

Dark Energy Survey Year 1 results: weak lensing mass calibration of redMaPPer galaxy clusters

T. McClintock,¹★ T. N. Varga,^{2,3} D. Gruen,^{4,5}† E. Rozo,¹ E. S. Rykoff,^{4,5} T. Shin,⁶ P. Melchior,⁷ J. DeRose,^{8,4} S. Seitz,^{2,3} J. P. Dietrich,^{9,10} E. Sheldon,¹¹ Y. Zhang,¹² A. von der Linden,¹³ T. Jeltema,¹⁴ A. B. Mantz,⁴ A. K. Romer,¹⁵ S. Allen,⁸ M. R. Becker,^{8,4} A. Bermeo,¹⁵ S. Bhargava,¹⁵ M. Costanzi,³ S. Everett,¹⁴ A. Farahi,¹⁶ N. Hamaus,³ W. G. Hartley,^{17,18} D. L. Hollowood,¹⁴ B. Hoyle,^{2,3} H. Israel,¹⁰ P. Li,¹⁹ N. MacCrann,^{20,21} G. Morris,⁵ A. Palmese,^{17,12} A. A. Plazas,²² G. Pollina,^{9,3} M. M. Rau,^{19,3} M. Simet,^{22,23} M. Soares-Santos,²⁴ M. A. Troxel,^{20,21} C. Vergara Cervantes,¹⁵ R. H. Wechsler,^{8,4,5} J. Zuntz,²⁵ T. M. C. Abbott,²⁶ F. B. Abdalla,^{17,27} S. Allam,¹² J. Annis,¹² S. Avila,²⁸ S. L. Bridle,²⁹ D. Brooks,¹⁷ D. L. Burke,^{4,5} A. Carnero Rosell,^{30,31} M. Carrasco Kind,^{32,33} J. Carretero,³⁴ F. J. Castander,^{35,36} M. Crocce,^{35,36} C. E. Cunha,⁴ C. B. D’Andrea,⁶ L. N. da Costa,^{30,31} C. Davis,⁴ J. De Vicente,³⁷ H. T. Diehl,¹² P. Doel,¹⁷ A. Drlica-Wagner,¹² A. E. Evrard,^{38,16} B. Flaugher,¹² P. Fosalba,^{35,36} J. Frieman,^{12,39} J. García-Bellido,⁴⁰ E. Gaztanaga,^{35,36} D. W. Gerdes,^{38,16} T. Giannantonio,^{41,42,3} R. A. Gruendl,^{32,33} G. Gutierrez,¹² K. Honscheid,^{20,21} D. J. James,⁴³ D. Kirk,¹⁷ E. Krause,^{44,22} K. Kuehn,⁴⁵ O. Lahav,¹⁷ T. S. Li,^{12,39} M. Lima,^{46,30} M. March,⁶ J. L. Marshall,⁴⁷ F. Menanteau,^{32,33} R. Miquel,^{48,34} J. J. Mohr,^{9,10,2} B. Nord,¹² R. L. C. Ogando,^{30,31} A. Roodman,^{4,5} E. Sanchez,³⁷ V. Scarpine,¹² R. Schindler,⁵ I. Sevilla-Noarbe,³⁷ M. Smith,⁴⁹ R. C. Smith,²⁶ F. Sobreira,^{50,30} E. Suchyta,⁵¹ M. E. C. Swanson,³³ G. Tarle,¹⁶ D. L. Tucker,¹² V. Vikram,⁵² A. R. Walker,²⁶ and J. Weller^{9,2,3} DES Collaboration

Affiliations are listed at the end of the paper

Accepted 2018 September 25. Received 2018 September 14; in original form 2018 May 2

ABSTRACT

We constrain the mass–richness scaling relation of redMaPPer galaxy clusters identified in the Dark Energy Survey Year 1 data using weak gravitational lensing. We split clusters into 4×3 bins of richness λ and redshift z for $\lambda \geq 20$ and $0.2 \leq z \leq 0.65$ and measure the mean masses of these bins using their stacked weak lensing signal. By modelling the scaling relation as $\langle M_{200m} | \lambda, z \rangle = M_0 (\lambda/40)^F ((1+z)/1.35)^G$, we constrain the normalization of the scaling relation at the 5.0 per cent level, finding $M_0 = [3.081 \pm 0.075(\text{stat}) \pm 0.133(\text{sys})] \cdot 10^{14} M_\odot$ at $\lambda = 40$ and $z = 0.35$. The recovered richness scaling index is $F = 1.356 \pm 0.051(\text{stat}) \pm 0.008(\text{sys})$ and the redshift scaling index $G = -0.30 \pm 0.30(\text{stat}) \pm 0.06(\text{sys})$. These are the tightest measurements of the normalization and richness scaling index made to date from a weak lensing experiment. We use a semi-analytic covariance matrix to characterize the statistical errors in the recovered weak lensing profiles. Our analysis accounts for the following sources of systematic error: shear and photometric redshift errors, cluster miscentring, cluster member

* E-mail: tmclintock@email.arizona.edu

† Einstein Fellow.

dilution of the source sample, systematic uncertainties in the modelling of the halo–mass correlation function, halo triaxiality, and projection effects. We discuss prospects for reducing our systematic error budget, which dominates the uncertainty on M_0 . Our result is in excellent agreement with, but has significantly smaller uncertainties than, previous measurements in the literature, and augurs well for the power of the DES cluster survey as a tool for precision cosmology and upcoming galaxy surveys such as LSST, Euclid, and WFIRST.

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

1 INTRODUCTION

Galaxy clusters have the potential to be the most powerful cosmological probe (Dodelson et al. 2016). Current constraints are dominated by uncertainties in the calibration of cluster masses (e.g. Rozo et al. 2010; Mantz et al. 2015a; Planck Collaboration XXIV 2016). Weak lensing allows us to determine the mass of galaxy clusters: gravitational lensing of background galaxies by foreground clusters induces a tangential alignment of the background galaxies around the foreground cluster. This alignment is a clear observational signature predicted from clean, well-understood physics. Moreover, the resulting signal is explicitly sensitive to all of the cluster mass, not just its baryonic component, and is insensitive to the dynamical state of the cluster. For all these reasons, weak lensing is the most robust method currently available for calibrating cluster masses. It is therefore not surprising that the community has invested in a broad range of weak lensing experiments specifically designed to calibrate the masses of galaxy clusters (Applegate et al. 2014; von der Linden et al. 2014a,b; Hoekstra et al. 2015; Mantz et al. 2015b; Okabe & Smith 2016; Dietrich et al. 2017; Melchior et al. 2017; Simet et al. 2017; Medezinski et al. 2018b; Miyatake et al. 2018; Murata et al. 2018).

The Dark Energy Survey (DES) is a 5000 square degree photometric survey of the southern sky. It uses the 4-m Blanco Telescope and the Dark Energy Camera (Flaugher et al. 2015) located at the Cerro Tololo Inter-American Observatory. As its name suggests, the primary goal of the DES is to probe the physical nature of dark energy, in addition to constraining the properties and distribution of dark matter. Owing to its large area, depth, and image quality, at its conclusion DES will support optical identification of $\sim 100\,000$ galaxy clusters and groups up to redshift $z \approx 1$. We use galaxy clusters identified using the redMaPPer algorithm (Rykoff et al. 2014), which assigns each cluster a photometric redshift and optical richness λ of red galaxies. To fully utilize these clusters, one must understand mass-observable relations (MORs), such as that between cluster mass and optical richness. Weak lensing can establish this relation – with high statistical uncertainty for individual clusters, but low systematic uncertainty in the mean mass scale derived from the joint signal of large samples.

In this work, we use stacked weak lensing to measure the mean galaxy cluster mass of redMaPPer galaxy clusters identified in DES Year 1 (Y1) data. We use these data to calibrate the mass–richness–redshift relation of these clusters. In Melchior et al. (2017), we provided a first calibration of this relation using DES Science Verification (SV) data. There, we were able to achieve a 9.2 per cent statistical and 5.1 per cent systematic uncertainty. Here, we update that result using the first year of regular DES observations, incorporating a variety of improvements to the analysis pipeline. Our results provide the tightest, most accurate calibration of the richness–mass

relation of galaxy clusters to date, at 2.4 per cent statistical and 4.3 per cent systematic uncertainty.

The structure of this paper is as follows. In Section 2, we introduce the DES Y1 data used in this work. In Section 3, we describe our methodology for obtaining ensemble cluster density profiles from stacked weak lensing shear measurements, with a focus on updates relative to Melchior et al. (2017). A comprehensive set of tests and corrections for systematic effects is presented in Section 4. The model of the lensing data and the inferred stacked cluster masses are given in Section 5. The main result, the mass–richness–redshift relation of redMaPPer clusters in DES, is presented in Section 6. We compare our results to other published works in the literature in Section 7, discuss systematic improvements made in this work compared to Melchior et al. (2017) in Section 8, and conclude in Section 9. In Appendix A, we present the DES Y1 redMaPPer catalogue used in this work for public use. Supplementary information on the analysis is given in additional appendices.

Unless otherwise stated, we assume a flat Λ 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. Finally, unless otherwise noted all cluster masses refer to M_{200m} . That is, cluster mass is defined as the mass enclosed within a sphere whose average density is 200 times higher than the mean cosmic matter density $\bar{\rho}_m$ at the cluster’s redshift, matching the mass definition used in the cosmological analyses that make use of our calibration.

2 THE DES YEAR 1 DATA

DES started its main survey operations in 2013, with the Year One (Y1) observational season running from August 31, 2013 to February 9, 2014 (Drlica-Wagner et al. 2018). During this period, 1839 deg² of the southern sky were observed in three to four tilings in each of the four DES bands g, r, i, z , as well as $\sim 1800 \text{ deg}^2$ in the Y -band. The resulting imaging is shallower than the SV data release but covers a significantly larger area. In this study, we utilize approximately 1500 deg² of the main survey, split into two large non-contiguous areas. This is a reduction from the 1800 deg² area due to a series of veto masks. These masks include masks for bright stars and the Large Magellanic Cloud, among others. The two non-contiguous areas are the ‘SPT’ area (1321 deg²), which overlaps the footprint of the South Pole Telescope Sunyaev-Zel’dovich Survey (Carlstrom et al. 2011), and the ‘S82’ area (116 deg²), which overlaps the Stripe-82 deep field of the Sloan Digital Sky Survey (SDSS; Annis et al. 2014). The DES Y1 footprint is shown in Fig. 1.

In the following, we briefly describe the main data products used in this analysis, and refer the reader to the corresponding papers for more details. The input photometric catalog, as well as the photometric redshift and weak lensing shape catalogues used in this study have already been employed in the cosmological analysis combin-

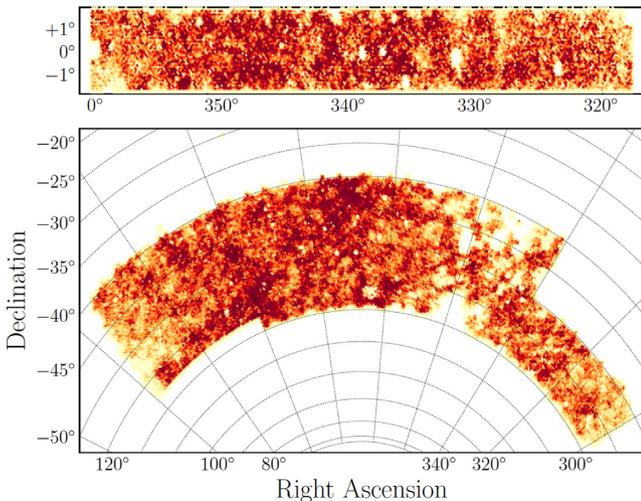


Figure 1. Surface density of source galaxies in the METACALIBRATION catalogue within the DES Y1 footprint in the ‘S82’ field (top) and the ‘SPT’ field (bottom).

ing galaxy clustering and weak lensing by the DES collaboration (DES Collaboration et al. 2018).

2.1 Photometric catalogue

Input photometry for the redMaPPer cluster finder (Section 2.2) and photometric redshifts (Section 2.4) were derived from the DES Y1A1 Gold catalogue (Drlica-Wagner et al. 2018). Y1A1 Gold is the science-quality internal photometric catalogue of DES created to enable cosmological analyses. This data set includes a catalogue of objects as well as maps of survey depth and foreground masks, and star-galaxy classification. In this work, we make use of the multi-epoch, multi-object fitting (MOF) composite model (CM) galaxy photometry. The MOF photometry simultaneously fits a point spread function (psf)-convolved galaxy model to all available epochs and bands for each object, while subtracting and masking neighbours. The typical 10σ limiting magnitude inside $2''$ diameter apertures for galaxies in Y1A1 Gold using MOF CM photometry is $g \approx 23.7$, $r \approx 23.5$, $i \approx 22.9$, and $z \approx 22.2$. Due to its low depth and calibration uncertainty, we do not use Y -band photometry for shape measurement or photometric redshift estimation.

The galaxy catalogue used for the redMaPPer cluster finder is constructed as follows. Bad objects that are determined to be catalogue artefacts, including having unphysical colours, astrometric discrepancies, and PSF model failures are rejected (Section 7.4 Drlica-Wagner et al. 2018). Galaxies are then selected via the more complete MODEST_CLASS classifier (Section 8.1 Drlica-Wagner et al. 2018). Only galaxies that are brighter in z band than the local 10σ limiting magnitude are used by redMaPPer. The average survey limiting magnitude is deep enough to image a $0.2L_*$ galaxy at $z \approx 0.7$. Finally, we remove galaxies in regions that are contaminated by bright stars, bright nearby galaxies, globular clusters, and the Large Magellanic Cloud.

2.2 Cluster catalog

We use a volume-limited sample of galaxy clusters detected in the DES Y1 photometric data using the redMaPPer cluster finding algorithm v6.4.17 (Rykoff et al. 2014, 2016). This redMaPPer version is

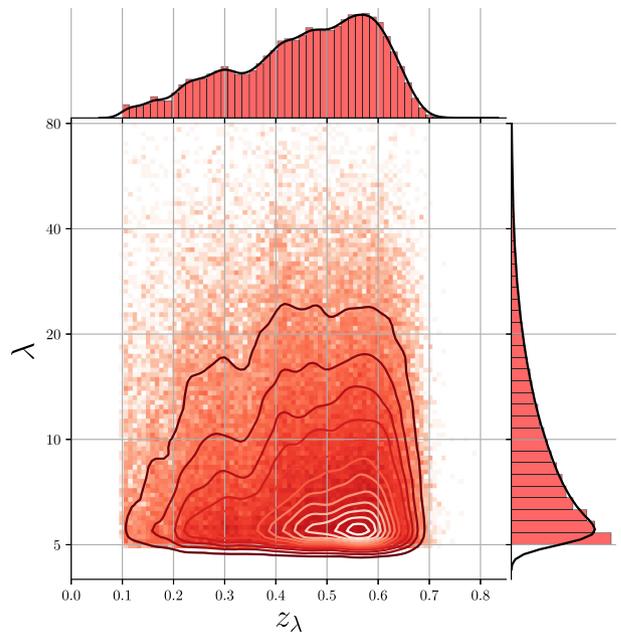


Figure 2. Redshift-richness distribution of redMaPPer clusters in the volume-limited DES Y1 cluster catalogue, overlaid with density contours to highlight the densest regions. At the top and on the right are histograms of the projected quantities, z_λ and λ , respectively, with smooth kernel density estimates overlaid.

fundamentally the same as the v6.3 algorithm described in Rykoff et al. (2016), with minor updates.

Two versions of the redMaPPer cluster catalog are generated: a ‘flux limited’ version, which includes high-redshift clusters for which the richness requires extrapolation along the cluster luminosity function, and one that is locally volume-limited. By ‘locally volume-limited’ we mean that at each point in the sky, a galaxy cluster is included in the sample if and only if all cluster galaxies brighter than the luminosity threshold used to define cluster richness in redMaPPer lie above 10σ in z , 5σ in i and r , and 3σ in g according to the survey MOF depth maps (Drlica-Wagner et al. 2018). That is, no extrapolation in luminosity is required when estimating cluster richness. At the threshold, the galaxy sample is >90 – 95 per cent complete. It is this volume-limited cluster sample that is used in follow-up work deriving cosmological constraints from the abundance of galaxy clusters. Consequently, we focus exclusively on this volume-limited sample in this work. It contains more than 76 000 clusters down to $\lambda > 5$, of which more than 6500 are above $\lambda = 20$. The format of the catalogues are described in Appendix A.

redMaPPer identifies galaxy clusters as overdensities of red-sequence galaxies. Starting from an initial set of spectroscopic seed galaxies, the algorithm iteratively fits a model for the local red-sequence, and finds cluster candidates while assigning a membership probability to each potential member. Clusters are centred on bright galaxies selected using an iteratively self-trained matched-filter method. The method allows for the inherent ambiguity of selecting a central galaxy by assigning a probability to each galaxy of being the central galaxy of the cluster. The final membership probabilities of all galaxies in the field are assigned based on spatial, colour, and magnitude filters.

The distribution of cluster richness and redshift of the DES volume-limited cluster sample is shown in Fig. 2. The richness estimate λ is the sum over the membership probabilities of all

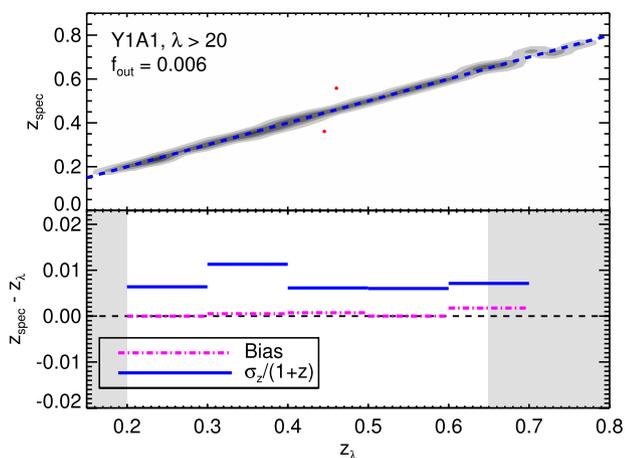


Figure 3. Photometric redshift performance of the DES Y1 redMaPPer cluster catalogue, as evaluated using available spectroscopy (333 clusters). *Upper panel:* Gray contours are 3σ confidence intervals, and the two red dots are the only 4σ outliers, caused by miscentring on a foreground/background galaxy. *Lower Panel:* photo- z bias and uncertainty. The comparatively large uncertainty from $0.3 < z < 0.4$ is due to a filter transition.

galaxies within a pre-defined, richness-dependent projected radius R_λ . The radius R_λ is related to the cluster richness via $R_\lambda = 1.0(\lambda/100)^{0.2} h^{-1} \text{Mpc}$. This relation was found to minimize the scatter between richness and X-ray luminosity in Rykoff et al. (2012). A redshift estimate for each cluster is obtained by maximizing the probability that the observed colour-distribution of likely members matches the self-calibrated red-sequence model of redMaPPer.

Fig. 3 shows the photometric redshift performance of the DES Y1 volume-limited redMaPPer cluster sample. The photometric redshift bias and scatter are calculated by comparing the photometric redshift of the clusters to the spectroscopic redshift of the central galaxy of the cluster, where available. Unfortunately, the small overlap with existing spectroscopic surveys means that our results are limited by small-number statistics: there are only 333 galaxy clusters with a spectroscopic central galaxy, and only 34 (six) with redshift $z \geq 0.6$ ($z \geq 0.65$). Nevertheless, the photometric redshift performance is consistent with our expectations: our redshifts are very nearly unbiased, and have a remarkably tight scatter — the median value of $\sigma_z/(1+z)$ is ≈ 0.006 . An upper limit for the photometric redshift bias of 0.003 is consistent with our data.

Of particular importance to this work is the distribution of miscentred clusters — both the frequency and severity of their miscentring. Based on the redMaPPer centring probabilities, we would expect ≈ 80 percent of the clusters to be correctly centred, meaning the most likely redMaPPer central galaxy is at the centre of the potential well of the host halo. In practice, the fraction of correctly centred galaxy clusters is closer to ≈ 70 per cent, as estimated from a detailed comparison of the redMaPPer photometric centres to the X-ray centres of redMaPPer clusters for which high-resolution X-ray data is available (von der Linden et al., in preparation; Zhang et al., in preparation). The expected impact of this miscentring effect, and the detailed model for the miscentred distribution from Zhang et al. (in preparation) and von der Linden et al. (in preparation) is described in Section 5.2.

2.3 Shear catalogs

Our work uses the DES Y1 weak lensing galaxy shape catalogues presented in Zuntz et al. (2018). Two independent catalogues were

created: METACALIBRATION (Sheldon & Huff 2017; Huff & Mandelbaum 2017) based on NGMIX (Sheldon 2015), and IM3SHAPE (Zuntz et al. 2013). Both pass a multitude of tests for systematics, making them suitable for cosmological analyses. While the Y1 data is shallower than the DES SV data, improvements in the shear estimation pipelines and overall data quality enabled us to reach a number density of sources similar to that from DES SV data (Jarvis et al. 2016).

In this study, we will focus exclusively on the METACALIBRATION shear catalogue because of its larger effective source density (6.28 arcmin^{-2}) compared to the IM3SHAPE catalogue (3.71 arcmin^{-2}). The difference mainly arises because METACALIBRATION utilizes images taken in r, i, z bands, whereas IM3SHAPE relies exclusively on r -band data. In the METACALIBRATION shear catalogue, the fiducial shear estimates are obtained from a single Gaussian fit via the NGMIX algorithm. As a supplementary data product, METACALIBRATION provides (g, r, i, z) -band fluxes and the corresponding error estimates for objects using its internal model of the galaxies.

Galaxy shape estimators, such as the NGMIX model-fitting procedure used for METACALIBRATION, are subject to various sources of systematic errors. For a stacked shear analysis, the dominant problem is a multiplicative bias, i.e. an over- or underestimation of gravitational shear as inferred from the mean tangential ellipticity of lensed galaxies. This bias needs to be characterized and corrected. Traditionally, this is done using simulated galaxy images — with the critical limitation that simulations never fully resemble the observations.

The METACALIBRATION catalogue, in contrast, uses the galaxy images themselves to de-bias shear estimates. Specifically, each galaxy image is deconvolved from the estimated PSF, and a small positive and negative shear is applied to the deconvolved image in both the \hat{e}_1 and \hat{e}_2 directions. The resulting images are then convolved once again with a representation of the PSF, and an ellipticity \mathbf{e} is estimated for these new images (Zuntz et al. 2018). These new measurements can be used to directly estimate the response of the ellipticity measurement to a gravitational shear γ using finite difference derivatives:

$$R_\gamma = \frac{\partial \mathbf{e}}{\partial \gamma}. \quad (1)$$

Selection effects can also be accounted for by examining the response of the selections to shear. The application of a weight when calculating the mean shear over an ensemble is effectively a type of smooth selection, and is accounted for in the same way. We describe this effect with a selection response \mathbf{R}_{sel} , which leads to the response-corrected mean shear estimate

$$\langle \gamma \rangle \approx \langle \mathbf{R} \rangle^{-1} \langle \mathbf{R} \cdot \gamma_{\text{true}} \rangle \approx \langle \mathbf{R} \rangle^{-1} \langle \mathbf{e} \rangle \quad (2)$$

from biased measurements \mathbf{e} with a joint response $\mathbf{R} \approx \mathbf{R}_\gamma + \mathbf{R}_{\text{sel}}$ (Sheldon & Huff 2017). Here, the left-hand side represents our estimate of the mean shear, while γ_{true} refers to the actual value.

\mathbf{R} is a 2×2 Jacobian matrix for the two ellipticity components e_1, e_2 in a celestial coordinate system. For the METACALIBRATION mean shear measurements in this work, we calculate the response of mean tangential shear on mean tangential ellipticity. \mathbf{R} is close to isotropic on average, which is why other recent weak lensing analyses (Prat et al. 2018; Troxel et al. 2018; Chang et al. 2018; Gruen et al. 2018) have assumed it to be a scalar. For the larger tangential shears measured on small scales around clusters, however, we account for the fact that the response might not be quite isotropic by explicitly rotating it to the tangential frame.

The tangential ellipticity e_T is related to e_1 , e_2 (and likewise γ_T to γ_1 and γ_2) by

$$e_T = -e_1 \cos(2\phi) - e_2 \sin(2\phi), \quad (3)$$

where ϕ is the polar angle of the source in a coordinate system centred on the lens. For the shear response, the corresponding rotation is derived from equation (1) and equation (3) as

$$\begin{aligned} R_{\gamma,T} &= R_{\gamma,11} \cos^2(2\phi) + R_{\gamma,22} \sin^2(2\phi) \\ &+ (R_{\gamma,12} + R_{\gamma,21}) \sin(2\phi) \cos(2\phi). \end{aligned} \quad (4)$$

For the METACALIBRATION selection response, no such rotation can be performed as the term itself is only meaningful for ensembles of galaxies. In this case, we exploit that the orientation of source galaxies should be random relative to the clusters, which suggest a symmetrized version of the response in the tangential frame:

$$\langle R_{\text{sel}}^{(T)} \rangle \approx \frac{1}{2} \text{Tr} \langle R_{\text{sel}} \rangle \quad \text{where} \quad \langle R_{\text{sel}} \rangle_{i,j} \approx \frac{\langle e_i \rangle^{S^+} - \langle e_i \rangle^{S^-}}{\Delta\gamma_j}. \quad (5)$$

In the above equation $\langle e_i \rangle^{S^\pm}$ denotes the mean *un-sheared* ellipticity of galaxies when selected based on quantities measured on their artificially *sheared* images. Four such sheared images are created by applying positive (+) and negative (−) shears of magnitude $\Delta\gamma_j = 0.01$ along the $j \in \{\hat{e}_1; \hat{e}_2\}$ directions separately. The response in the tangential and cross directions are consistent; however, both depend on the cluster-centric distance and richness of the lensing clusters. These dependencies will be investigated in a future work. Errors introduced from this approximation are sub-dominant due to the already small bias associated with source galaxy selection. A detailed discussion of additional possible systematics in our specific analysis is presented in Section 4.1.

Blinding procedure

As a precaution against unintentional confirmation bias in the scientific analyses, both weak lensing shape catalogues produced for DES Y1 had an unknown *blinding* factor in the magnitude of \mathbf{e} (Zuntz et al. 2018) applied to them. This unknown factor was constrained between 0.9 and 1.1. While we made initial blinded measurements for this work, the factor was revealed as part of unblinding the cosmology results of DES Collaboration et al. (2018).

In accordance with the practices of other DES Y1 cosmology analyses, we have further adopted a secondary layer of blinding. Specifically, we blindly transform the chains from our MCMCs to hide our in-progress results, and to prevent comparison between our cluster masses and those estimated using mass–observable relations from the literature. Chains of the parameters in the modelled lensing profiles and the mass–richness relation were unaltered after unblinding.

2.4 Photometric redshift catalogue

In interpreting the weak gravitational lensing signal of galaxy clusters as physical mass profiles, we need to employ information about the geometry of the source–lens systems by considering the relevant angular–diameter distances. To calculate these distances, we rely on estimates of the overall redshift distribution of source galaxies, and also on information about the individual $P(z)$ of source galaxies.

We use the DES Y1 photometric redshifts estimated and validated by Hoyle et al. (2018) using the template-based BPZ algorithm (Benítez 2000; Coe et al. 2006). It was found by Hoyle et al. (2018) that these photo- z estimates were modestly biased, introducing an

overall multiplicative systematic correction in the recovered weak lensing profiles. We determine this correction and its systematic uncertainty in Section 4.3.

In order to be able to correct selection effects due to the change of photo- z with shear while utilizing the highest signal-to-noise flux measurements for determining the source redshift distribution, we use two separate BPZ catalogues: one generated from METACALIBRATION-measured photometry (for selecting and weighting sources), and one from MOF (see Section 2.1) photometry (for determining the resulting source redshift distributions). Details of this are described in the following section.

3 STACKED LENSING MEASUREMENTS

3.1 Mass density profiles

Gravitational lensing induces distortions in the images of background ‘source’ galaxies. In the limit of weak gravitational lensing, these are characterized by the ‘reduced shear’

$$\mathbf{g} \equiv \frac{\gamma}{1 - \kappa}, \quad (6)$$

where γ is the shear and κ is the convergence. In the presence of non-negligible convergence, the ellipticity estimator \mathbf{e} introduced in Section 2.3 relates to the reduced shear as $\langle \mathbf{g} \rangle \approx \langle \mathbf{R} \rangle^{-1} \langle \mathbf{e} \rangle$.

The gravity of a localized mass distribution, such as a galaxy cluster, induces positive shear along the tangential direction with respect to the centre of the overdensity. This net tangential shear results in the stretching and preferential alignment of the images of background galaxies along the tangential direction. The magnitude of the azimuthally averaged tangential shear γ_T at projected radius R can be predicted from the line-of-sight projected surface mass density Σ of the lens mass distribution by the relation

$$\gamma_T = \frac{\overline{\Sigma}(< R) - \overline{\Sigma}(R)}{\Sigma_{\text{crit}}} \equiv \frac{\Delta\Sigma(R)}{\Sigma_{\text{crit}}}. \quad (7)$$

Here, $\overline{\Sigma}(< R)$ represents the average surface mass density within projected radius R , and $\overline{\Sigma}(R)$ represents the (azimuthal) average of the surface mass density at R . For the case of reduced shear, this equation holds only in linear order; therefore, we account for the effect of κ in our model described in Section 5.3.2.

The geometry of the source–lens system modulates the amplitude of the induced shear signal, and is characterized by the critical surface mass density

$$\Sigma_{\text{crit}}(z_s, z_l) = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}} \quad (8)$$

in equation (7). Here, D_s , D_l and D_{ls} are the angular diameter distances to the source, to the lens, and between the lens and the source. Estimating the $\Delta\Sigma$ signal thus relies on robustly estimating the redshifts of the galaxy clusters and the source galaxies. The lens redshifts are the photometric redshift estimates from the redMaPPer algorithm. The statistical uncertainty on these estimates is found to be $\Delta z_l \approx 0.01$ (Rykoff et al. 2016), which is negligible compared to other sources of error in the lensing measurement, allowing us to treat these redshifts as point estimates.

Source redshifts are also estimated from photometry, and are described by a probability distribution $p_{\text{phot}}(z_s)$ for each source galaxy. We can therefore only estimate an *effective* critical surface density

$$\langle \Sigma_{\text{crit}}^{-1} \rangle_{i,j} = \int dz_s p_{\text{phot}}(z_{s,i}) \Sigma_{\text{crit}}^{-1}(z_{l,j}, z_{s,i}), \quad (9)$$

where i and j index the source and the lens in a lens–source pair. Note that here we choose to express the *inverse* critical surface density, which is the predicted amplitude of the lensing signal in equation (7). We consistently define it as zero if $z_s \leq z_l$. For reasons of data compression, we will in fact *not* use the full integral over $p_{\text{phot}}(z)$ later, but rather replace equation (9) by $\Sigma_{\text{crit}}^{-1}$ evaluated at a random sample of the $p_{\text{phot}}(z)$.

3.1.1 The lensing estimator

Due to the low signal-to-noise of individual source–lens pairs, we measure the stacked (mean) signal of many source galaxies around a selection of clusters.

Sheldon et al. (2004) show that the minimum variance estimator for the weak lensing signal is

$$\widetilde{\Delta\Sigma} = \frac{\sum_j \sum_i^{\text{lens src}} w_{i,j} e_{T,i,j} / \langle \Sigma_{\text{crit}}^{-1} \rangle_{i,j}}{\sum_{j,i} w_{i,j}}, \quad (10)$$

where the summation goes over all source–lens pairs in some radius bin and $e_{T,i,j}$ is the tangential component of the ellipticity of source i relative to lens j . The optimal weights, proportional to the inverse variance of $e_{T,i,j} / \langle \Sigma_{\text{crit}}^{-1} \rangle$, are

$$w_{i,j} = \langle \Sigma_{\text{crit}}^{-1} \rangle_{i,j}^2 / \sigma_{\gamma,i}^2, \quad (11)$$

where $\sigma_{\gamma,i}^2$ is the estimate on the variance of the measured shear estimate of galaxy i relating to both the intrinsic variance of shapes and also to the uncertainty originating from shear estimation.

3.1.2 Practical lensing estimator

This estimator can be equivalently understood as a mean tangential ellipticity, weighted by the expected shear signal amplitude of each galaxy $\langle \Sigma_{\text{crit}}^{-1} \rangle$. It is normalized by the expected signal per unit $\Delta\Sigma$, i.e. the $\langle \Sigma_{\text{crit}}^{-1} \rangle$ -weighted mean of the $\langle \Sigma_{\text{crit}}^{-1} \rangle$. With this in mind, and including shear and selection response (see Section 2.3), we define the estimator we use in practice as

$$\widetilde{\Delta\Sigma} \equiv \frac{\sum_{j,i} \omega_{i,j} e_{T,i,j}}{\sum_{j,i} \omega_{i,j} \Sigma_{\text{crit},i,j}'^{-1} \mathbf{R}_{\gamma,i}^T + \left(\sum_{j,i} \omega_{i,j} \Sigma_{\text{crit},i,j}'^{-1} \right) \langle \mathbf{R}_{\text{sel}}^T \rangle}. \quad (12)$$

In the above, $\langle \mathbf{R}_{\text{sel}}^T \rangle$ is calculated via equation (5) separately for source galaxies selected in each radial bin and each richness–redshift bin, where the corresponding selections were defined by the photometric redshift estimates derived from the *sheared* METACALIBRATION photometries. The small number of source galaxies at small radii introduces some noise to the estimated response, however, due to the intrinsic environmental dependence of $\mathbf{R}_{\text{sel}}^T$, this cannot be readily substituted or approximated with other, less noisy quantities. By considering the expectation value

$$\langle e_{T,i,j} \rangle = \Delta\Sigma \Sigma_{\text{crit},i,j}'^{-1} \mathbf{R}_i^T, \quad (13)$$

it is easy to see that the definition of equation (12) yields an unbiased estimate of $\Delta\Sigma$.

Equation (12) includes two simplifications to make calculations less computationally demanding. First, for the normalization, we replace the expectation value of $\Sigma_{\text{crit}}^{-1}$ by a Monte Carlo estimate

$$\Sigma_{\text{crit},i,j}'^{-1} = \Sigma_{\text{crit}}(z_{l_j}, z_{s_i}^{\text{MC}}), \quad (14)$$

where $z_{s_i}^{\text{MC}}$ is a random sample from the $p_{\text{phot}}(z_s)$ distribution estimated with BPZ using MOF photometry. Second, the weights are chosen as

$$\omega_{i,j} \equiv \Sigma_{\text{crit}}^{-1}(z_{l_j}, \langle z_{s_i}^{\text{MCAL}} \rangle) \text{ if } \langle z_{s_i}^{\text{MCAL}} \rangle > z_{l_j} + \Delta z, \quad (15)$$

with $\langle z_s^{\text{MCAL}} \rangle$ being the mean redshift of the source galaxy estimated from METACALIBRATION photometry. Given the width of our photometrically estimated $p(z)$, this is close to the optimal weight. We use a padding of $\Delta z = 0.1$ for source selection. We found that including the source weights provided by METACALIBRATION does not introduce a significant improvement in the signal-to-noise of the measurement.

The use of two different photometric estimators is necessary because when calculating the selection response, the internal photometry of the METACALIBRATION, with measurements on sheared images, must be used for all selection and weighting of sources. Hoyle et al. (2018) find this photometric redshift estimate to have a greater scatter than the default MOF photometry. We therefore opt to use the METACALIBRATION photo- z estimates only for *selecting and weighting* source–lens pairs. When normalizing the shear signal to find $\Delta\Sigma$, we utilize the MOF-based photo- z estimates.

3.1.3 Data vector binned in redshift and richness

In estimating the lensing signal through equation (12), we utilize a modified version of the publicly available XSHEAR code¹ and the custom-built XPIPE python package.² The core implementation of the measurement code is identical to the one used by Melchior et al. (2017).

We group the clusters into three bins in redshift: $z \in [0.2; 0.4)$, $[0.4; 0.5)$, and $[0.5; 0.65)$, as well as seven bins in richness: $\lambda \in [5; 10)$, $[10; 14)$, $[14; 20)$, $[20; 30)$, $[30; 45)$, $[45; 60)$, and $[60; \infty)$. The redshift limit $z = 0.65$ of our highest redshift corresponds roughly to the highest redshift for which the redMaPPer cluster catalogue remains volume-limited across the full DES Y1 survey footprint. The $\Delta\Sigma$ profiles were measured in 15 logarithmically spaced radial bins ranging from 0.03 Mpc to 30 Mpc. For our later results, we will only utilize the radial range above 200 kpc. Scales below this cut are included only in our figures and for reference purposes, and are excluded from the analysis to avoid systematic effects such as obscuration, significant membership contamination, and blending. This radial binning scheme yields similar S/N across all bins. The measured shear profiles are shown in Fig. 4.

We find a mild radial dependence in the typical value for METACALIBRATION shear response $\langle \mathbf{R}_{\gamma,T} \rangle$, the asymptotic values are 0.6, 0.58, and 0.55 as a function of increasing cluster redshift. For the selection response, we find an asymptotic value of $\langle \mathbf{R}_{\text{sel}} \rangle \approx 0.013$, 0.014, and 0.015.

3.2 Covariance matrices

The $\Delta\Sigma$ profiles estimated in the previous section deviate from the true signal due to statistical uncertainties and systematic biases. We construct a description for the covariance of our data vector below and calibrate the influence of systematic effects in Section 4.

Statistical uncertainties originate from the large intrinsic scatter in the shapes of source galaxies, the uncertainty in estimating

¹<https://github.com/esheldon/xshear>

²<https://github.com/vargatn/xpipe>

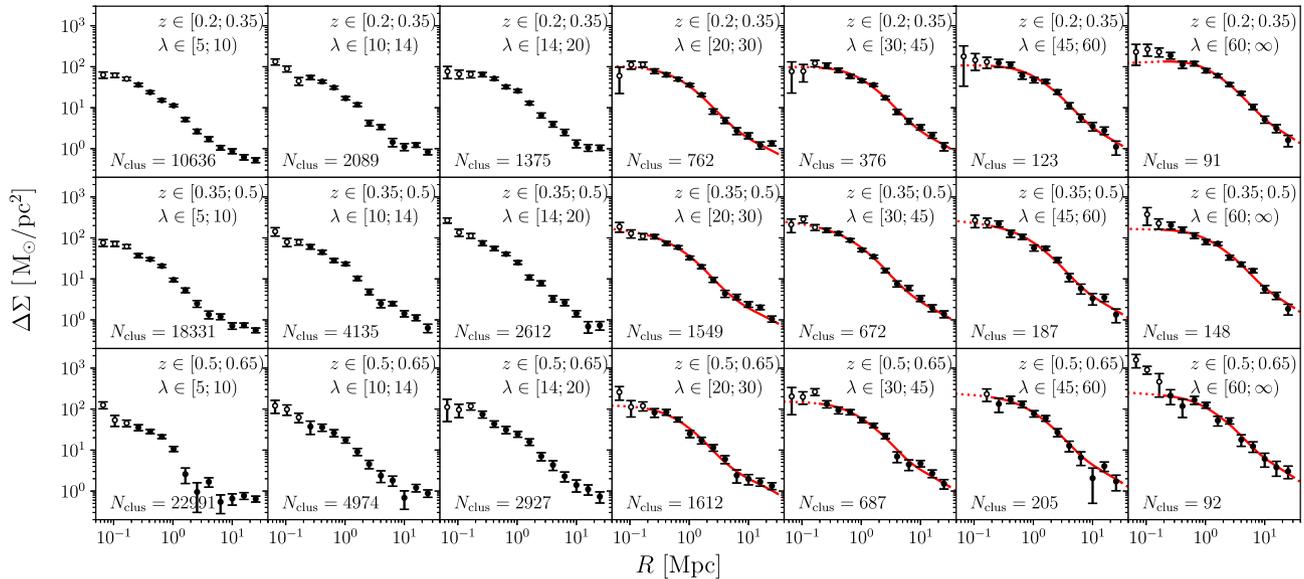


Figure 4. Mean $\Delta\Sigma$ for cluster subsets split in redshift z_1 (increasing from top to bottom) and λ (increasing from left to right), as labelled. The error bars shown are the diagonal entries of our semi-analytic covariance matrix estimate (see Section 3.2) for bins with $\lambda > 20$ and the jackknife-estimated covariance matrix for bins with $\lambda < 20$. The best-fit model (red curve) is shown for bins with $\lambda > 20$, and includes dilution from cluster member galaxies (Section 5.3.1) and miscentring (Section 5.2); see Section 5 for details. Semi-analytic covariances were not computed for stacks with $\lambda < 20$ due to the significant computational cost. Below 200 kpc, we consider data points unreliable and therefore exclude them from our analysis; these are indicated by open symbols and dashed lines. The profiles and jackknife errors are calculated after the subtraction of the random-point shear signal (see Section 4.1.3).

their photometric redshifts, and due to the intrinsic variations in the properties and environments of galaxy clusters. Furthermore, our typical maximum radii: 2, 1.5, and 1.3 degrees for the different redshift ranges, respectively, are much larger than the 0.22 degree median separation between clusters in the catalogue. This means that source galaxies are paired with multiple clusters, possibly generating covariance between different radial ranges and/or across different cluster bins in richness and redshift.

To quantify the correlation and uncertainty involved in the measurement, we construct a semi-analytic model for the data covariance matrix following the framework developed by Gruen et al. (2015). Our use of a semi-analytic covariance (SAC) matrix is motivated by explicit covariance estimators exhibiting non-negligible uncertainty and possible biases, for instance from jackknife regions that are not completely independent. Both of these problems lead to a biased estimate of the precision matrix (i.e. the inverse covariance matrix), which, in turn, will bias the posteriors of likelihood inference (Friedrich et al. 2016).

Instead, we *predict* several key contributions of the observed covariances, namely those due to correlated and uncorrelated large-scale structure, stochasticity in cluster centring, the intrinsic scatter in cluster concentrations at fixed mass, cluster ellipticity, and the scatter in the richness–mass relation of galaxy clusters. Only the shape noise contribution is estimated directly from the data, as detailed below.

While we rely on the SAC matrix estimates in the remainder of our analysis, we compare the SAC matrices to those derived using a standard jackknife method. We use jackknife (JK) resampling with $K = 100$ simply-connected spatial regions \mathcal{R}_k selected via a *k-means* algorithm on the sphere.³ The jackknife covariance is

defined following Efron (1982):

$$C_{\Delta\Sigma} = \frac{K-1}{K} \sum_k \left(\widetilde{\Delta\Sigma}_{(k)} - \widetilde{\Delta\Sigma}_{(\cdot)} \right)^T \cdot \left(\widetilde{\Delta\Sigma}_{(k)} - \widetilde{\Delta\Sigma}_{(\cdot)} \right), \quad (16)$$

where $\widetilde{\Delta\Sigma}_{(\cdot)} = \frac{1}{K} \sum_k \widetilde{\Delta\Sigma}_{(k)}$ and $\widetilde{\Delta\Sigma}_{(k)}$ denotes the lensing signal estimated via equation (12) using all lenses except those in region \mathcal{R}_k . Using this method, we calculate the covariance between all radial bins in a single richness and redshift bin, as well as the covariance between adjacent richness and redshift bins.

Fig. 5 shows an example of the structure of the jackknife-estimated correlation matrix between neighbouring bins in richness and redshift. We find no significant correlation between richness/redshift bin and therefore treat each bin independently, even though some systematic parameters may be shared between bins.

3.2.1 Shape noise

The large intrinsic variations of the shapes of galaxies (*shape noise*) in the source catalogue constitute a dominant source of uncertainty in lensing measurements. We estimate the covariance originating from both the random intrinsic alignments and also the stochastic positions of source galaxies. In order to do so, we make use of the measurement setup outlined in Section 3.1.3, but each source is randomly *rotated* to create a new source catalogue. We generated 1000 such independent *rotated* source catalogues, and performed the lensing measurement with each. The resulting data vectors are consistent with zero, as the random rotation washes away the imprint of the weak lensing signal. However, their scatter is indicative of the covariance due to shape noise.

We estimate the shape noise covariance matrix for each of the 1000 realizations using the spatial jackknife scheme outlined above in Section 3.2. The final shape noise covariance matrix estimate is

³https://github.com/esheldon/kmeans_radc

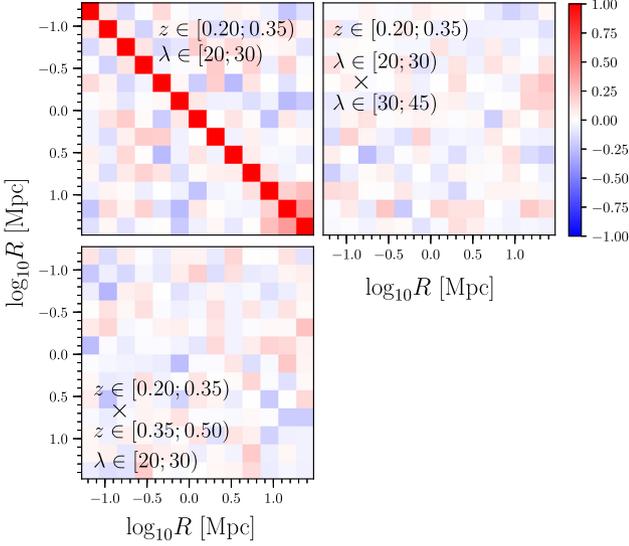


Figure 5. Jackknife-estimated correlation matrix of $\widetilde{\Delta\Sigma}$ of a single richness-redshift selection with $\lambda \in [20; 30]$ and $z \in [0.2; 0.35]$ (upper left panel). The off-diagonal blocks display the correlation matrix between the reference profile and the neighbouring richness bin $\lambda \in [30; 45]$ (upper right panel), and the neighbouring redshift bin $z \in [0.35; 0.5]$ (lower left panel).

obtained by averaging all 1000 of these jackknife covariance matrices. We expect this method to be less noisy compared to estimating the covariance matrix from the 1000 independent measurements of the rotated $\Delta\Sigma$ vector only.

3.2.2 Uncorrelated LSS

Line of sight structures which are not physically connected to the cluster leave an impact on the lensing signal. We cannot remove them from the signal, but we can estimate their expected contribution to the covariance. For an individual cluster, the covariance of the $\Delta\Sigma$ profile at radii θ_i, θ_j due to uncorrelated large-scale structure (uLSS) can be written as (e.g. Schneider et al. 1998; Hoekstra 2003; Umetsu et al. 2011)

$$C_{ij}^{\text{uLSS}} = \langle \Sigma_{\text{crit}}^{-1} \rangle^{-2} \int \frac{\ell d\ell}{2\pi} P_\kappa(\ell) J_2(\ell\theta_i) J_2(\ell\theta_j), \quad (17)$$

where P_κ is the power spectrum of the convergence, and J_2 is the Bessel function of the first kind of order 2.

Naively, one would expect that the variance of a cluster stack due to uncorrelated large-scale structure to scale simply as $1/N_{\text{clusters}}$. In practice, however, the positions of galaxy clusters are correlated, and the area around them overlaps on large scales. Consequently, we expect the variance due to uncorrelated structure to decrease somewhat more slowly than $1/N_{\text{clusters}}$.

We estimate this source of noise by measuring random realizations of the signal due to shear fields induced by lognormal density fields with the appropriate power spectra and skewness. We calculate the latter with the perturbation theory model of Friedrich et al. (2018) for the Buzzard cosmology (DeRose et al., in preparation; Wechsler, DeRose & Busha, in preparation), using the lognormal parameter κ_0 at a $10'$ aperture radius. As our cluster sample spans a range in redshift, a different shear field is calculated for each of the three redshift bins. This is done such that the shear fields are calculated at the lens-weighted mean source galaxy redshifts found during the initial measurement in Section 3.1.3. We then pass these shear fields through the measurement pipeline using a spatial mask

reflecting the actual source number density variations across the footprint, and estimate the covariance matrix for each realization using 100 spatial jackknife regions for each bin in richness and redshift.

This above procedure was repeated 300 times, and the final covariance matrix due to uncorrelated LSS is taken to be the mean of the 300 jackknife covariance estimates.

3.2.3 Correlated LSS and halo ellipticity

Following Gruen et al. (2015), we model correlated large-scale structure using a halo model approach. We assume correlated haloes can be adequately described using only two parameters: the mass M of the correlated halo, and the projected distance R_h from the cluster. The mass distribution of the haloes is assumed to follow the halo mass function, while their spatial distribution is modelled as a Poisson realization of the density field defined by the appropriate halo-cluster correlation function. That is, the excess density of haloes of mass M a distance R from the halo is

$$\rho_h(M, R_h) = b(M_{\text{cl}})b(M)w_{\text{mm}}(R_h)\frac{dn}{dM}, \quad (18)$$

where w_{mm} is the projected linear correlation function at the redshift of the cluster, $b(M)$ is the halo bias, and dn/dM is the halo mass function, or the number of haloes per unit volume per unit mass (Tinker et al. 2008).

Given a model for the halo profile $\Sigma(R|M)$, the contribution of a halo at location R_h to the mean surface density Σ of the cluster in radial bin R_i is $\Sigma_i(M, R_h) = \Sigma_{\text{misc}}(R_i|M, R_h)$, where Σ_{misc} is a miscentred halo profile. Because of the Poisson-sampling assumption, the covariance matrix is generated by the mass profiles of individual haloes, so that the correlated large-scale structure contribution to the covariance matrix can be written as

$$C_{ij}^{\text{cLSS}} = \int (2\pi R_h dR_h) dM \rho_h(M, R_h) \Delta\Sigma_i(M, R_h) \Delta\Sigma_j(M, R_h). \quad (19)$$

In practice, the above predicted covariance matrix is further rescaled by a constant factor calibrated on simulations. This is meant to account for additional variance not captured by linear bias and Poisson noise, due to filamentary structure and higher order statistics in the spatial distribution of the correlated haloes (e.g. the non-zero three-point function). A more detailed derivation of the above equation and its calibration is found in Gruen et al. (2015).

A very similar calculation can be made for characterizing the contribution due to halo ellipticity (to the covariance matrix – for the effect on the mean signal, see Section 5.4.2). If $\rho_{\text{ell}}(q, \mu)$ is the distribution of the halo axial ratio q and the line-of-sight orientation angle θ relative to the major axis such that $\cos \theta = \mu$, then one finds (Gruen et al. 2015)

$$C_{ij}^{\text{ell}} = \int dq d\mu \rho_{\text{ell}}(q, \mu) \Delta\Sigma_i \Delta\Sigma_j, \quad (20)$$

where $\Delta\Sigma_i$ is the contribution to the bin R_i under the assumption that the halo has an axial ratio q and an orientation μ .

3.2.4 M-c scatter; M-λ scatter and miscentring

Halos at a given mass have some intrinsic scatter in their M - λ relation. A rough estimate of the intrinsic scatter in the mass–richness (M - λ) relation is ~ 25 per cent (Rozo & Rykoff 2014; Farahi et al., in preparation), and it causes an increase in the variance of stacked

measurements of $\Delta\Sigma$. This scatter causes an even larger increase in the variance, since it propagates into quantities that depend directly on the mass, including the M - c relation. In addition, concentration (e.g. Bhattacharya et al. 2013; Diemer & Kravtsov 2015) and miscentring possess some intrinsic scatter from halo to halo themselves.

Scatter in the M - λ relation causes variance on all scales, since the bias $b(M)$ directly depends on the mass. By comparison, scatter in the M - c relation primarily affects small scales where the 1-halo term dominates. Similarly, some cluster centres are misidentified in our stacks, which creates additional covariance at small scales where the signal is substantially suppressed.

We modelled the combined contribution to the SAC from scatter in M - λ , scatter in concentration at fixed mass, and miscentring of individual clusters in our stacks by doing the following:

(i) For each cluster in our stack, assign a mass by inverting a fiducial M - λ relation (Melchior et al. 2017) and assuming 25 per cent scatter. This is not identical to 25 per cent scatter in the M - λ relation; however, this choice negligibly affects this component of the covariance matrix.

(ii) For each cluster, assign a concentration (including scatter) based on Diemer & Kravtsov (2015).

(iii) For each cluster, make a draw from our centring prior described in Section 5.2. In other words, some fraction f_{mis} of clusters in the stack are miscentred, and the distribution of the amount of miscentring is given by $p(R_{\text{mis}})$.

(iv) Calculate $\Delta\Sigma$ for each cluster and average these signals to generate a signal for the entire stack.

(v) Repeat this process many times, and use these independent realizations to estimate the corresponding covariance matrix between the various radial bins.

Using Simet et al. (2017) as our fiducial M - λ relation or using the Bhattacharya et al. (2013) mass–concentration relation had no impact on the final SAC matrix. We have also verified that using half as many realizations as our fiducial choice (1000) did not appreciably change the resulting covariance matrix. The same is true for changes in the richness scatter or miscentring model parameters within reasonable ranges.

3.2.5 Semi-analytic covariance matrix

Following Gruen et al. (2015), the full SAC matrix is obtained by adding each of the above contributions. The individual components described in the previous subsections are shown in Fig. 6. Fig. 7 demonstrates the differences between the SAC and jackknife covariance matrices. The top two panels show the correlation matrix R of the SAC and C^{JK} , respectively, where the correlation matrix is defined via

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \quad (21)$$

Similarly, to visualize the difference between C^{SA} and $C^{\text{JK,Z}}$, we define the residual matrix

$$Q_{ij} = \frac{C_{ij}^{\text{SA}} - C_{ij}^{\text{JK}}}{\sqrt{C_{ii}^{\text{SA}}C_{jj}^{\text{SA}}}}. \quad (22)$$

We show this residual matrix in the bottom right-hand panel of Fig. 7. Finally, in the lower left-hand panel, we show the difference between the errors along the diagonal between the C^{JK} and the SAC, along with each of the contributions to the SAC; the lower panel shows the fractional difference between the diagonal entries.

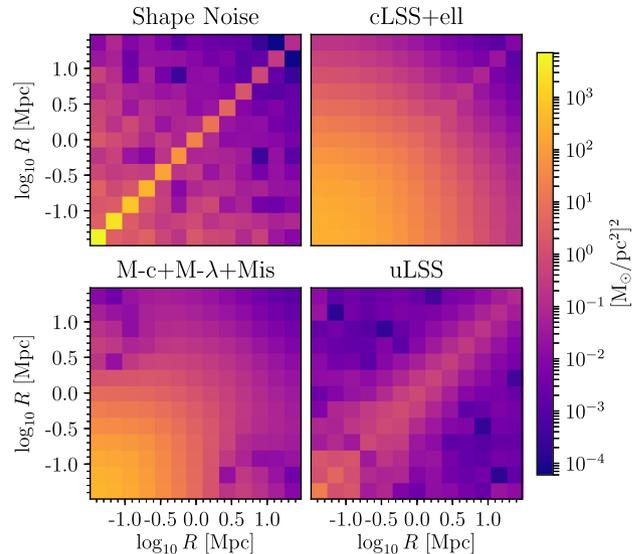


Figure 6. The four individual components to our semi-analytic covariance matrix (SAC). *Clockwise from top left:* Shape noise component from randomly rotating sources, correlated LSS and ellipticity component from integrating over configurations of the host cluster and its correlated halos, uncorrelated LSS from integrating over large-scale structure, and finally scatter in the M - λ relation, M - c relation, and miscentring distribution. Dark colours correspond to low covariance and the colours are log-scaled to show trends. Light colours are normalized to the total covariance in the SAC. See Section 3.2 for details.

As expected, shape noise is the dominant contributor to the SAC matrix, with uncorrelated LSS becoming important at the largest scales. This explains why the choices we had to make in modeling the non-shape noise components did not significantly affect the resulting SAC matrix or the posteriors analysis.

Using the SACs in our analysis provides two major improvements: minimal bias from inverting the covariance matrix, and less overall noise in the off-diagonal elements which improves the mass measurement. In Melchior et al. (2017), we demonstrated that noise in the jackknife covariance matrix led to an increase of ≈ 30 per cent in the uncertainty of the mass of the stack. Using the SACs reduces the contribution of the covariance to the error budget by 10 per cent compared to the jackknife-estimated covariance.

4 SYSTEMATICS

4.1 Shear systematics

The METACALIBRATION shear catalogue and the associated calibration of the source redshift distributions (Hoyle et al. 2018) passed a large number of tests performed by Zuntz et al. (2018) and Prat et al. (2018). Here, we briefly enumerate the constraints on the most relevant systematics, and refer the reader to the corresponding papers for a more detailed analysis.

We parametrize the various potential biases in the data set as:

$$g_i = (1 + m_i)g_i^{\text{tr}} + \alpha^{\text{PSF}}e_i^{\text{PSF}} + c_i, \quad (23)$$

where g_i^{tr} is the true shear, while g_i is the shear estimate, and α^{PSF} relates to the contamination from the PSF ellipticity e_i^{PSF} .

In weak lensing surveys, the three main sources of bias are commonly found to be model bias, noise bias, and selection bias (or representativeness bias). In order to account and correct for these sources of error, the METACALIBRATION algorithm performs a self-

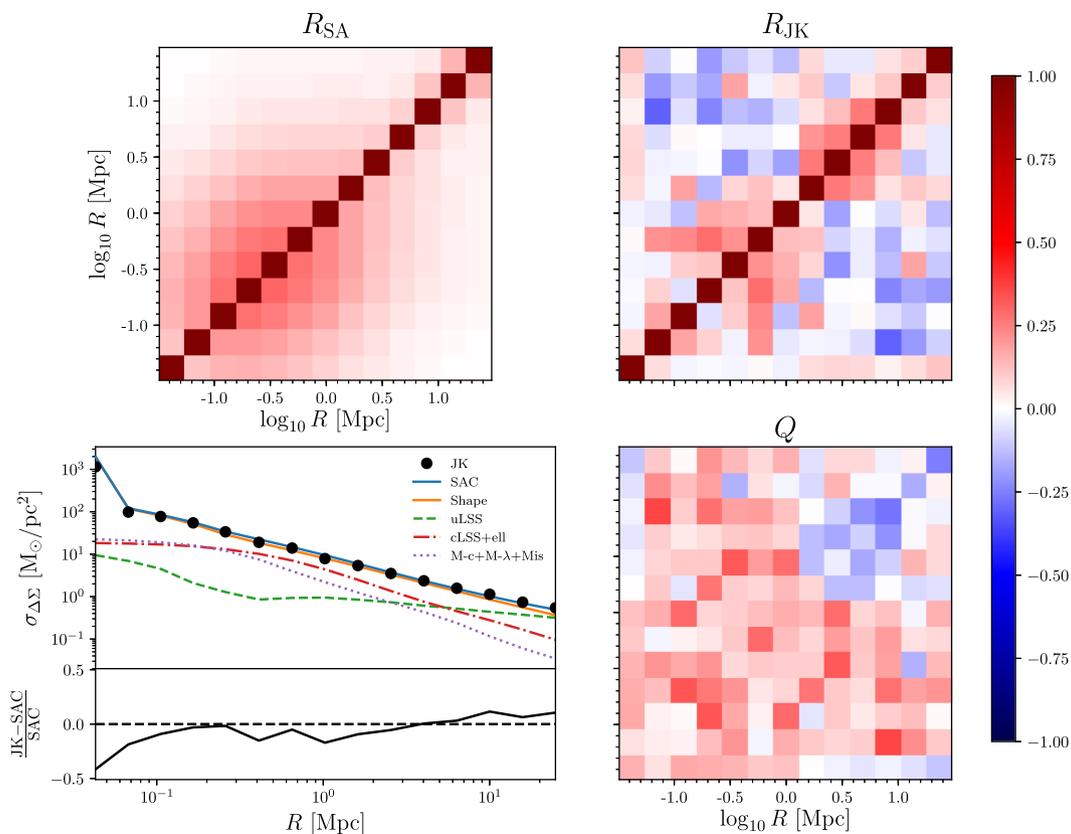


Figure 7. Comparison between the semi-analytic covariance matrix and the jackknife-estimated covariance matrix. *Top left:* Correlation matrix of the SAC matrix. *Top-right:* Correlation matrix of the jackknife estimate. *Bottom left:* Comparison of the errors in the SAC and jackknife estimate along with the contributions to the SAC error from each individual component. The line showing the SAC errors lies almost on top of the shape noise contribution, confirming that it is the dominant source of covariance. *Bottom right:* Residual matrix Q (see equation 22) that represents the difference between the SAC and jackknife covariance matrices. See Section 3.2 for details.

calibration on the actual data by shearing the galaxy images during the measurement, and using the thus-calculated responses to correct the shear estimates. To quantify the effectiveness of this self-calibration, Zuntz et al. (2018) ran the METACALIBRATION pipeline on a set of simulated galaxy images using GALSIM (Rowe et al. 2015). The images were produced from high-resolution galaxy images from the COSMOS sample, and processed to resemble the actual DES Y1 observations both in noise and PSF properties. Based on this test scenario, Zuntz et al. (2018) found no significant multiplicative bias m or additive bias c present in the data set.

Zuntz et al. (2018) further investigated the multiplicative biases due to blending of galaxy images, due to the potential leakage of stellar objects into the galaxy sample, and due to potential errors in the modelling of the PSF. They found blending as the only component with a net bias, with the other sources being consistent with zero, although contributing to the uncertainty on the value of m . The final multiplicative bias estimates were found to be $m = [1.2 \pm 1.3] \cdot 10^{-2}$ with a 1σ Gaussian error. They found no evidence of a significant additive bias term.

Prat et al. (2018) tested for the presence of residual shear calibration biases in the DES Y1 galaxy-galaxy lensing analysis by splitting the source sample by various galaxy properties and parameters of the observational data. They showed that within the statistical uncertainty of the respective galaxy-galaxy lensing signals, and including the differences in redshift distributions induced

by the splitting, no differential multiplicative biases between any of the splits were significantly detected.

In addition to the above calibrations during the construction of the shear catalogue, we perform additional sanity checks relevant to stacked weak lensing measurements in the subsections below.

4.1.1 Second order shear bias

Due to the larger tangential shear near massive clusters, this analysis is more strongly affected by non-linear shear response than previous DES Y1 lensing analyses (see the discussion in section 9 of Sheldon & Huff 2017). This response biases cluster masses higher than they would be otherwise. To test this effect, we modify the measured $\Delta\Sigma$ profiles by adding the leading non-linear shear bias term, at third order in $\gamma_t = \Delta\Sigma \times \Sigma_{\text{crit}}^{-1}$, as

$$\Delta\Sigma'_{\text{obs}} = \frac{\Delta\Sigma_{\text{obs}} \langle \Sigma_{\text{crit}}^{-1} \rangle - \alpha_{\text{NL}} \left(\Delta\Sigma_{\text{Model}} \langle \Sigma_{\text{crit}}^{-1} \rangle \right)^3}{\langle \Sigma_{\text{crit}}^{-1} \rangle}, \quad (24)$$

where $\Delta\Sigma_{\text{Model}}$ is the optimized model discussed in Section 5. For the amplitude of non-linear shear bias, we choose $\alpha_{\text{NL}} = 0.6$ (Sheldon & Huff 2017). We model the profile of the highest richness stack at $z \in [0.2, 0.35]$ where, for the source redshift distribution of DES Y1, this effect is strongest. The recovered mass changes by less than 1 per cent, demonstrating that our recovered mass–richness–redshift relation is robust to non-linear shear bias within our error budget.

The choice of $\alpha = 0.6$ in our test is motivated by the image simulations used in Sheldon & Huff (2017). Other simulations find a range of values of similar magnitude. Since the effect is smaller than the overall shear uncertainty, yet its calibration is uncertain, we choose not to implement a correction in our final model.

4.1.2 B-modes

Gravitational lensing due to localized mass distributions can only produce a net E-mode signal in the shear field, which corresponds to the tangential shear γ_t . This allows for a simple null test for the presence of systematics: any non-zero cross-shear (i.e. a non-zero B-mode) must be due to systematics. We compute the cross-shear by projecting the shears to the direction 45° from the tangential direction. We estimate the stacked B-mode signal for all richness and redshift bins, and calculate the corresponding χ^2 values using the jackknife estimate of the covariance matrix. We find $\chi^2/11 < 18/11$ for all richness bins with $\lambda > 20$, indicating that our measurement is consistent with no systematics at a $p > 0.1$ level.

4.1.3 Random point test

In spite of not being detected by Zuntz et al. (2018) and Section 4.1.2, additive shear systematics may be present in the data, which could manifest as net signals visible on all radial scales. In order to test for such potential systematics, we measure the lensing signal around a set of random points chosen by the redMaPPer algorithm (Rykoff et al. 2016). These points are selected via weighted random draws to mirror the distribution of DES Y1 redMaPPer clusters both in angular distribution, as well as in redshift and richness.

As additive systematics would affect the lensing profiles of galaxy clusters and random points the same way, the systematic effect can be calibrated out by subtracting the profile of random points from the profile of clusters. While we find no significant net signal around random points, we nevertheless apply this calibration, and subtract the signal of 10^5 random points from the $\widetilde{\Delta\Sigma}$ of each bin in richness and redshift. Thanks to the large number of random points used, this subtraction does not introduce significant noise to the measurement.

A motivation for subtracting the signal around random points from the measurement, regardless of the presence of systematics, is presented by Singh et al. (2016). They found that the random subtracted signal relates to the *matter over-density field* around the lenses, while the un-subtracted lensing signal traces the *matter density field*, which carries additional variance on large scales. Indeed, the precursor study of the present paper (Melchior et al. 2017) found a similar trend. We note that when constructing our SAC matrix we always apply the random point subtraction described above to ensure that our covariance matrix properly accounts for the reduced covariance that this estimator enables.

4.2 Correction for cluster members in the shear catalog

Due to uncertainty in photometric redshift estimates, foreground galaxies can be included in the source catalogue used in our lensing measurements. So long as the ensemble redshift distribution dn/dz of the sources is properly estimated, this is accounted for in our analysis. In the projected vicinity of galaxy clusters, there is however a systematic effect biasing the naive redshift estimates of galaxies: the presence of a large cluster member population and the associated large-scale matter overdensity localized at the cluster redshift. For rich clusters, these member galaxies could make

up a significant fraction of all detected galaxies in a particular line of sight. Consequently, due to intrinsic imperfections in the selection, some of these galaxies leak into the source catalog used in the weak lensing measurement. Cluster member galaxies are randomly aligned (Sifón et al. 2015), meaning their contamination results in a measured lensing signal which is biased low due to the dilution of actual source galaxies within the catalogue.

It is therefore important for weak lensing studies to characterize and correct this dilution when interpreting the measurements.⁴ There have been several approaches in the literature to correct for the net effect of cluster member contamination. For instance, Sheldon et al. (2004) estimated the correction factor from the transverse correlation of source galaxies around galaxy clusters, while Gruen et al. (2014) and Melchior et al. (2017) estimated the contamination rate based on the colour or photometric redshift $p(z)$ information of galaxies in different radial separations from the cluster. One can also make simple colour cuts (Medezinski et al. 2010, 2018a; Schrabback et al. 2018) or photo- z cuts (Applegate et al. 2014) on the source population to mitigate the contamination, or estimate its effect based on the increased galaxy number density around the lenses (Hoekstra et al. 2015; Simet et al. 2015; Dietrich et al. 2017).

In this study, we adopt the approach of our precursor analysis from the Science Verification data release of DES (Melchior et al. 2017), in which we make use of the estimated $p(z)$ of the source galaxy sample to calculate the cluster contamination fraction f_{cl} along with a corresponding covariance matrix $C_{f_{cl}}$ estimated via spatial jackknifing, and use this quantity to recover the contamination corrected lensing profile:

$$\widetilde{\Delta\Sigma}_{\text{corr}}(R) = \frac{\widetilde{\Delta\Sigma}(R)}{1 - f_{cl}(R)}. \quad (25)$$

Using this $p(z)$ decomposition approach is further motivated by the complexity of the shear selection function in our analysis, which limits our ability to measure the correlation function of source galaxies. A detailed description of this method, along with validation on simulated DES-like mock observations is presented in an accompanying paper (Varga et al., in preparation). The robustness of the method to the presence of intra-cluster light is confirmed in Gruen & Brimiouille (2017, their appendix A).

4.3 Photometric redshift systematics

The redshift distribution of our selected source galaxies was estimated using BPZ (Benítez 2000) in the implementation of Hoyle et al. (2018). In BPZ or similar photometric redshift estimation procedures, one assumes a variety of galaxy spectral energy distribution (SED) templates and priors for the relative abundance of galaxies as a function of luminosity and redshift. Any deviation from these assumptions in the DES source galaxy sample can cause biases in photometric redshift estimates which must be calibrated.

For the cosmology analyses of the lensing two-point functions (Dark Energy Survey Collaboration 2016; Troxel et al. 2018), this calibration was performed in two independent ways, and with consistent results: by the redshift distributions of samples of galaxies with high-quality 30-band photo- z s from COSMOS, matched to DES lensing source galaxies (Hoyle et al. 2018), and by the clustering of lensing source galaxies with redMaGiC (Roza et al. 2016)

⁴This correction is also referred to as a *boost factor* as the measured signal should be boosted to correct for the contamination.

galaxies as a function of the redshift of the latter (Davis et al. 2018; Gatti et al. 2018).

For this work, we adapt the COSMOS calibration of Hoyle et al. (2018) to estimate the bias of our $\Delta\Sigma$ measurements, and the uncertainty in that bias. To this end, we select and weight galaxies from COSMOS in the same manner as for our measurements of the cluster $\Delta\Sigma$ profiles.

Following Hoyle et al. (2018, their section 4.1), we randomly sample 200 000 galaxies in our data and match them to COSMOS galaxies according to their flux in each band and their intrinsic size. From this *COSMOS resampling*, we select and weight galaxies as per Sections 3.1.2 and 3.1.3. From the COSMOS 30-band, we calculate the weighted mean true $\Sigma'_{\text{crit,TRUE}}^{-1}$. From noisy MOF *griz* BPZ redshift distribution samples, we get a mean $\Sigma'_{\text{crit,MEAS}}^{-1}$ that relates the weighted mean tangential shear to the $\Delta\Sigma$ profile. As in the denominator of equation (12), we use a weight $\omega \times R$ for the means. Because the source selection, ω weight, and Σ'_{crit} depends on lens redshift, we repeat this exercise for the range of cluster redshifts sampled by our catalogue, $z_l = 0.2 \dots 0.65$. A bias in $\Sigma'_{\text{crit}}^{-1}$ translates directly into a multiplicative bias in $\Delta\Sigma$.

We estimate four sources of uncertainty in the calibration of photometric redshift distributions (see Hoyle et al. 2018, their sections 4.2–4.5): 1) an uncertainty due to cosmic variance from the relative scatter of average $\Sigma'_{\text{crit,TRUE}}^{-1}$ in the resampling of the 368 simulated COSMOS footprints, to which we add the (subdominant) statistical uncertainty due to the limited sample size from bootstrap resamplings in quadrature; 2) an uncertainty due to photometric zeropoint offsets from realizations of photometric zeropoint calibration offsets; 3) an uncertainty due to the morphology matching, which we estimate as half the difference between the estimated $\Sigma'_{\text{crit,TRUE}}^{-1}$ of the sample with size + flux matching and that obtained without the size matching; and 4) a systematic uncertainty of the matching algorithm by a comparison between the fiducial $\Sigma'_{\text{crit,TRUE}}^{-1}$ value and that of the aforementioned 368 resampled simulated COSMOS fields. Effects 1, 3, and 4 contribute to the systematic uncertainties with similar size, while effect 2 is smaller but not quite negligible.

We define our model for the bias as

$$\frac{\Sigma'_{\text{crit,MEAS}}^{-1}}{\Sigma'_{\text{crit,TRUE}}^{-1}} \equiv 1 + \delta, \quad (26)$$

with the mean value given from the COSMOS analysis, and an uncertainty due to the four effects mentioned above. This ratio depends on lens redshift through the selection/weighting of sources and the source redshift dependence of photo- z bias. It is plotted across the entire lens redshift range in Fig. 8. The red points show the ratio at the mean redshifts of the bins used in our analysis. This multiplicative factor is fully degenerate with shear systematics (see Section 4.1) and is assumed to be correlated across redshift bins. δ is incorporated in our analysis by a prior that varies between each stack. The variation between richness bins is small compared to the variation across cluster redshift bins.

$$\delta = \begin{cases} 0.009 \pm 0.021 & \text{for } z \in [0.2, 0.35] \\ 0.002 \pm 0.020 & \text{for } z \in [0.35, 0.5] \\ 0.004 \pm 0.022 & \text{for } z \in [0.5, 0.65]. \end{cases} \quad (27)$$

This prior is combined with the prior on m and included in the final likelihood as described in Section 5.3.3.

An additional concern is the effect of intra-cluster light leaking into source photometry used for redshift estimation. We test for this in Gruen et al. (in preparation) with the intra-cluster light measure-

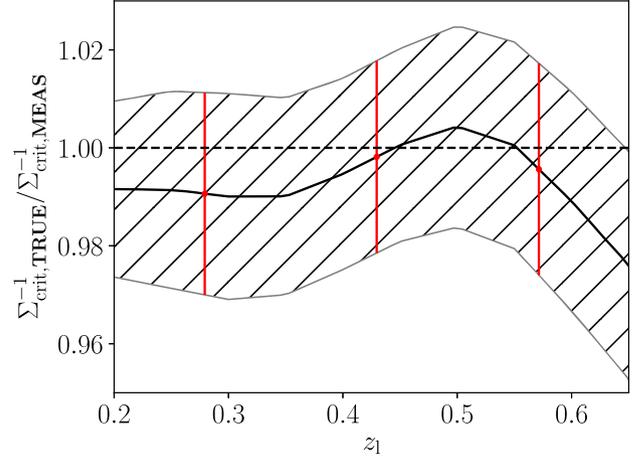


Figure 8. The photo- z correction factor to $\Sigma'_{\text{crit}}^{-1}$ as described in Section 4.3. The grey hatched region indicates the 1σ range of the correction factor. Red points with error bars show the correction factors applied to each redshift bin.

ments of Zhang et al. (in preparation), finding negligible effects in the regimes relevant for this study.

5 THE STACKED LENSING SIGNAL

5.1 Surface density model

Our surface density model remains unchanged from Melchior et al. (2017). The lensing signal is given by equation (7). The quantities $\bar{\Sigma}(R)$ and $\bar{\Sigma}(< R)$ are given by

$$\bar{\Sigma}(R) = \int_{-\infty}^{+\infty} d\chi \Delta\rho \left(\sqrt{R^2 + \chi^2} \right), \quad (28)$$

where R is the projected separation and χ the separation along the line of sight in comoving units and

$$\bar{\Sigma}(< R) = \frac{2}{R^2} \int_0^R dR' R' \bar{\Sigma}(R'). \quad (29)$$

If the shear signal is caused by haloes of mass M , the average excess 3D matter density is given by

$$\Delta\rho(r) = \rho(r) - \rho_m = \rho_m \xi_{\text{hm}}(r | M), \quad (30)$$

where $\rho_m = \Omega_m \rho_{\text{crit}}(1+z)^3$ is the mean matter density in physical units at the redshift of the sample, ρ_{crit} is the critical density at redshift zero, and $\xi_{\text{hm}}(r | M)$ is the halo–matter correlation function at the halo redshift.

At small scales, ξ_{hm} is dominated by the so-called ‘1-halo’ term while at large scales it is dominated by the ‘2-halo’ term. We use the Zu et al. (2014) update to the Hayashi & White (2008) model of ξ_{hm} . This is

$$\xi_{\text{hm}}(r | M) = \max \{ \xi_{1\text{h}}(r | M), \xi_{2\text{h}}(r | M) \}. \quad (31)$$

For the 1-halo term, we use a Navarro, Frenk & White (1996, hereafter *NFW*) density profile $\rho_{\text{NFW}}(r | M)$

$$\xi_{1\text{h}}(r | M, c) = \frac{\rho_{\text{NFW}}(r | M, c)}{\rho_m} - 1, \quad (32)$$

where

$$\rho_{\text{NFW}}(r | M, c) = \frac{\Omega_m \rho_{\text{crit}} \delta_c}{(r/r_s)(1+r/r_s)^2}. \quad (33)$$

Table 1. Parameters entering $\mathcal{L}(\Delta\Sigma)$ (equation 48) and $\mathcal{L}(\mathcal{B})$ (equation 49). Flat priors are specified with limits in square brackets, Gaussian priors with means \pm standard deviations.

| Parameter | Description | Prior |
|---------------------|-----------------------------------|-----------------|
| $\log_{10}M_{200m}$ | Halo mass | [11.0, 18.0] |
| c_{200m} | Concentration | [0, 20] |
| τ | Dimensionless miscentering offset | 0.17 ± 0.04 |
| f_{mis} | Miscentered fraction | 0.25 ± 0.08 |
| A_m | Shape & photo- z bias | equation (45) |
| B_0^{cl} | Boost magnitude | [0, ∞] |
| R_s^{cl} | Boost factor scale radius | [0, ∞] |

The concentration $c = r_{200m}/r_s$ is left as a free parameter, with a flat prior as per Table 1. This differs from the analysis in Melchior et al. (2017), in which we forced the halo concentration to follow the concentration–mass relation of Diemer & Kravtsov (2015).

For the two-halo term $\xi_{2h}(r|M)$, we use the non-linear matter correlation function ξ_{nl} scaled by the halo bias $b(M)$ of Tinker et al. (2008) as

$$\xi_{2h}(r|M) = b(M)\xi_{\text{nl}}(r). \quad (34)$$

ξ_{nl} is the 3D Fourier transform of the non-linear power spectrum P_{nl} (Smith et al. 2003; Takahashi et al. 2012), given by

$$\xi_{\text{nl}}(r) = \int_0^\infty \frac{dk}{k} \frac{k^3 P_{\text{nl}}(k)}{2\pi^2} j_0(kr), \quad (35)$$

where $j_0(z)$ is the 0th spherical Bessel function of the first kind. The power spectrum is computed using the CLASS code⁵ (Lesgourgues 2011; Blas, Lesgourgues & Tram 2011). We repeated our analysis using the linear matter correlation function ξ_{lin} and found nearly identical results as discussed later in Section 6.3.1.

5.2 Miscentering correction

We have thus far assumed that we can measure the stacked shear profile of clusters relative to the ‘centre’ of the halo as defined in an N -body simulation. Our simulations use the spherical overdensity algorithm ROCKSTAR as implemented in Behroozi, Wechsler & Wu (2013). If cluster centres are not properly identified, or are ‘miscentred’, then the observed weak lensing signal in annuli around these clusters will be suppressed. As in Melchior et al. (2017), we model the recovered weak lensing signal as a weighted sum of two independent contributions: a contribution $\Delta\Sigma_{\text{cen}}$ from properly centered clusters, and a contribution $\Delta\Sigma_{\text{mis}}$ from miscentred clusters,

$$\Delta\Sigma_{\text{model}} = (1 - f_{\text{mis}})\Delta\Sigma_{\text{cen}} + f_{\text{mis}}\Delta\Sigma_{\text{mis}}. \quad (36)$$

$\Delta\Sigma_{\text{cen}}$ is given by equation (7). When a cluster is miscentred by some radial offset R_{mis} , the corresponding azimuthally averaged surface mass density is (e.g. Yang et al. 2006; Johnston et al. 2007)

$$\bar{\Sigma}_{\text{mis}}(R|R_{\text{mis}}) = \int_0^{2\pi} \frac{d\theta}{2\pi} \bar{\Sigma} \left(\sqrt{R^2 + R_{\text{mis}}^2 - 2RR_{\text{mis}}\cos\theta} \right). \quad (37)$$

Letting $p(R_{\text{mis}})$ be the distribution of radial offsets for miscentred clusters, the corresponding mean miscentred profile $\bar{\Sigma}_{\text{mis}}$ is

$$\bar{\Sigma}_{\text{mis}}(R) = \int dR_{\text{mis}} p(R_{\text{mis}}) \bar{\Sigma}_{\text{mis}}(R|R_{\text{mis}}). \quad (38)$$

⁵<http://class-code.net/>

It is this quantity that we use to model the miscentered profile term in equation (36).

Zhang et al. (in preparation) and von der Linden et al. (in preparation) measure the centring fraction and centring distribution of redMaPPer clusters by comparing the reported centres to those derived from high-resolution X-ray data (where available). Here, we summarize their findings. The miscentering distribution $p(R_{\text{mis}})$ has the form

$$p(R_{\text{mis}}) = \frac{R_{\text{mis}}}{(\tau R_\lambda)^2} \exp\left(-\frac{R_{\text{mis}}}{\tau R_\lambda}\right), \quad (39)$$

where R_λ is the cluster radius assigned by redMaPPer, and $\tau = R_{\text{mis}}/R_\lambda$. Note that this is a Gamma distribution, which is more heavily tailed than the Rayleigh distribution used in Simet et al. (2017) and Melchior et al. (2017). This model choice is justified in Zhang et al. (in preparation) and von der Linden et al. (in preparation). In the latter, the cluster sample is complete for $\lambda \in [25, 35]$ and $\in [0.08, 0.12]$, and was selected to be representative of the SDSS redMaPPer cluster population. We use a combination of the posteriors from those two works to set the priors of the miscentering parameters f_{mis} and τ , as detailed in Table 1. The prior uncertainties conservatively encompass the spread in best-fitting values and the confidence intervals of both works. It corresponds to a miscentering fraction $f_{\text{mis}} = 0.25 \pm 0.08$, that is, roughly ≈ 75 per cent of the redMaPPer clusters are correctly centred. Because the variation in R_λ within each richness bin is mild, we ignore variations in R_λ across the bin, and set R_λ to the radius of a cluster whose richness is equal to the mean richness of the clusters in the bin. We have explicitly verified that if use the median rather than the mean cluster richness, the difference between our predicted profiles is insignificant relative to our statistical errors.

5.3 Multiplicative corrections

Multiplicative corrections to our model include boost factors, reduced shear, and shear + photo- z biases. These adjust our model according to

$$\Delta\Sigma_{\text{full}}(R) = \frac{\mathcal{A}_m \mathcal{G}(R)}{\mathcal{B}(R)} \Delta\Sigma_{\text{model}}. \quad (40)$$

In this equation, \mathcal{A}_m is the multiplicative correction due to shear and photometric redshift biases, $\mathcal{G}(R)$ is the multiplicative correction due to using reduced shear, and $\mathcal{B}(R)$ is the boost factor correction.

5.3.1 Boost factor model

In Section 4.2, we discussed how membership dilution biases the recovered weak lensing profile by a factor $1 - f_{\text{cl}}$. This factor is known as a boost factor because to correct for it in the lensing profile, one would increase the measured signal. We decided to not apply this factor to our data, and instead dilute the theoretical profile. To marginalize over the statistical uncertainty in our boost factor measurements, we parametrize the boost factor $\mathcal{B} \equiv (1 - f_{\text{cl}})^{-1}$ by constructing a model for the cluster-member contamination using a two component (B_0 and R_s) NFW profile:

$$\mathcal{B}_{\text{model}}(R) = 1 + B_0 \frac{1 - F(x)}{x^2 - 1}, \quad (41)$$

where $x = R/R_s$, and

$$F(x) = \begin{cases} \frac{\tan^{-1} \sqrt{x^2-1}}{\sqrt{x^2-1}} & : x > 1 \\ 1 & : x = 1 \\ \frac{\tanh^{-1} \sqrt{1-x^2}}{\sqrt{1-x^2}} & : x < 1 \end{cases} \quad (42)$$

We fit the boost factors measured in each bin along with the lensing profile of that bin. This introduces two additional parameters in our model, B_0 and R_s , for each richness and redshift bin.

5.3.2 Reduced Shear

We account for the fact that we measure the reduced shear g rather than true shear γ , as seen in equation (6), by multiplying our lensing model by

$$\mathcal{G}(R) = \frac{1}{1-\kappa} = \frac{1}{1-\Sigma(R)\Sigma_{\text{crit}}^{-1}}. \quad (43)$$

Here, $\Sigma_{\text{crit}}^{-1}$ is the same as that in equation (4.3) and $\Sigma(R)$ is

$$\Sigma(R) = (1 - f_{\text{mis}})\Sigma_{\text{cen}} + f_{\text{mis}}\Sigma_{\text{mis}}, \quad (44)$$

where Σ_{cen} is given by equation (28) and Σ_{mis} is given by equation (38). This adjustment has a negligible effect on our results, and introduces no new free parameters to our analysis.

5.3.3 Shear + photo- z bias

The factor $\mathcal{A}_m = 1 + m + \delta$ combines the effects of shear (m , Section 4.1) and photo- z (δ , Section 4.3) systematic uncertainties. Zuntz et al. (2018) found a shear calibration of $m = 0.012 \pm 0.013$. The photo- z bias comes from Hoyle et al. (2018) and varies between cluster stacks.

Since both m and δ are assigned Gaussian priors, the width of the prior on \mathcal{A}_m is obtained by adding the widths of the priors on m and δ in quadrature. We arrive at

$$\mathcal{A}_m = \begin{cases} 1.021 \pm 0.025 & \text{for } z \in [0.2, 0.35] \\ 1.014 \pm 0.024 & \text{for } z \in [0.35, 0.5] \\ 1.016 \pm 0.025 & \text{for } z \in [0.5, 0.65]. \end{cases} \quad (45)$$

The typical systematic uncertainty of ≈ 2.5 per cent represents a significant improvement over the typical systematic uncertainty of ≈ 3.8 per cent we achieved in Melchior et al. (2017). This dramatic improvement in accuracy is primarily driven by the improved shear calibration achieved in the DES Y1 data with METACALIBRATION.

For the following data releases of DES, we anticipate that improvements in the treatment of blended objects can further reduce the multiplicative shear bias. This implies that uncertainties in the calibration of photometric redshift estimates will likely be our dominant measurement related systematic. Significant improvements on this will require either extended calibration data sets or a hierarchical treatment that uses survey data to inform redshift estimation consistently.

5.4 Stacked mass corrections

We expect the masses we measure in Section 5.5 to be biased with respect to the true mean mass of the stacks. This bias arises from two sources: our model presented above is not a true description of cluster lensing profiles, and effects due to triaxiality and projection. We account for both sources of bias by calculating a correction \mathcal{C} applied to the expected mass of the stack $\mathcal{M}_{\text{true}} = \mathcal{C}\langle M \rangle$, as detailed

in the section below. This is applied after the lens modelling is complete, but before modelling the mass–observable relation from our stacked masses in Section 6.

5.4.1 Modelling systematics

The model presented above for $\Delta\Sigma$ is not perfect; our analytic model for the halo–mass correlation function in equation (31) does not match density profiles in simulations (Melchior et al. 2017; Murata et al. 2018), in particular, in the transition between the 1-halo and 2-halo regimes. In lieu of a fully calibrated model, we correct for any bias imparted by our choice of model by using our likelihood analysis to estimate halo masses of synthetic data generated from N -body simulations. The haloes are drawn from an N -body simulation of a flat Λ CDM cosmology run with GADGET (Springel 2005). The simulation uses 1400^3 particles in a box with $1050 h^{-1}$ Mpc on a side with periodic boundary conditions and for softening of $20 h^{-1}$ kpc. The simulation was run with the cosmology $\Omega_m = 0.318$, $h = 0.6704$, $\Omega_b = 0.049$, $\tau = 0.08$, $n_s = 0.962$, and $\sigma_8 = 0.835$. Halos of mass $10^{13} h^{-1} M_\odot$ are resolved with 100 particles. We discard all information below five softening lengths, and verified that the choice of extrapolation scheme for describing the correlation function below this scale does not impact our results. Halos were defined using a spherical overdensity mass definition of 200 times the background density and were identified with the ROCKSTAR halo finder (Behroozi et al. 2013).

The simulation is used to construct the synthetic $\Delta\Sigma$ profiles of galaxy clusters at four different snapshots: $z \in [0, 0.25, 0.5, 1]$. There were $\sim 420\,000$ haloes at $z = 1$ and $\sim 830\,000$ haloes at $z = 0$. We used snapshots instead of lightcones for two main reasons: we wanted to maximize the number of haloes we had available to perform the calibration, and we found that the synthetic profiles to only weakly depend on redshift. Instead of splitting haloes into mass subsets as in Melchior et al. (2017), we assigned a richness to each halo by inverting the mass–richness relation of Melchior et al. (2017) and adding 25 per cent scatter. We then grouped our haloes into richness subsets identical to how we grouped our clusters. For each of these halo subsets, we measured the halo–matter correlation function with the Landy & Szalay (1993) estimator as implemented in CORRFUNC⁶ (Sinha & Garrison 2017). We numerically integrate the halo–matter correlation function to obtain the $\Delta\Sigma$ profile as described in Section 5.1.

This $\Delta\Sigma$ profile contains none of the systematics that exist in the real data. To incorporate them, we modified this profile with the multiplicative corrections described in Section 5.3 and miscentering corrections in Section 5.2. We took the central values of our priors in Table 1 as well as values for B_0 and R_s from modelling the boost factors independently and modified the simulated $\Delta\Sigma$ profile according to equation (47). The observed mass M_{obs} for this simulated profile was obtained by using the same pipeline that we apply on the real data. When evaluating the likelihood in equation (50), we used the semi-analytic covariance matrix corresponding to the nearest cluster subset in redshift.

Denoting M_{true} as the mean mass of the haloes in the simulated stack, the calibration for each simulated profile is seen in Fig. 9. The calibration $\mathcal{C} = M_{\text{true}}/M_{\text{obs}}$ was then modelled as a function of

⁶<https://github.com/manodeep/Corrfunc>

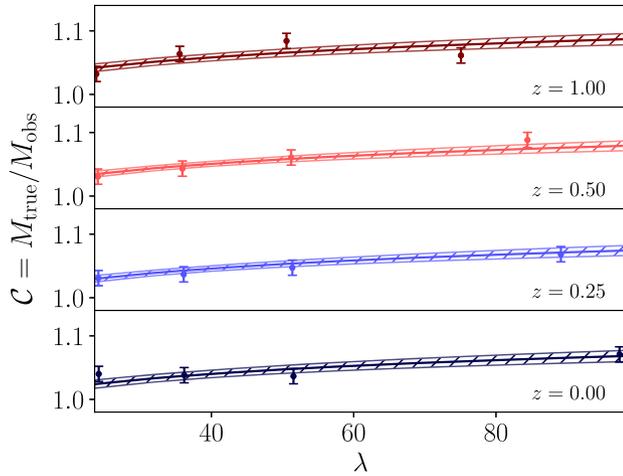


Figure 9. The mass calibration $C = M_{\text{true}}/M_{\text{obs}}$ from adopting our model of the correlation function in equation (31) as a function of λ and redshift. The solid line and hatched region are the best-fit model and 1σ uncertainty of the calibration. Error bars on the measured calibrations are the fitted intrinsic scatter σ_C .

the mean richness of the simulated stack $\bar{\lambda}$ in the form

$$C(\bar{\lambda}, z) = C_0 \left(\frac{\bar{\lambda}}{\lambda_0} \right)^\alpha \left(\frac{1+z}{1+z_0} \right)^\beta. \quad (46)$$

The free parameters in our fit are C_0 , α , and β with pivot values at $z_0 = 0.5$ and $\lambda_0 = 30$, as well as the intrinsic scatter σ_C of the calibration. The mean model bias for our cluster stacks is ≈ 4 per cent with $C_0 = 1.042 \pm 0.004$, $\alpha = 0.03 \pm 0.006$, and $\beta = 0.025 \pm 0.012$ as well as $\sigma_C = 0.01$. We repeated this process while assuming different amounts of intrinsic scatter in the M - λ relation from 10 per cent up to 45 per cent, as well as with no intrinsic scatter which is equivalent to the treatment of Melchior et al. (2017). We found that the amount of model bias increased with scatter in the M - λ relation. The model bias from Melchior et al. (2017) was recovered when no intrinsic scatter was present and using covariance matrices from that analysis.

We incorporated the dependence of the calibration on the intrinsic scatter in the M - λ relation as follows. We took the calibration described above at 25 per cent scatter to be our fiducial model as estimated in Rozo & Rykoff (2014). In addition to the covariance between C_0 , α , and β , we add additional uncorrelated uncertainty to each of these terms equal to half of the difference between the mean values obtained for these parameters assuming 15 per cent and 45 per cent scatter. This increased the variance of all three parameters C_0 , α , and β slightly. As discussed further in 8, the calibration contributed 0.73 per cent to the overall systematic uncertainty on the normalization of the M - λ relation.

One effect which we have not explicitly accounted for is the impact of baryonic physics on the recovered weak lensing masses. Baryonic cooling and feedback leads to mass redistribution in the central regions of galaxy clusters, with the impact of baryonic physics decreasing with increasing radius. Given that our fits allowed the concentration parameter of each cluster stack to float with no informative priors, we naively expect the impact of baryonic physics can be absorbed into the concentration of each cluster stack (Schaller et al. 2015). This naive expectation is confirmed in the work of Henson et al. (2017), who found that while baryonic physics impact the recovered weak lensing masses, the relative bias between the recovered weak lensing mass and the true mass

is roughly constant, independent of baryonic physics. Given that we measure masses at large radii (R_{200m}), and in light of the above results, we believe the impact of baryonic physics is likely to be less than ~ 3 per cent or so. Future work in which we explicitly test our fitting routines using hydrodynamic simulations is clearly desirable.

5.4.2 Triaxiality and projection effects

Photometric cluster selection preferentially selects haloes that are oriented with their major axis along the line of sight. Similarly, cluster selection is affected by other objects along the line of sight, which increases both the observed cluster richness and the recovered lensing mass. These two effects have been studied closely elsewhere (White et al. 2011; Angulo et al. 2012; Noh & Cohn 2012; Dietrich et al. 2014), and have competing effects on the recovered cluster masses. Dietrich et al. (2014) determined that triaxiality leads to an overestimation of the weak lensing mass and requires a correction factor of 0.96 ± 0.02 , while Simet et al. (2017) argued projection effects require that the recovered masses be multiplied by a factor of 1.02 ± 0.02 . Together, triaxiality and projection effects modify the recovered weak lensing masses by a multiplicative factor of 0.98 ± 0.03 . Our treatment is identical to that of Melchior et al. (2017), where additional details are provided. Although these two effects mildly depend on richness and redshift, we assume them to be constant in this analysis. We show the cumulative effect in Table 6. For reference, we have estimated the number of galaxy clusters that have another *cluster* within a 500 kpc radius along the line of sight. The number of such cases with $\lambda \geq 20$ is about 30, or 0.4 per cent of our sample, and thus negligible.

These effects as well as the correction for model bias are applied to the masses after fitting the lensing and boost factor data as described in Section 5.5, but before modelling the M - λ relation in Section 6.

5.5 The complete likelihood

The full model of the weak lensing profile is

$$\Delta\Sigma = \frac{\mathcal{A}_m \mathcal{G}(R)}{\mathcal{B}(r)} [(1 - f_{\text{mis}})\Delta\Sigma_{\text{cen}} + f_{\text{mis}}\Delta\Sigma_{\text{mis}}]. \quad (47)$$

Written this way the model includes the multiplicative bias \mathcal{A}_m , the boost factor $\mathcal{B}(r)$, the reduced shear correction $\mathcal{G}(R)$, and miscentring effects f_{mis} and $\Delta\Sigma_{\text{mis}}$. Using the semi-analytic covariance matrices $C_{\Delta\Sigma}$ described in Section 3.2.5, the log-likelihood of the k th $\Delta\Sigma$ profile is

$$\ln \mathcal{L}(\Delta\Sigma_k | M_k, c, \mathcal{A}_m, R_{\text{mis}}, f_{\text{mis}}, B_0, R_s) \propto -\frac{1}{2} \mathbf{D}_k^T C_{\Delta\Sigma}^{-1} \mathbf{D}_k \quad (48)$$

where $\mathbf{D} = (\widetilde{\Delta\Sigma} - \Delta\Sigma)$ and $\widetilde{\Delta\Sigma}$ is the measurement from equation (12).

The boost factor covariance matrix $C_{f_{\text{cl}}}$ is estimated from jackknifing. With this the corresponding log-likelihood of the measured $f_{\text{cl},k}$ in cluster subset k given the parameters in equation (41) is

$$\ln \mathcal{L}(f_{\text{cl},k} | B_0, R_s) \propto -\frac{1}{2} \mathbf{B}_k^T C_{f_{\text{cl}}}^{-1} \mathbf{B}_k \quad (49)$$

where $\mathbf{B}_k = (\mathcal{B} - \mathcal{B}_{\text{model}})_k$. Each boost factor \mathcal{B}_k is fit in conjunction with the associated lensing profile for the k th subset, allowing us to account for any degeneracies between the parameters in the respective models.

The total log-likelihood for a single cluster subset is

$$\begin{aligned} \ln \mathcal{L}_k = & \ln \mathcal{L}(\Delta \Sigma_k | M_k, c, \mathcal{A}_m, R_{\text{mis}}, f_{\text{mis}}, B_0, R_s) \\ & + \ln \mathcal{L}(f_{\text{cl},k} | B_0, R_s). \end{aligned} \quad (50)$$

Our goal is to constrain masses of independent cluster subsets M_k and boost factor parameters. Constraints on the latter are informed by both their effect on the $\Delta \Sigma$ profile as well as independent measurements of f_{cl} . The weak lensing and boost factor profiles are fit simultaneously, but each cluster subset is fit independently of other subsets.

5.6 Stacked cluster masses

A complete list of the model parameters describing each cluster stack as well as their corresponding priors are summarized in Table 1. The likelihood is sampled using the package *emcee*⁷ (Foreman-Mackey et al. 2013), which enables a parallelized exploration of the parameter space. We use 32 walkers with 10000 steps each, discarding the first 1000 steps of each walker as burn-in. We checked the convergence with independent runs of 5000 steps per walker that yielded identical results. After 14 steps the chains of single walkers become uncorrelated (with a correlation coefficient $|r| < 0.1$). This is much shorter than the total length of the chain. As a result, the number of independent draws between all walkers is ≈ 20500 .

The calibration correction described in 5.4 was applied to the recorded chains for each cluster subset. Specifically, for each point in the chain, we randomly sample the mass calibration factor $\mathcal{C}(\lambda, z)$ from its posteriors to adjust the mass. Further, we also apply the effect of triaxiality and projection effects (Section 5.4.2), both of which add 2 per cent to the uncertainty of each mass. In practice, this is written as

$$M = \mathcal{C}(\lambda, z) \mathcal{G}(0.96, 0.02) \mathcal{G}(1.02, 0.02) \times M_0, \quad (51)$$

where \mathcal{G} is a Gaussian and M_0 are the ‘uncalibrated’ masses in the chains. In this fashion, our final posteriors are properly marginalized over the uncertainty in the calibration factor \mathcal{C} as well as triaxiality and projection effects.

In order to characterize the contribution of both statistical and systematic uncertainties to our final results we perform our analysis three different times with three different sets of assumptions. These three analyses we run are:

- (i) **FULL**: All parameters (concentration, Shear + photo- z , boost factors, miscentering) are allowed to vary within their priors. This is our fiducial analysis.
- (ii) **FIXED \mathcal{A}_m** : \mathcal{A}_m is set to the center of its prior distribution but all other parameters are allowed to vary. This determines the contribution from the shape and photo- z uncertainties.
- (iii) **ONLY \mathcal{M}** : Only mass and concentration are free. All other parameter priors are set to δ -functions at their central values. This represents our statistical uncertainty on the mass.

Table 2 contains the results of the **FULL** analysis. Full posteriors from the cluster subset $z \in [0.2, 0.35]$ and $\lambda \in [20, 30]$ are shown in Fig. 10. The corresponding data and best-fit model are shown in Fig. 11, where we also demonstrate the combined effects of miscentering, boost factors, reduced shear, and multiplicative bias. The best-fit model for each richness and redshift bin is over-plotted on top of the weak lensing data in Fig. 4.

From the **FULL** analysis we can see the contribution of the various systematics to our final results. The boost factors amount to a correction of ≈ 2 per cent to $\Delta \Sigma$ at $R = 1$ Mpc. The posteriors on the miscentering parameters are equal to the priors, demonstrating that these parameters are only weakly constrained by the weak lensing data. In our earlier analysis (Melchior et al. 2017), we found a weak correlation between f_{mis} and M , which did not appear in this work. This was due to our use of the Diemer & Kravtsov (2014) M - c relation. We determined this by running one additional configuration in which the concentration was fixed by the Diemer & Kravtsov (2014) M - c relation, thus increasing the correlation between f_{mis} and M . At present, the contribution of miscentering to the mass is sub-dominant to other sources of systematic uncertainties in our final error budget (cf. Table 6). The multiplicative bias \mathcal{A}_m follows the prior and is degenerate with mass.

The **ONLY \mathcal{M}** likelihood evaluation allows us to quantify the statistical and systematic uncertainties of the fiducial analysis. The difference in quadrature between the uncertainties in the **FULL** and **ONLY \mathcal{M}** configurations represents the total systematic contribution to the error budget, while the **ONLY \mathcal{M}** alone provides the statistical contribution. The central values for each cluster subset along with statistical and systematic contributions to the uncertainties are presented in Table 2.

6 THE MASS-RICHNESS-REDSHIFT RELATION

The quantity we aim to constrain in this paper is the mean mass $\mathcal{M}(\lambda, z)$ of clusters of galaxies at a given observed richness λ and redshift z , similar to what was done in Melchior et al. (2015). Note that this is different from constraining the mean (and possibly distribution) of richness at given mass, or the full distribution of mass at given richness, as done in, e.g. Simet et al. (2017) and Murata et al. (2018). In particular, we neither constrain nor require a model of the intrinsic scatter in richness, hence making this analysis largely independent from the choices in subsequent cluster cosmology studies based upon it.

We note that an assumed value of the intrinsic scatter is used in two places: the semi-analytic covariance matrices described in Section 3.2.5 and the calibration described in Section 5.4.1. For the covariance, this assumption had a negligible effect compared to the shape noise. While the overall calibration did depend on the amount of scatter, we took a conservative approach by treating the difference in calibration between assuming 15 per cent and 45 per cent scatter as a systematic uncertainty. In this way, our final results are not sensitive to the amount of assumed intrinsic scatter.

6.1 Modelling the mass–richness relation

We fit a redshift-dependent power-law relation between cluster richness and cluster mass. Specifically, we set

$$\mathcal{M}(\lambda, z) \equiv \langle M | \lambda, z \rangle = M_0 \left(\frac{\lambda}{\lambda_0} \right)^{F_\lambda} \left(\frac{1+z}{1+z_0} \right)^{G_z}, \quad (52)$$

where M_0 , F_λ , and G_z are our model parameters. We select pivot values $\lambda_0 = 40$ and $z_0 = 0.35$, which are very near the median values of the cluster sample. Note \mathcal{M} is the expectation value of the mass of a halo as a function of richness and redshift.

The expected mass of a given cluster subset k represents a weighted mean of the masses of the individual clusters in that stack.

⁷<http://dan.iel.fm/emcee>

Table 2. Calibrated masses for each richness–redshift stack. All masses are in units of $\log_{10}M_{\odot}$ using the M_{200m} definition. Listed uncertainties are split into the symmetric 68 per cent confidence intervals of the systematic and statistical components, in that order. Adding the two in quadrature gives the total uncertainty on the mass.

| λ | $z \in [0.2, 0.35)$ | $z \in [0.35, 0.5)$ | $z \in [0.5, 0.65)$ |
|-----------------|------------------------------|------------------------------|------------------------------|
| [20, 30) | $14.191 \pm 0.013 \pm 0.032$ | $14.162 \pm 0.013 \pm 0.033$ | $14.083 \pm 0.015 \pm 0.048$ |
| [30, 45) | $14.477 \pm 0.014 \pm 0.031$ | $14.446 \pm 0.014 \pm 0.031$ | $14.456 \pm 0.015 \pm 0.041$ |
| [45, 60) | $14.608 \pm 0.011 \pm 0.044$ | $14.643 \pm 0.011 \pm 0.044$ | $14.648 \pm 0.016 \pm 0.056$ |
| [60, ∞) | $14.913 \pm 0.014 \pm 0.038$ | $14.899 \pm 0.015 \pm 0.038$ | $14.879 \pm 0.023 \pm 0.061$ |

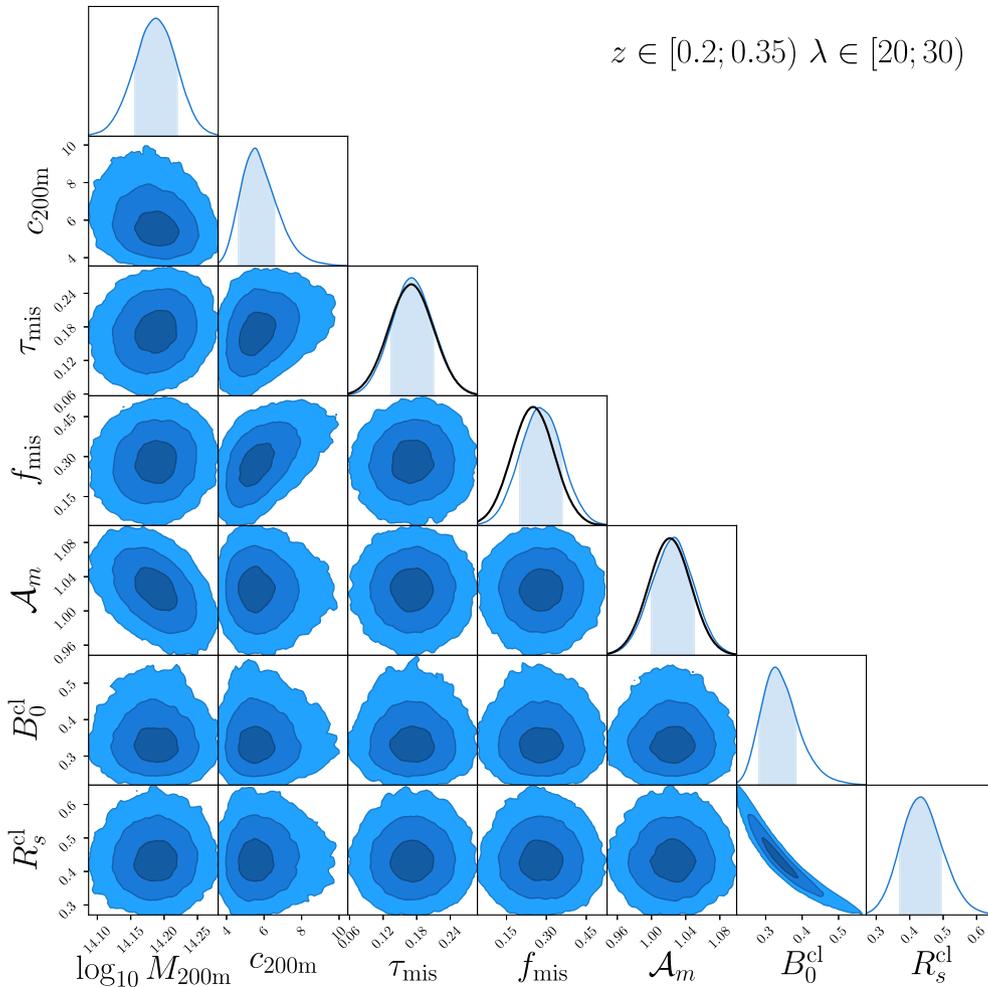


Figure 10. Posteriors for the parameters describing the lensing profile $\Delta\Sigma$ and the boost factor profile B for the bin $z \in [0.2, 0.35)$, $\lambda \in [20, 30)$. Contours show the 1σ , 2σ , and 3σ confidence areas. Black lines show the prior distributions. The mass presented here is uncalibrated, meaning it has not been corrected for modelling systematics, projection effects, or cluster triaxiality (see Section 5.4).

We then have

$$\mathcal{M}_k = \frac{\sum_{j \in k} \hat{w}_j \mathcal{M}(\lambda_j, z_j)}{\sum_{j \in k} \hat{w}_j}. \quad (53)$$

We take the weight \hat{w}_j of the j th cluster to be the sum of the weights of all lens–source pairs $w_{j,i}$ around that cluster from 0.3 Mpc and above and verified that the choice of radial range does not affect our recovered masses. Individual cluster weights \hat{w}_j differ from unity. This is because 1) the lensing weight of each lens–source pair given by equation (11) depends on the cluster redshift, and 2) in a given radial bin there are more sources associated with low redshift clusters since that bin subtends a larger angle on the sky compared

to the same bin around a high redshift cluster. In other words, the mass in the bin is skewed toward the average mass of the lower redshift clusters in the bin.

The impact of the weak lensing weights on the stacked mass estimates can be characterized by the ratio

$$a = \frac{\mathcal{M}_0}{\mathcal{M}_{\hat{w}}}. \quad (54)$$

We report the quantity $\log a$ in Table 3. We chose to report $\log a$ rather than a which has the computational advantage that one need only add $\log a$ to the mass values in Table 2 to arrive at an estimate of the mean cluster masses of cluster in a bin with unit weighting (as

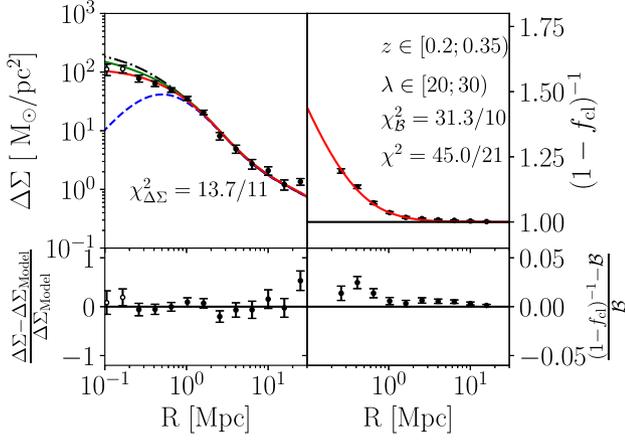


Figure 11. Fit with all components of the $\Delta\Sigma$ and B models for the cluster subset $z \in [0.2, 0.35]$ and $\lambda \in [20, 30]$. The top two panels show the best-fit models in red compared to the data. Unfilled points are not included in the fit. *Top left:* the black dot-dashed line is $\Delta\Sigma_{\text{cen}}$ while the blue-dashed line is $\Delta\Sigma_{\text{mis}}$. The weighted mean of these two yields the green solid line, and then applying the boost factor model, reduced shear, and multiplicative bias yields the final model in red. *Top right:* the red line is our NFW model for the boost factors. *Bottom:* the fractional difference between our data and models. The total χ^2 is 45 with 21 degrees of freedom, which is acceptable despite the imperfect fit of our simple model to the boost factors. The boost factors are measured from the data with small uncertainty, which is why the small mismatch with respect to the best-fit model causes a relatively large χ^2 but negligible effect on the recovered mass.

Table 3. The logarithm of the mean mass correction factor $\log_{10}a$ from equation (54). This represents a correction to the stacked cluster masses due to the fact that different clusters contribute to the measured mass in a different way than they contribute to $\Delta\Sigma$.

| λ | $z \in [0.2, 0.35]$ | $z \in [0.35, 0.5]$ | $z \in [0.5, 0.65]$ |
|-----------------|-------------------------|-------------------------|-------------------------|
| [20, 30) | -1.372×10^{-3} | -8.744×10^{-4} | -4.501×10^{-4} |
| [30, 45) | -2.979×10^{-3} | -3.278×10^{-3} | -6.660×10^{-4} |
| [45, 60) | -8.258×10^{-4} | -7.856×10^{-5} | -1.903×10^{-3} |
| [60, ∞) | 3.043×10^{-3} | -4.061×10^{-3} | 6.264×10^{-3} |

opposed to mean weak lensing weighted masses). The corrections in Table 3 are used to correct the recovered cluster masses to unit-weighted masses in the DES Y1 analysis of cluster abundances (DES Collaboration 2018).

6.2 Mass covariance

The purpose of our different chain configurations (Full, FixedAm, and OnlyMc) is to allow us to estimate the contribution of each systematic to the final uncertainty on the mass calibration parameters M_0 , F_λ , and G_z . In our analysis, there are seven sources of systematic uncertainty: multiplicative shear bias, multiplicative photo- z bias, miscentering, boost factors, modelling systematics, triaxiality and projection.

We discuss how we combine all systematics to arrive at a full covariance matrix between our bins that respects the covariance in our systematic error budget. The reader will recall that the original chains we obtain from fitting the weak lensing profiles are processed via equation (51) to account for the effect of calibration, triaxiality, and projection effects. If we let M_0 denote our unprocessed chains, and M denote the chains post-processing, in order to derive the statistical uncertainty only in our mass measurement we generate a

new chain \tilde{M} via

$$\tilde{M} = \bar{C} \times 0.96 \times 1.02 \times M_0. \quad (55)$$

The difference in the variance between chain M in equation (51) and that of chain \tilde{M} represents the uncertainty associated with calibration, triaxiality, and projection effects. We will use the M without a \sim to denote the chains post-processed as per equation (51), and \tilde{M} for chains post-processed as above.

The OnlyMc chain configuration contains only the statistical uncertainty in our analysis. For this reason, the covariance matrix describing the masses in this configuration is diagonal. We define the statistical uncertainty of the i th mass $\sigma_{i,\text{stat}}^2$

$$\sigma_{i,\text{stat}}^2 = \text{Var}(\tilde{M}_i^{\text{OnlyMc}}). \quad (56)$$

We also isolate the uncertainty associated with shear and photo- z systematics. To do so, we note the Full chain configuration includes all sources of uncertainty. Consequently, the difference in the variance between this chain and the FixedAm chain represents the uncertainty associated with shear and photo- z systematics. Therefore, we define

$$\sigma_{i,S+Pz}^2 = \text{Var}(M_i^{\text{Full}}) - \text{Var}(M_i^{\text{FixedAm}}). \quad (57)$$

Finally, the uncertainty associated with modelling systematics (calibration, triaxiality, and projection effects) takes the form

$$\sigma_{i,\text{mod}}^2 = \text{Var}(M_i^{\text{Full}}) - \text{Var}(\tilde{M}_i^{\text{Full}}). \quad (58)$$

By construction, the full uncertainty $\sigma_{i,\text{Tot}}^2$ satisfies

$$\text{Var}(M_i^{\text{Full}}) = \sigma_{i,\text{stat}}^2 + \sigma_{i,S+Pz}^2 + \sigma_{i,\text{mod}}^2. \quad (59)$$

We now define three different covariance matrices. First, \mathbf{C}^{stat} is diagonal, with $\mathbf{C}_{ii}^{\text{stat}} = \sigma_{i,\text{stat}}^2$. When we fit the weak lensing masses using this covariance matrix, we recover the statistical uncertainty in our scaling relation parameters. Second, $\mathbf{C}^{\text{S+Pz}}$ is defined via

$$\mathbf{C}_{ii}^{\text{S+Pz}} = \sigma_{i,\text{stat}}^2 + \sigma_{i,S+Pz}^2 \quad (60)$$

$$\mathbf{C}_{ij}^{\text{S+Pz}} = [\sigma_{i,S+Pz}^2 \sigma_{j,S+Pz}^2]^{1/2}. \quad (61)$$

Note the shear and photo- z components of the uncertainty are perfectly correlated across all bins. Fitting the weak lensing mass with this covariance matrix, and subtracting the statistical uncertainty in quadrature, enables us to calculate the uncertainty in our scaling relation parameters associated with shear and photo- z systematics. Third, \mathbf{C}^{Full} is defined via

$$\mathbf{C}_{ii}^{\text{Full}} = \sigma_{i,\text{stat}}^2 + \sigma_{i,S+Pz}^2 + \sigma_{i,\text{mod}}^2 \quad (62)$$

$$\mathbf{C}_{ij}^{\text{Full}} = [(\sigma_{i,S+Pz}^2 + \sigma_{i,\text{mod}}^2)(\sigma_{j,S+Pz}^2 + \sigma_{j,\text{mod}}^2)]^{1/2}. \quad (63)$$

Just like the shear and photo- z calibration uncertainties, modelling systematics are assumed to be perfectly correlated across all bins. The posteriors for the scaling relation parameters derived with this covariance matrix represent our full error budget, and the difference in quadrature between these errors and those obtained using the covariance matrix $\mathbf{C}^{\text{S+Pz}}$ give us the error budget associated with modelling systematics.

6.3 Likelihood for the mass-observable relation

We model the likelihood of the recovered weak lensing masses as Gaussian in the log. This is illustrated in Fig. 12, which shows the

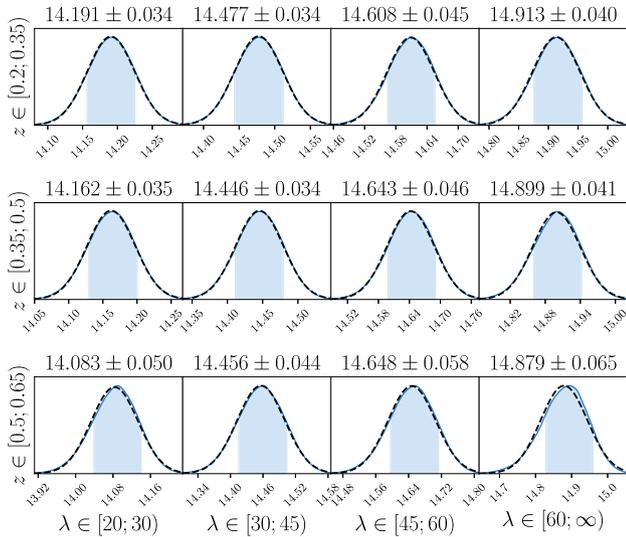


Figure 12. The calibrated posteriors of the masses for each cluster stack. Uncertainties appear above each panel, and are highlighted by the blue regions. Gaussian approximations to these posteriors appear as black dashed lines.

Table 4. Parameters of the M - λ - z relation from equation (64) with their posteriors. The mass is defined as M_{200m} in units of M_{\odot} . The pivot richness and pivot redshift correspond to the median values of the cluster sample. Uncertainties are the 68 percent confidence intervals and are split into statistical (first) and systematic (second).

| Parameter | Description | Posterior |
|-----------------|------------------|------------------------------|
| $\log_{10} M_0$ | Mass pivot | $14.489 \pm 0.011 \pm 0.019$ |
| F_{λ} | Richness scaling | $1.356 \pm 0.051 \pm 0.008$ |
| G_z | Redshift scaling | $-0.30 \pm 0.30 \pm 0.06$ |

posterior for each of the 12 cluster bins with $\lambda \geq 20$, along with the corresponding Gaussian approximation. We rely on the $\lambda \geq 20$ cluster bins exclusively as it is only these systems for which we are confident we can unambiguously map haloes to clusters and vice-versa. The full likelihood function is

$$\ln \mathcal{L}(M_{\text{obs}} | M_0, F_{\lambda}, G_z) \propto -\frac{1}{2} (\Delta \log M)^T \mathbf{C}_M^{-1} (\Delta \log M), \quad (64)$$

where \mathbf{C}_M is the covariance between the mass bins for a particular configuration (see Section 6.2). In the above equation, $\Delta \log M$ is the difference between the measured mass of each cluster subset and our model for the expected mass given as per equation (52). Thus, for the k th bin

$$\Delta \log M_k = \log M_k - \log \mathcal{M}_k. \quad (65)$$

We sample the posterior of the MOR parameters using *emcee* with 48 walkers taking 10000 steps each, discarding the first 1000 steps of each walker as burn-in. Table 4 summarizes the posteriors of our model parameters, while Fig. 13 shows the corresponding confidence contours. All parameters in the M - λ - z relation have flat priors.

We explicitly enforce correlated uncertainties of shear, photo- z , modelling systematics, and triaxiality and projection effects. Mis-centering and boost factors are considered independent across cluster subsets. These independent uncorrelated systematics will tend to average out across bins.

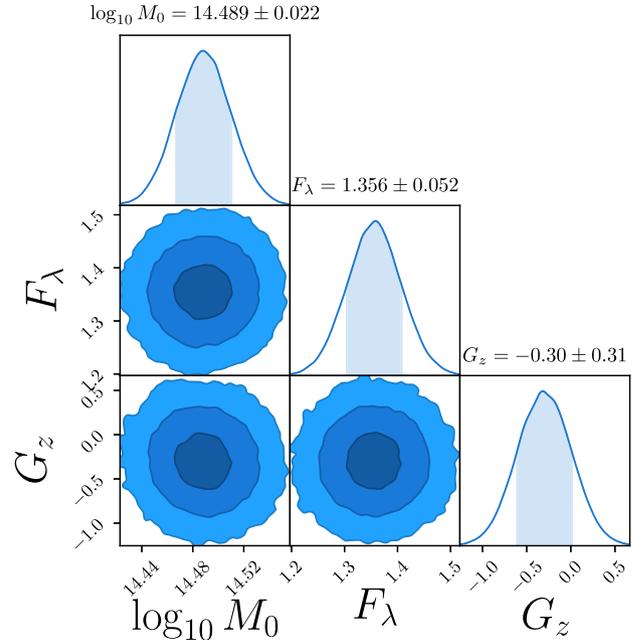


Figure 13. Parameters of the M - λ - z relation. Contours show the 1 σ , 2 σ and 3 σ confidence areas from the Full run.

In order to distinguish between the systematic and statistical contribution to the error budget on the M - λ - z relation parameters, we repeat the analysis using the statistical errors from the OnlyMc run. That is, we calculate equation (64) using only the uncertainties measured from the OnlyMc run, or the $\mathbf{C}_M^{\text{stat}}$ covariance matrix. The central values of the measured masses from the OnlyMc run are nearly identical to the Full run, as are the parameters in the M - λ - z relation. The difference in quadrature between the two uncertainties represents the systematic contribution while the excess uncertainty from the OnlyMc run is the statistical contribution. These uncertainties are reported in Table 4.

Our results imply that galaxy clusters of richness $\lambda = 40$ at redshift $z = 0.35$ have a mean mass of $\log_{10} \mathcal{M} = 14.489 \pm 0.011$ (stat) ± 0.019 (sys). The richness scaling is slightly steeper than linear at $F_{\lambda} = 1.356 \pm 0.051$ (stat) ± 0.008 (sys), while the mass shows a weak redshift dependence of $G_z = -0.30 \pm 0.30$ (stat) ± 0.06 (sys) consistent with no evolution. This amounts to a 5.0 per cent calibration (2.4 per cent statistical, 4.3 per cent systematic), of the normalization of the M - λ - z relation.

In Melchior et al. (2017), we found that the dominant systematic uncertainty stemmed from shear and photo- z systematics, as was the case in Simet et al. (2017). By repeating our analysis with the FixedAm run, which includes all systematics *except* \mathcal{A}_m , we are able to quantify the contribution from these sources. We found that the posterior distributions from the M - λ - z relation are significantly reduced, and that shear and photo- z systematics alone account for 48 per cent of the systematic uncertainty. This means that the remaining 52 per cent of the systematic uncertainty is due to modelling systematics, projection effects, and cluster triaxiality.

6.3.1 Alternative model using ξ_{lin}

Hayashi & White (2008) used a similar model to ours, but with the *linear* matter correlation function for their 2-halo term. This causes a very different behaviour near the 1-halo to 2-halo transition region, which can affect the fitting procedure, as discussed in Melchior et al.

Table 5. redMaPPer scaling relation comparisons from the literature. Of note, the Simet et al. (2017) results have changed slightly (Simet, private communication). We evaluate $\log_{10}\langle M|\lambda = 40, z = 0.35 \rangle$ of the other scaling relations in order to compare them to our result, applying the richness correction given by equation (66). When necessary, we use the method presented in Evrard et al. (2014) to convert from richness–mass relations to mass–richness relations. All masses are M_{200m} .

| Authors | Description | $\log(M \lambda = 40, z = 0.35) [M_{\odot}]$ | Richness scaling index F_{λ} |
|------------------------|---|--|--------------------------------------|
| This work | Weak lensing calibration using DES Y1 | 14.489 ± 0.022 | 1.356 ± 0.052 |
| Melchior et al. (2017) | Weak lensing calibration using DES SV | 14.540 ± 0.067 | 1.12 ± 0.21 |
| Baxter et al. (2018) | CMB lensing calibration using DES Y1 | 14.49 ± 0.31 | 1.24 ± 0.30 |
| Baxter et al. (2016) | Cluster clustering using SDSS | 14.7 ± 0.1 | 1.18 ± 0.16 |
| Simet et al. (2017) | Weak lensing calibration using SDSS | 14.48 ± 0.03 | 1.30 ± 0.09 |
| Murata et al. (2018) | Weak lensing calibration using SDSS | 14.533 ± 0.013 | 1.167 ± 0.052 |
| Farahi et al. (2016) | Pairwise velocity dispersion using SDSS | 14.42 ± 0.10 | 1.31 ± 0.14 |
| Saro et al. (2015) | SZE mass calibration using SPT and DES SV | 14.44 ± 0.05 | 0.91 ± 0.18 |
| Mantz et al. (2016) | Weak lensing of individual WtG clusters | 14.42 ± 0.11 | 1.36 ± 0.21 |
| Geach & Peacock (2017) | CMB lensing calibration using SDSS | 14.37 ± 0.05 | 0.74 ± 0.3 |

(2017). We repeated our entire analysis, including recomputing the calibration, using ξ_{lin} in place of ξ_{nl} . The masses of the stacks changed by less than 1 per cent, as did the normalization of the M – λ relation $\log_{10}M_0$. This means that our approach of calibrating the masses is largely robust to our choice of model.

6.3.2 Additional tests

We performed additional tests to verify our results. To ensure against possible small-scale systematic effects, we repeated our analysis with a more conservative radial cut of 500 kpc rather than 200 kpc. The resulting M – λ – z relation changed only in the mass scale, with M_0 changing by 0.2σ .

We also tested against possible differences in modelling systematics between large and small scales. By dividing each $\Delta\Sigma$ profile at 2 Mpc into large and small scale samples we could fit these regimes independently. While the constraining power was greatly diminished, the recovered masses were consistent with each other and the fiducial value within errors. No trend was observed in the differences between the recovered masses in any of these tests compared to the fiducial masses in Table 2.

Lastly, we tested an extension of equation (52) where $F_{\lambda}(z) = F_{\lambda,0} + zF_{\lambda,1}$ and found $F_{\lambda,1}$ consistent with 0 at the 1.2σ level. Therefore, if any redshift evolution exists in the richness scaling, we are unable to resolve the behaviour at present.

7 COMPARISON TO RESULTS IN THE LITERATURE

We compare our calibration of the M – λ relation to previous results from the literature. The specific richness–mass relations we consider are summarized in Table 5, and we describe below the origin of each of these.

(i) Melchior et al. (2017) was the precursor to this analysis. In that work, we calibrated the mass–richness relation of redMaPPer clusters in the DES Science Verification data. A detailed description of the changes between that analysis and this one appears in the next section.

(ii) Baxter et al. (2018) used the lensing of the Cosmic Microwave Background as measured by the South Pole Telescope to measure the mass–richness relation of DES Y1 redMaPPer clusters. Their analysis focused on 7066 clusters with richness $20 \leq \lambda \leq 40$. The upper limit was set to avoid potential biases in the recovered masses

Table 6. Systematic error budget on the amplitude of the mass–richness relation as measured with the DES Y1 data compared to the DES SV result of Melchior et al. (2017). The shear (Section 4.1), photo- z (Section 4.3), modelling systematics (Section 5.4.1), triaxiality, and projection effects (Section 5.4.2) are conservatively taken to be perfectly correlated between cluster stacks. Membership dilution (Section 4.2) and miscentering (Section 5.2) are assumed to be independent. Statistical and systematic errors are added in quadrature to arrive at the total error.

| Source of systematic | SV Amplitude uncertainty | Y1 Amplitude Uncertainty |
|------------------------------------|--------------------------|--------------------------|
| Shear measurement | 4% | 1.7% |
| Photometric redshifts | 3% | 2.6% |
| Modelling systematics | 2% | 0.73% |
| Cluster triaxiality | 2% | 2.0% |
| Line-of-sight projections | 2% | 2.0% |
| Membership dilution + miscentering | $\leq 1\%$ | 0.78% |
| Total Systematics | 6.1% | 4.3% |
| Total Statistical | 9.4% | 2.4% |
| Total | 11.2% | 5.0% |

from contamination by thermal Sunyaev-Zel’dovich emission by the clusters.

(iii) Simet et al. (2017) measured the mass–richness relation of redMaPPer clusters found in the Sloan Digital Sky Survey (SDSS). While their analysis is similar in spirit to ours, there are numerous methodological differences, including modelling choices (Simet et al. only fit the 1-halo term in the lensing profile), different radial scales used in the fit, a different shape catalog, and different photometric redshift catalogs.

(iv) Murata et al. (2018) measured the richness–mass relation of SDSS redMaPPer clusters assuming a *Planck* cosmology. We compute the mean mass at $\lambda = 40$ as well as the local slope at this point in the scaling relation. As demonstrated in Murata et al. (2018), their work and Simet et al. (2017) are consistent with each other, despite the fact that they used different models for $\Delta\Sigma$, different radial scales and slightly different richness bins. Of special note is the fact that while Simet et al. (2017) modelled only the 1-halo term using an NFW profile (along with a calibration step to correct for any biases introduced by this choice), Murata et al. (2018) used an emulator approach to simultaneously model the 1-halo and 2-halo terms of the lensing profile. The authors constrained the richness–mass relation using both lensing and cluster abundance data, and the use of the emulator effectively fixed the concentration–mass

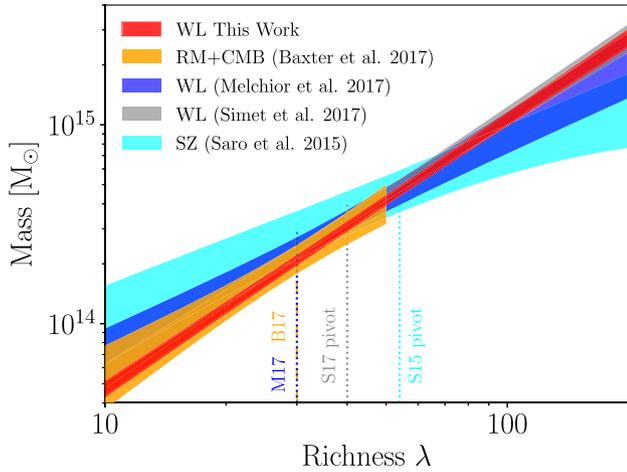


Figure 14. Best fit model for M – λ relation evaluated at the pivot redshift of our model, $z_0 = 0.35$, compared to other measurements. Our pivot richness is at $\lambda_0 = 40$. The previous DES result is in blue, from Melchior et al. (2017), while the relation measured in this analysis is in red. The analysis by Baxter et al. (2018) in orange used the same clusters as this work and found a consistent scaling relation over the richness range it probed.

relation. These differences add significant information relative to a lensing-only analysis. Finally, the posteriors we had available did not include the effects of photo- z or shear uncertainty in the error budget. Together, these differences result in error bars that are tighter than our own.

(v) Baxter et al. (2016) analysed the cluster clustering of SDSS redMaPPer clusters. By measuring the angular correlation function of clusters they were able to constrain the amplitude of the mass scaling relation to 18 per cent, in which their dominant systematic was uncertainty in the bias–mass relation.

(vi) Farahi et al. (2016) measured masses using stacked pairwise velocity dispersion measurements of SDSS redMaPPer clusters. Their measurements serve as a good cross check against other analyses of SDSS clusters, but found that they are ultimately less precise due to large uncertainties in velocity bias.

(vii) Saro et al. (2015) measured the mass–richness relation of galaxy clusters by assuming a *Planck* cosmology to determine the observable–mass relation of clusters from the South Pole Telescope (Bleem et al. 2015). They then matched these SPT clusters to redMaPPer clusters from the DES Science Verification data, and use the overlap sample to determine the richness–mass relation. We invert the relation using the method of Evrard et al. (2014) in order to show the comparison in Fig. 14.

(viii) Mantz et al. (2016) compared the scaling relation measured from the Weighting the Giants mass estimates for individual redMaPPer clusters in SDSS from Applegate et al. (2014) to that of the Simet et al. (2017) analysis. They found the two scaling relations in good agreement, which is also the case when compared to our measurement.

(ix) Geach & Peacock (2017) constrained the mass–richness relation of redMaPPer clusters found in SDSS using convergence profiles measured from *Planck* data. They constrain the normalization of the scaling relation at the ~ 11.5 per cent level, but are unable to reach similar precision for the scaling index. Mass calibrations from this type of measurement are expected to improve in the future as both optical cluster catalogs expand and CMB lensing maps improve.

(x) van Uitert et al. (2016) focused on significantly lower richness clusters found with a different algorithm. For these reasons, a direct comparison is not possible, however they were able to constrain the mass–richness scaling relation at low redshifts for their cluster finder at the 5 per cent level using lensing and cluster–satellite correlations.

Table 5 summarizes these scaling relations. Critically, the richness definition λ is sensitive to the details of image processing, source detection, choice of magnitudes, etc., and can therefore vary systematically from one survey to the next. We explicitly correct for this impact cross-matching DES Y1 clusters to DES SV and SDSS redMaPPer clusters, and measuring the richness offset. We find

$$\lambda_{\text{DES SV}} = (1.08 \pm 0.16)\lambda_{\text{DES Y1}} \quad (66)$$

$$\lambda_{\text{SDSS}} = (0.93 \pm 0.14)\lambda_{\text{DES Y1}}. \quad (67)$$

In these equations, the error is the standard deviation in the richness ratio, not the error on the mean. We apply these corrections to the SDSS and DES SV scaling relations before comparing to our result. So, for instance, if the scaling relation for data set X takes the form

$$\langle M | \lambda_X \rangle = A \lambda_X^\alpha \quad (68)$$

and the ratio $\lambda_X / \lambda_{\text{DES Y1}} = r$, then the scaling relation for Y1 richnesses is

$$\langle M | \lambda_{\text{DES Y1}} \rangle = A r^\alpha \lambda_{\text{DES Y1}}^\alpha. \quad (69)$$

Finally, all scaling relations that do not explicitly incorporate redshift evolution are transported from their quoted pivot redshift to our chosen pivot redshift $z = 0.35$ using our best fit redshift evolution.

Fig. 15 and Table 5 show the mass at $\lambda = 40$ and $z = 0.35$ as well as the richness scaling index for each of the scaling relations described above.

8 SYSTEMATIC IMPROVEMENTS FROM DES SV TO DES YEAR 1, AND FROM YEAR 1 TO YEAR 5

Our current analysis has multiple significant improvements relative to Melchior et al. (2017), the precursor to this work. Specifically:

(i) Shear calibration related errors on mass decreased from 4 per cent to 1.7 per cent, based primarily on the data driven correction of shear biases with METACALIBRATION. This implies that the shear calibration uncertainty is no longer the dominant source of systematic error in our weak lensing analysis.

(ii) The largest contribution to the systematic uncertainty is now photo- z errors. In the purely COSMOS-based calibration applied in this work, we find only a minimal improvement between SV and Y1.

(iii) The ≈ 15 per cent increase in uncertainty due to using noisy jackknife estimates of the covariance matrix (Dodelson & Schneider 2013) of the weak lensing profiles was entirely removed through the use of a semi-analytic covariance matrix.

(iv) Uncertainty from modelling systematics decreased from 2 per cent to 0.73 per cent. In Melchior et al. (2017), the model calibration corrections were computed by stacking haloes in mass bins. By contrast, our current analysis assigns richness according to a fiducial richness–mass relation, allowing us to stack in richness bins and to therefore accurately compute the correction for a richness bin. Relative to our SV analysis, the amplitude of this correction increased, while its uncertainty was reduced, from 1.00 ± 0.02 to 1.042 ± 0.004 . The increase in the correction is primarily due to the richness binning of the haloes as well as the fact that by using

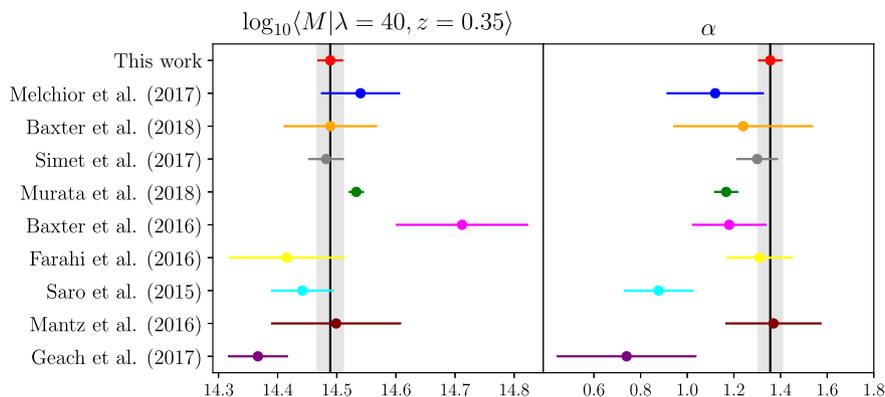


Figure 15. Comparison of the predicted mass at $\lambda = 40$ and $z = 0.35$ as well as the richness scaling relation between this work (*grey bands*) and other results from the literature.

the Y1 covariance matrix, the impact of small differences on large scales between our analytical model and numerical simulations is amplified, leading to larger correction factors. More importantly, however, the uncertainty on this correction was greatly reduced. This is due to the semi-analytic covariance matrix as well as allowing for intrinsic scatter of the calibration factors. While we use the semi-analytic covariance matrix for calibration on the simulated profiles, the matrix does not adequately describe the uncertainty in any systematic differences between the model and real data. Additionally, we now fit for the associated systematic uncertainty from the dispersion in the calibration data. Both of these factors result in a decrease in the systematic error budget.

(v) The mass–concentration relation of galaxy clusters is allowed to float in this analysis, while it was held fixed in Melchior et al. (2017). The fixed mass–concentration relation in our SV analysis was necessitated by the relatively low S/N of the weak lensing measurements. By contrast, our current analysis enables us to marginalize over concentration, which in turn should make our results significantly more robust to the impact of baryonic physics in the cores of galaxy clusters (see Section 5.4.1).

All in all, the reduced statistical and systematic uncertainty in our analysis has reduced the error in the amplitude of the mass–richness relation from 11.2 per cent to 5.0 per cent. Unlike our analysis in Melchior et al. (2017), our current constraints are close to systematics limited. Without improved systematics between now and the end of the survey, the improved statistics of the Year 5 data will only decrease our total error budget from 5.0 per cent to ≈ 4 per cent. Evidently, further reducing systematic uncertainties in future weak lensing mass calibration analysis is imperative.

Photometric redshift errors currently dominate the systematic error budget. Significant improvements in the weak lensing mass calibration of galaxy clusters will require new algorithms for calibrating photometric redshifts. A joint constraint from high-fidelity photo- z s of matched reference samples and clustering redshifts (Davis et al. 2018; Gatti et al. 2018), as done in Hoyle et al. (2018), is not feasible for higher redshift sources due to the limited redshift range of available spectroscopic or redMaGiC catalogs. Alternatively, source selection criteria that take into consideration photometric redshift uncertainty could lead to a desirable trade off between statistical and systematic uncertainty. With some combination of these approaches, reducing the photometric redshift uncertainty by a factor of two for a year 5 analysis seems plausible.

Related to this, in our current analysis, we have assumed that all systematics are perfectly correlated across all redshift and rich-

ness bins. This is likely *too* conservative. In particular, photometric redshift systematics are unlikely to be perfectly correlated across all redshift bins: the selection of the source population, and their relative signal contribution as a function of source redshift, differ as a function of lens redshift. Adequately characterizing the covariance in the systematic uncertainties associated with photometric redshift errors seems like a relatively simple way to significantly decrease our systematic error budget. For instance, if one were to assume that the photometric redshift systematics were entirely uncorrelated, the associated systematic would be reduced by a factor of $1/\sqrt{3}$, rendering photometric redshift errors sub-dominant. This is clearly unrealistic, but it does illustrate that characterizing the covariance in the systematics may lead to significant reductions in the total error budget.

Following photometric redshift uncertainties, three different effects come in at the ≈ 2 per cent level: shear systematics, triaxiality effects, and projection effects. Of these, shear systematics are the least problematic. We fully expect shear calibration uncertainties will continue to decrease over the coming years, and they will no longer be a major source of error for cluster mass calibration. By contrast, the current systematic error estimates for triaxiality and projection effects clearly demonstrate that there is a significant need for a detailed study of these on weak lensing mass profiles, such as in the recent work of Osato et al. (2018).

Additional, but less urgent, upgrades to our analysis are also possible. For instance, following Murata et al. (2018), an emulator based approach to modelling the halo-matter correlation function or $\Delta\Sigma$ directly can potentially greatly reduce the modelling calibration and its contribution to the uncertainty. Centring errors will also continue to decrease as the availability of multi-wavelength data continues to increase.

Finally, systematics that we have thus far ignored need to be better addressed. For instance, intrinsic alignment by cluster member galaxies even if its effect is very small (Sifón et al. 2015), which impact membership dilution estimates. Likewise, a study of the impact on baryonic physics on our weak lensing calibration methodology is necessary. While we expect these sources of error to be subdominant in our present study, quantifying the systematic error associated with these effects will be increasingly important in the future.

9 SUMMARY AND CONCLUSIONS

We measured the stacked weak lensing signal of redMaPPer clusters in the DES Y1 data. The clusters were divided into 21 subsets of

richness and redshift. The mean mass of each cluster stack was estimated for those subsets with $\lambda \geq 20$ and $0.2 \leq z \leq 0.65$. Our model incorporated:

- (i) Shear measurement systematics (Section 4.1),
- (ii) Source photometric redshift uncertainties (Section 4.3),
- (iii) Source sample dilution by cluster members (Section 4.2, Section 5.3.1),
- (iv) Cluster miscentering (Section 5.2),
- (v) Model calibration systematics (Section 5.4),
- (vi) Triaxiality & projection effects (Section 5.4.2).

The mean masses of the cluster subsets were used to determine the mean cluster mass as a function of richness and redshift according to equation (52). We emphasize that the full analysis was performed blindly: the paper underwent internal review by the DES collaboration prior to unblinding, and no changes to the analysis were made post-unblinding.

We summarize our constraints on the scaling relation as follows: for clusters at our pivot richness of $\lambda_0 = 40$ and pivot redshift of $z_0 = 0.35$, the mean cluster mass is

$$M_0 = [3.081 \pm 0.075 \pm 0.133] \cdot 10^{14} M_{\odot}. \quad (70)$$

The slope F_{λ} for the mass–richness relation is

$$F_{\lambda} = 1.356 \pm 0.051 \pm 0.008, \quad (71)$$

and the slope G_z governing the redshift evolution of the mass–richness relation is

$$G_z = -0.30 \pm 0.30 \pm 0.06, \quad (72)$$

where the first and second set of errors correspond to statistical and systematic errors, respectively. The full-scaling relation is given by equation (52). This scaling relation is in excellent agreement with, while being significantly more precise and accurate than, previous results from the literature: Saro et al. (2015); Mantz et al. (2016); Melchior et al. (2017); Simet et al. (2017); Baxter et al. (2018); Murata et al. (2018).

The 5.0 per cent constraint on the amplitude of the mass–richness relation is systematics-dominated, with our systematic error alone reaching 4.3 per cent. We stress the systematic uncertainty in the shear and photometric redshift catalogs have been extensively tested and validated, so we are confident our systematic error budget is robust. Halo triaxiality and line of sight projections are now important contributors to the total systematic error, and represent a critical path for minimizing the overall error budget for future analyses

Mass calibration remains the limiting factor for the ability of testing cosmological models with cluster counts. Nevertheless, this work represents a significant step forward: we were able to reduce the systematic error budget from 6.1 per cent in DES Science Verification to 4.3 per cent in DES Year 1. While we will need to achieve similar level of improvements for future analyses including DES Year 5 and LSST Year 1 to significantly improve upon our results, we are confident that we will be able to rise to the challenge: the story of weak lensing mass calibration is one of ever-decreasing systematic errors, a trend that to this day shows no signs of abating.

ACKNOWLEDGEMENTS

This paper has gone through internal review by the DES collaboration. TM and ER are supported by DOE grant DE-SC0015975. 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’. ER ac-

knowledges additional support by the Sloan Foundation, grant FG-2016-6443. Support for DG was provided by NASA through Einstein Postdoctoral Fellowship grant number PF5-160138 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060. TJ, DH, and SE are supported by DOE grant DE-SC0010107. KR and SB acknowledge support from the UK Science and Technology Facilities Council via Research Training Grants ST/P000252/1 and ST/N504452/1, respectively. CV is supported by the Mexican National Council for Science and Technology via grant number 411117. AvdL is supported by DoE grant 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.

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

The 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 No. 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.

This work used simulations and computations performed using computational resources at SLAC and at NERSC.

REFERENCES

- Angulo R. E. et al., 2012, *MNRAS*, 426, 2046
 Annis J. et al., 2014, *ApJ*, 794, 120
 Applegate D. E. et al., 2014, *MNRAS*, 439, 48
 Baxter E. J., Rozo E., Jain B., Rykoff E., Wechsler R. H., 2016, *MNRAS*, 463, 205
 Baxter E. J. et al., 2018, *MNRAS*, in press
 Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, *ApJ*, 762, 109
 Benítez N., 2000, *ApJ*, 536, 571
 Bhattacharya S. et al., 2013, *ApJ*, 766, 32
 Blas D., Lesgourgues J., Tram T., 2011, *J. Cosmol. Astropart. Phys.*, 7, 034
 Bleem L. E. et al., 2015, *ApJS*, 216, 27
 Carlstrom J. E. et al., 2011, *PASP*, 123, 568
 Chang C. et al., 2018, *MNRAS*, 475, 3165
 Coe D. et al., 2006, *AJ*, 132, 926
 Dark Energy Survey Collaboration, 2016, *Phys. Rev. D*, 94, 022001
 Davis C. et al., 2018, *MNRAS*, 477, 2196
 DES Collaboration et al., 2018, *Phys. Rev. D*, 98, 043526
 Diemer B., Kravtsov A. V., 2014, *ApJ*, 789, 1
 Diemer B., Kravtsov A. V., 2015, *ApJ*, 799, 108
 Dietrich J. P. et al., 2014, *MNRAS*, 443, 1713
 Dietrich J. P. et al., 2017, *MNRAS*, preprint (arXiv:1711.05344)
 Dodelson S., Schneider M. D., 2013, 88, 063537
 Dodelson S., Heitmann K., Hirata C., Honscheid K., Roodman A., Seljak U., Slosar A., Trodden M., 2016, preprint (arXiv:1604.07626)
 Drlaca-Wagner A. et al., 2018, *ApJS*, 235, 33
 Efron B., 1982, *The Jackknife, the Bootstrap and other Resampling Plans*, CBMS-NSF Regional Conf. Ser. Applied Mathematics (SIAM). Society for Industrial and Applied Mathematics (SIAM), Philadelphia
 Evrard A. E. et al., 2014, *MNRAS*, 441, 3562
 Farahi A. et al., 2016, *MNRAS*, 460, 3900
 Flaugher B. et al., 2015, *AJ*, 150, 150
 Foreman-Mackey D. et al., 2013, *PASP*, 125, 306
 Friedrich O. et al., 2016, *MNRAS*, 456, 2662
 Friedrich O. et al., 2018, *Phys. Rev. D*, 98, 023508
 Gatti M. et al., 2018, *MNRAS*, 477, 1664
 Geach J. E., Peacock J. A., 2017, *Nature Astron*, 1, 795
 Gruen D., Brimiouille F., 2017, *MNRAS*, 468, 769
 Gruen D. et al., 2014, *MNRAS*, 442, 1507
 Gruen D. et al., 2015, *MNRAS*, 449, 4264
 Gruen D. et al., 2018, *Phys. Rev. D*, 98, 023507
 Hayashi E., White S. D. M., 2008, *MNRAS*, 388, 2
 Henson M. A. et al., 2017, *MNRAS*, 465, 3361
 Hoekstra H., 2003, *MNRAS*, 339, 1155
 Hoekstra H. et al., 2015, *MNRAS*, 449, 685
 Hoyle B. et al., 2018, *MNRAS*, 478, 592
 Huff E., Mandelbaum R., 2017, *ApJ*, preprint (arXiv:1702.02600)
 Jarvis M. et al., 2016, *MNRAS*, 460, 2245
 Johnston D. E. et al., 2007, preprint (arXiv:0709.1159)
 Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
 Lesgourgues J., 2011, preprint (arXiv:1104.2932)
 Mantz A. B. et al., 2015a, *MNRAS*, 446, 2205
 Mantz A. B. et al., 2015b, *MNRAS*, 446, 2205
 Mantz A. B. et al., 2016, *MNRAS*, 463, 3582
 Medezinski E. et al., 2010, *MNRAS*, 405, 257
 Medezinski E. et al., 2018a, *PASJ*, 70, 30
 Medezinski E. et al., 2018b, *PASJ*, 70, S28
 Melchior P. et al., 2015, *MNRAS*, 449, 2219
 Melchior P. et al., 2017, *MNRAS*, 469, 4899
 Miyatake H. et al., 2018, preprint (arXiv:1804.05873)
 Murata R. et al., 2018, *ApJ*, 854, 120
 Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
 Noh Y., Cohn J. D., 2012, *MNRAS*, 426, 1829
 Okabe N., Smith G. P., 2016, *MNRAS*, 461, 3794
 Osato K. et al., 2018, *MNRAS*, 477, 2141
 Planck Collaboration XXIV, 2016, *A&A*, 594, A24
 Prat J. et al., 2018, *Phys. Rev. D*, 98, 042005
 Rowe B. T. P. et al., 2015, *Astron. Comput.*, 10, 121
 Rozo E., Rykoff E. S., 2014, *ApJ*, 783, 80
 Rozo E. et al., 2010, *ApJ*, 708, 645
 Rozo E. et al., 2016, *MNRAS*, 461, 1431
 Rykoff E. S. et al., 2012, *ApJ*, 746, 178
 Rykoff E. S. et al., 2014, *ApJ*, 785, 104
 Rykoff E. S. et al., 2016, *ApJS*, 224, 1
 Saro A. et al., 2015, *MNRAS*, 454, 2305
 Schaller M. et al., 2015, *MNRAS*, 451, 1247
 Schneider P. et al., 1998, *MNRAS*, 296, 873
 Schrabback T. et al., 2018, *A&A*, 610, A85
 Sheldon E., 2015, *Astrophysics Source Code Library*, record ascl:1508.008
 Sheldon E. S., Huff E. M., 2017, *ApJ*, 841, 24
 Sheldon E. S. et al., 2004, *AJ*, 127, 2544
 Sifón C. et al., 2015, *A&A*, 575, A48
 Simet M., Battaglia N., Mandelbaum M., Seljak U., 2015, preprint (arXiv:1502.01024)
 Simet M. et al., 2017, *MNRAS*, 466, 3103
 Singh S., Mandelbaum R., Seljak U., Slosar A., Vazquez Gonzalez J., 2016, preprint (arXiv:1611.00752)
 Sinha M., Garrison L., 2017, *Astrophysics Source Code Library*, record ascl:1703.003
 Smith R. E. et al., 2003, *MNRAS*, 341, 1311
 Springel V., 2005, *MNRAS*, 364, 1105
 Takahashi R. et al., 2012, *ApJ*, 761, 152
 Tinker J. et al., 2008, 688, 709
 Troxel M. A. et al., 2018, *Phys. Rev. D*, 98, 043528
 Umetsu K. et al., 2011, *ApJ*, 738, 41
 van Uitert E. et al., 2016, *A&A*, 586, A43
 von der Linden A. et al., 2014a, *MNRAS*, 439, 2
 von der Linden A. et al., 2014b, *MNRAS*, 443, 1973
 White M. et al., 2011, *ApJ*, 728, 126
 Yang X. et al., 2006, *MNRAS*, 373, 1159
 Zu Y. et al., 2014, *MNRAS*, 439, 1628
 Zuntz J. et al., 2013, *MNRAS*, 434, 1604
 Zuntz J. et al., 2018, *MNRAS*, 481, 1149

APPENDIX A: THE REDMAPPER V6.4.17 CLUSTER CATALOG

The full redMaPPer DES Y1A1 catalogues will be available at <http://risa.stanford.edu/redmapper/> in FITS format. The catalogs will also be available from the online journal in machine-readable formats. We note that this is of the same format as Rykoff et al. (2016), and we point the reader to

Table A1. redMaPPer Y1A1 cluster catalog format.

| Name | Format | Description |
|------------------|-------------|---|
| ID | INT(4) | redMaPPer Cluster Identification Number |
| VLIM | INT(2) | One if in cosmology catalog, 0 otherwise |
| NAME | CHAR(20) | redMaPPer Cluster Name |
| RA | FLOAT(8) | Right ascension in decimal degrees (J2000) |
| DEC | FLOAT(8) | Declination in decimal degrees (J2000) |
| Z_LAMBDA | FLOAT(4) | Cluster photo- z_{λ} |
| Z_LAMBDA_ERR | FLOAT(4) | Gaussian error estimate for z_{λ} |
| LAMBDA | FLOAT(4) | Richness estimate λ |
| LAMBDA_ERR | FLOAT(4) | Gaussian error estimate for z_{λ} |
| S | FLOAT(4) | Richness scale factor |
| Z_SPEC | FLOAT(4) | Spectroscopic redshift for most likely center (-1.0 if not available) |
| COADD_OBJECTS_ID | INT(8) | DES COADD_OBJECTS_ID identification number |
| MAG_CM_G | FLOAT(4) | g MAG_CM magnitude for most likely central galaxy (SLR corrected) |
| MAGERR_CM_G | FLOAT(4) | error on g MAG_CM magnitude |
| MAG_CM_R | FLOAT(4) | r MAG_CM magnitude for most likely central galaxy (SLR corrected) |
| MAGERR_CM_R | FLOAT(4) | error on g MAG_CM magnitude |
| MAG_CM_I | FLOAT(4) | i MAG_CM magnitude for most likely central galaxy (SLR corrected) |
| MAGERR_CM_I | FLOAT(4) | error on g MAG_CM magnitude |
| MAG_CM_Z | FLOAT(4) | z MAG_CM magnitude for most likely central galaxy (SLR corrected) |
| MAGERR_CM_Z | FLOAT(4) | error on g MAG_CM magnitude |
| ZLUM | FLOAT(4) | Total membership-weighted z -band luminosity (units of L_*) |
| P_CEN[5] | 5×FLOAT(4) | Centering probability P_{cen} for 5 most likely centrals |
| RA_CEN[5] | 5×FLOAT(8) | R.A. for 5 most likely centrals |
| DEC_CEN[5] | 5×FLOAT(8) | Decl. for 5 most likely centrals |
| ID_CEN[5] | 5×INT(8) | DES COADD_OBJECTS_ID identification number for 5 most likely centrals |
| PZBINS[21] | 21×FLOAT(4) | Redshift points at which $P(z)$ is evaluated |
| PZ[21] | 21×FLOAT(4) | $P(z)$ evaluated at redshift points given by PZBINS |

Table A2. redMaPPer DES Y1A1 member catalog format.

| Name | Format | Description |
|------------------|----------|--|
| ID | INT(4) | redMaPPer Cluster Identification Number |
| RA | FLOAT(8) | Right ascension in decimal degrees (J2000) |
| DEC | FLOAT(8) | Declination in decimal degrees (J2000) |
| R | FLOAT(4) | Distance from cluster center (h^{-1} Mpc) |
| P | FLOAT(4) | Membership probability |
| P_FREE | FLOAT(4) | Probability that member is not a member of a higher ranked cluster |
| THETA_L | FLOAT(4) | Luminosity (z -band) weight |
| THETA_R | FLOAT(4) | Radial weight |
| MAG_CM_G | FLOAT(4) | g MAG_CM magnitude (SLR corrected) |
| MAGERR_CM_G | FLOAT(4) | error on g MAG_CM magnitude |
| MAG_CM_R | FLOAT(4) | r MAG_CM magnitude (SLR corrected) |
| MAGERR_CM_R | FLOAT(4) | error on r MAG_CM magnitude |
| MAG_CM_I | FLOAT(4) | i MAG_CM magnitude (SLR corrected) |
| MAGERR_CM_I | FLOAT(4) | error on i MAG_CM magnitude |
| MAG_CM_Z | FLOAT(4) | z MAG_CM magnitude (SLR corrected) |
| MAGERR_CM_Z | FLOAT(4) | error on z MAG_CM magnitude |
| Z_SPEC | FLOAT(4) | Spectroscopic redshift (-1.0 if not available) |
| COADD_OBJECTS_ID | INT(8) | DES COADD_OBJECTS_ID identification number |

Table A3. redMaPPer z_{\max} map format.

| Name | Format | Description |
|----------|----------|--|
| HPIX | INT(8) | HEALPIX ring-ordered pixel number (NSIDE = 4096) |
| ZMAX | FLOAT(4) | Maximum redshift of a cluster centered in this pixel |
| FRACGOOD | FLOAT(4) | Fraction of pixel area that is not masked |

Table A4. redMaPPer random points catalogue format.

| Name | Format | Description |
|--------|----------|--|
| RA | FLOAT(8) | Right ascension in decimal degrees (J2000) |
| DEC | FLOAT(8) | Declination in decimal degrees (J2000) |
| Z | FLOAT(4) | Redshift of random point |
| LAMBDA | FLOAT(4) | Richness of random point |
| WEIGHT | FLOAT(4) | Weight of random point |

that paper for further details. The cluster catalogue is described in Table A1, and the associated members in Table A2. The catalogue is the ‘full’ catalogue, with all clusters with $\lambda > 20$, and the volume-limited subset is flagged with the VLIM flag. The map of the maximum redshift of the volume-limited catalogue is described in Table A3, and the random points are described in Table A4.

APPENDIX B: PARAMETER POSTERIORS

When fitting the weak lensing profiles, some parameters are not constrained by a prior and are also not shared between cluster bins. These are the halo concentration c , the boost factor amplitude B_0 , and the boost factor scale radius R_s . Table B1 shows the posteriors for these three parameters for each cluster bin. As seen in Fig. 10 B_0 and R_s are highly degenerate.

Table B1. Lensing profile parameters not constrained by priors or shared between cluster bins. Uncertainties are the 68 per cent confidence intervals. Note that in the highest redshift and richness bin, the boost factor profile model scale radius had a bimodal distribution, and is not well constrained. This did not affect the mass estimate at all.

| λ | $z \in [0.2, 0.35)$ | $z \in [0.35, 0.5)$ | $z \in [0.5, 0.65)$ |
|---------------------------------------|---------------------|---------------------|---------------------|
| Concentration c | | | |
| [20, 30) | 5.81 ± 1.03 | 5.68 ± 1.14 | 4.76 ± 1.62 |
| [30, 45) | 4.53 ± 0.74 | 6.24 ± 1.08 | 3.61 ± 0.72 |
| [45, 60) | 4.38 ± 0.96 | 5.41 ± 1.17 | 4.76 ± 1.21 |
| [60, ∞) | 4.65 ± 0.82 | 3.19 ± 0.56 | 3.73 ± 1.02 |
| Boost factor amplitude B_0 | | | |
| [20, 30) | 0.34 ± 0.05 | 0.05 ± 0.01 | 0.13 ± 0.05 |
| [30, 45) | 0.37 ± 0.06 | 0.14 ± 0.04 | 0.13 ± 0.08 |
| [45, 60) | 0.27 ± 0.05 | 0.05 ± 0.02 | 0.09 ± 0.06 |
| [60, ∞) | 0.23 ± 0.03 | 0.21 ± 0.17 | 0.03 ± 0.04 |
| Boost factor scale radius R_s [Mpc] | | | |
| [20, 30) | 0.44 ± 0.06 | 0.89 ± 0.24 | 0.38 ± 0.11 |
| [30, 45) | 0.50 ± 0.07 | 0.44 ± 0.10 | 0.44 ± 0.18 |
| [45, 60) | 0.80 ± 0.15 | 1.72 ± 0.95 | 0.85 ± 0.37 |
| [60, ∞) | 1.37 ± 0.21 | 0.51 ± 0.23 | 35.94 ± 29.69 |

- ¹Department of Physics, University of Arizona, Tucson, AZ 85721, USA
- ²Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany
- ³Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, D-81679 München, Germany
- ⁴Kavli Institute for Particle Astrophysics & Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA
- ⁵SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
- ⁶Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
- ⁷Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA
- ⁸Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA
- ⁹Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany
- ¹⁰Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany
- ¹¹Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
- ¹²Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA
- ¹³Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA
- ¹⁴Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
- ¹⁵Department of Physics and Astronomy, Pevensy Building, University of Sussex, Brighton BN1 9QH, UK
- ¹⁶Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
- ¹⁷Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK
- ¹⁸Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland
- ¹⁹Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15312, USA
- ²⁰Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
- ²¹Department of Physics, The Ohio State University, Columbus, OH 43210, USA
- ²²Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
- ²³University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
- ²⁴Physics Department, Brandeis University, 415 South Street, Waltham MA 02453, USA
- ²⁵Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK
- ²⁶Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile
- ²⁷Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown 6140, South Africa
- ²⁸Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK
- ²⁹Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK
- ³⁰Laboratório Interinstitucional de e-Astronomia – LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ – 20921-400, Brazil
- ³¹Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ – 20921-400, Brazil
- ³²Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
- ³³National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA
- ³⁴Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain
- ³⁵Institut d'Estudis Espacials de Catalunya (IEEC), E-08193 Barcelona, Spain
- ³⁶Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain
- ³⁷Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), E-28040 Madrid, Spain
- ³⁸Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
- ³⁹Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
- ⁴⁰Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, E-28049 Madrid, Spain
- ⁴¹Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- ⁴²Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
- ⁴³Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
- ⁴⁴Department of Astronomy/Steward Observatory, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA
- ⁴⁵Australian Astronomical Observatory, North Ryde, NSW 2113, Australia
- ⁴⁶Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP 05314-970, Brazil
- ⁴⁷George 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
- ⁴⁸Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain
- ⁴⁹School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK
- ⁵⁰Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas SP, Brazil
- ⁵¹Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
- ⁵²Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA

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