ApJS*, 224, 1
- Schrabback T. et al., 2018, *MNRAS*, 474, 2635
- Sheldon E. S., Huff E. M., 2017, *ApJ*, 841, 24
- Sheldon E. S. et al., 2004, *AJ*, 127, 2544
- Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R. F. J., Sand D. J., Graham M. L., 2015, *A&A*, 575, A48
- Simet M., Battaglia N., Mandelbaum R., Seljak U., 2017, *MNRAS*, 466, 3663
- Simet M., McClintock T., Mandelbaum R., Roze E., Rykoff E., Sheldon E., Wechsler R. H., 2017, *MNRAS*, 466, 3103
- Springel V., 2005, *MNRAS*, 364, 1105
- Stern C. et al., 2019, *MNRAS*, 485, 69
- Suchyta E. et al., 2016, *MNRAS*, 457, 786
- Sunyaev R. A., Zeldovich Y. B., 1970, Comment. Astrophys. Space Phys., 2, 66
- Sunyaev R. A., Zeldovich Y. B., 1972, Comment. Astrophys. Space Phys., 4, 173
- Tanaka M. et al., 2018, *PASJ*, 70, S9
- Truemper J., 1993, *Science*, 260, 1769
- von der Linden A. et al., 2014a, *MNRAS*, 439, 2
- von der Linden A. et al., 2014b, *MNRAS*, 443, 1973
- Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Roze E., 2013, *Phys. Rep.*, 530, 87
- Zuntz J. et al., 2018, *MNRAS*, 481, 1149

<sup>1</sup>Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany

<sup>2</sup>Fakultät für Physik, Universitäts-Sternwarte, Ludwig-Maximilians-Universität München, Scheinerstr. 1, D-81679 München, Germany

<sup>3</sup>Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA

- <sup>4</sup>Kavli Institute for Particle Astrophysics and Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA
- <sup>5</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
- <sup>6</sup>Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
- <sup>7</sup>Department of Physics, University of Arizona, Tucson, AZ 85721, USA
- <sup>8</sup>Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
- <sup>9</sup>Department of Physics, The Ohio State University, Columbus, OH 43210, USA
- <sup>10</sup>Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
- <sup>11</sup>Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
- <sup>12</sup>University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
- <sup>13</sup>Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA
- <sup>14</sup>Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA
- <sup>15</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- <sup>16</sup>CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France
- <sup>17</sup>Institut d'Astrophysique de Paris, Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, F-75014, Paris, France
- <sup>18</sup>Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- <sup>19</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
- <sup>20</sup>Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- <sup>21</sup>Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
- <sup>22</sup>National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA
- <sup>23</sup>Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona) Spain
- <sup>24</sup>Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
- <sup>25</sup>Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
- <sup>26</sup>Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India

- <sup>27</sup>Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany
- <sup>28</sup>Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany
- <sup>29</sup>Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
- <sup>30</sup>Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
- <sup>31</sup>Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain
- <sup>32</sup>Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain
- <sup>33</sup>Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
- <sup>34</sup>Instituto de Física Teórica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain
- <sup>35</sup>Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland
- <sup>36</sup>Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
- <sup>37</sup>Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
- <sup>38</sup>Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia
- <sup>39</sup>Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP 05314-970, Brazil
- <sup>40</sup>George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
- <sup>41</sup>Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA
- <sup>42</sup>Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain
- <sup>43</sup>Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton BN1 9QH, UK
- <sup>44</sup>School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK
- <sup>45</sup>Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, SP 13083-859 Campinas, Brazil
- <sup>46</sup>Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.