Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **485,** 69–87 (2019)
Advance Access publication 2019 January 22

doi:10.1093/mnras/stz234

Weak-lensing analysis of SPT-selected galaxy clusters using Dark Energy Survey Science Verification data

C. Stern,[1,2] J. P. Dietrich,[1,2]★ S. Bocquet,[3] D. Applegate,[4,5] J. J. Mohr,[1,2,6] S. L. Bridle,[7] M. Carrasco Kind,[8,9] D. Gruen,[10,11]† M. Jarvis,[12] T. Kacprzak,[13] A. Saro,[1,14] E. Sheldon,[15] M. A. Troxel,[16,17] J. Zuntz,[18] B. A. Benson,[19,20,21] R. Capasso,[1,2] I. Chiu,[1,2] S. Desai,[22] D. Rapetti,[23,24] C. L. Reichardt,[25] B. Saliwanchik,[26,27] T. Schrabback,[5] N. Gupta,[1,2] T. M. C. Abbott,[28] F. B. Abdalla,[29,30] S. Avila,[31] E. Bertin,[32,33] D. Brooks,[29] D. L. Burke,[10,11] A. Carnero Rosell,[34,35] J. Carretero,[36] F. J. Castander,[37,38] C. B. D'Andrea,[12] L. N. da Costa,[34,35] C. Davis,[10] J. De Vicente,[39] H. T. Diehl,[19] P. Doel,[29] J. Estrada,[19] A. E. Evrard,[40,41] B. Flaugher,[19] P. Fosalba,[37,38] J. Frieman,[19,21] J. García-Bellido,[42] E. Gaztanaga,[37,38] R. A. Gruendl,[8,9] J. Gschwend,[34,35] G. Gutierrez,[19] D. Hollowood,[43] T. Jeltema,[43] D. Kirk,[29] K. Kuehn,[44] N. Kuropatkin,[19] O. Lahav,[29] M. Lima,[34,45] M. A. G. Maia,[34,35] M. March,[12] P. Melchior,[46] F. Menanteau,[8,9] R. Miquel,[36,47] A. A. Plazas,[48] A. K. Romer,[49] E. Sanchez,[39] R. Schindler,[11] M. Schubnell,[41] I. Sevilla-Noarbe,[39] M. Smith,[50] R. C. Smith,[28] F. Sobreira,[34,51] E. Suchyta,[52] M. E. C. Swanson,[9] G. Tarle,[41] and A. R. Walker[28]

(DES and SPT Collaborations)

Affiliations are listed at the end of the paper

Accepted 2019 January 16. Received 2018 December 10; in original form 2018 February 13

ABSTRACT

We present weak-lensing (WL) mass constraints for a sample of massive galaxy clusters detected by the South Pole Telescope (SPT) via the Sunyaev–Zel'dovich effect (SZE). We use *griz* imaging data obtained from the Science Verification (SV) phase of the Dark Energy Survey (DES) to fit the WL shear signal of 33 clusters in the redshift range $0.25 \leq z \leq 0.8$ with NFW profiles and to constrain a four-parameter SPT mass–observable relation. To account for biases in WL masses, we introduce a WL mass to true mass scaling relation described by a mean bias and an intrinsic, lognormal scatter. We allow for correlated scatter within the WL and SZE mass–observable relations and use simulations to constrain priors on nuisance parameters related to bias and scatter from WL. We constrain the normalization of the $\zeta - M_{500}$ relation, $A_{\rm SZ} = 12.0^{+2.6}_{-6.7}$ when using a prior on the mass slope $B_{\rm SZ}$ from the latest SPT cluster cosmology analysis. Without this prior, we recover $A_{\rm SZ} = 10.8^{+2.3}_{-5.2}$ and $B_{\rm SZ} = 1.30^{+0.22}_{-0.44}$. Results in both cases imply lower cluster masses than measured in previous work with and without WL, although the uncertainties are large. The WL derived value of $B_{\rm SZ}$ is ≈20 per cent lower than the value preferred by the most recent SPT cluster cosmology analysis. The method demonstrated in this work is designed to constrain cluster masses and cosmological parameters simultaneously and will form the basis for subsequent studies that employ the full SPT cluster sample together with the DES data.

★ E-mail: astro@joergdietrich.com
† Einstein Fellow.

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

1 INTRODUCTION

Clusters of galaxies are the most massive collapsed objects in the Universe. Their abundance as a function of cluster mass and redshift is sensitive to the underlying cosmology and depends on both the expansion history of the Universe and the process of structure formation (Henry & Arnaud 1991; White, Efstathiou & Frenk 1993; Haiman, Mohr & Holder 2001). The main methods for identifying galaxy clusters include X-ray emission from the hot ($T \approx 10^8$ K) intra-cluster medium (ICM; e.g. Edge et al. 1990), spatial overdensities of galaxies (e.g. Abell 1958), and the Sunyaev–Zel'dovich effect (SZE; Sunyaev & Zeldovich 1972). The SZE results from the inverse Compton scattering of background cosmic microwave background (CMB) photons by energetic electrons in the ICM. Although number counts of galaxy clusters constitute a powerful cosmological probe that is complementary to other probes (e.g. Vikhlinin et al. 2009; Mantz et al. 2015; de Haan et al. 2016), there are two major obstacles for a cosmological analysis that need to be overcome.

The first obstacle is a precise understanding of the selection function. The interpretation of number counts is limited by the knowledge of completeness and contamination of the cluster sample to relate observed number counts to the underlying true distribution that is predicted by cosmological theories. The South Pole Telescope (SPT; Carlstrom et al. 2011) cluster sample has a very clean, uniform and well understood selection function that corresponds approximately to a mass selection that is almost redshift independent above redshifts $z \sim 0.25$. The 2500 deg^2 SPT-SZ survey is of sufficient depth to allow one to construct an approximately mass-limited sample of galaxy clusters above a lower limit of $M_{500,c} \approx 3 \times 10^{14}$ M$_\odot$[1] out to the highest redshifts where these systems exist ($z \sim 1.7$) (Bleem et al. 2015). It has been demonstrated that cluster high frequency radio galaxies, whose emission could mask the SZE decrement, have only a modest impact on the completeness of SZE-selected galaxy cluster samples (Gupta et al. 2017), and the contamination is well described simply by noise fluctuations arising from Gaussian noise in the SPT maps (Song et al. 2012; Bleem et al. 2015). The SPT SZE cluster selection therefore emphasizes the high-mass and high-redshift part of the mass function, which is of particular interest for cosmological studies (see Vanderlinde et al. 2010; Benson et al. 2013a; Reichardt et al. 2013; Bocquet et al. 2015; de Haan et al. 2016).

The second obstacle is measuring the cluster masses. Samples of galaxy clusters can be constructed using observables (e.g. X-ray luminosity or in the case of SPT the significance of the SZE detection), which often also serve as mass proxies. These mass proxies often depend on the morphological state of the galaxy cluster and their scaling to total mass is not clear a priori, leading to systematic uncertainties in mass determination. To avoid biases arising from these uncertainties, the mass–observable scaling relations need to be calibrated against a low bias observable. Because weak lensing (WL) is sensitive to the projected mass density, it is well suited for this task. In the context of SZE-selected cluster samples, a number of studies have tested the SZE-based mass estimates against the WL-derived masses (e.g. Gruen et al. 2014; Israel et al. 2014; von der Linden et al. 2014; Hoekstra et al. 2015; Battaglia et al. 2016). These analyses were in part motivated by an apparent tension between cosmological constraints based on Planck CMB anisotropy and those based on galaxy clusters (Planck Collaboration et al. 2016a,b, respectively).

To properly address the WL-calibrated SZE observable–mass scaling relation out to intermediate redshifts with a large sample of clusters, one needs a wide-field imaging survey of sufficient image quality over a part of the sky imaged by an SZE survey. For this purpose, we present results from the Dark Energy Survey (DES; DES Collaboration 2005). DES is a large *grizY*-band imaging survey covering a total area of 5000 deg^2 in the southern sky. It is estimated to yield about 300 million galaxies up to $z = 1.4$ when complete. The regular observations started in Fall 2013 and are planned to continue for 5 yr. The quality and depth of the DES data are superior to any other preceding survey of similarly large footprint, in particular the Sloan Digital Sky Survey (SDSS). Prior to the main survey, a smaller area was observed to approximately full survey depth. The ∼200 deg^2 with science quality imaging from this Science Verification (SV) period were meant as a test bed for the main survey. Because DES has by design almost complete overlap with the area observed by SPT, it is a natural choice for a WL analysis of large samples of intermediate redshift SPT-selected clusters where individual follow-up on larger ground or space-based telescopes would be too costly. To demonstrate the utility of DES for this task, we present a first WL analysis of SPT-selected galaxy clusters in the DES SV footprint.

Melchior et al. (2015) demonstrated the suitability of DES data for cluster WL using a sample of four very massive galaxy clusters and a precursor pipeline of the regular DES data processing software. A subsequent work (Melchior et al. 2017) measured stacked shear profiles for a large sample of optically selected clusters. In this work, we will extend the WL analysis of individual clusters to higher redshifts and lower masses using the regular DES pipelines and data taken in regular survey mode observations. As our main goal, we will use the individual shear profiles to calibrate the mass–observable relation for SPT- selected clusters of galaxies. Our method allows us to simultaneously constrain cosmological parameters and mass–observable relation parameters in a self-consistent way and can be used for larger samples of SPT-selected clusters to this end.

This paper is organized as follows. In Section 2, we give an overview of the DES and SPT observations as well as the associated shear catalogues and cluster sample used in this analysis. Section 3 contains a description of the measurement of the cluster shear profiles together with the corrections we apply and tests we carry out to ensure robustness. In Section 4, we present results of our efforts to constrain the SZE observable–mass scaling relation, using the shear profiles from the previous section. We review our conclusions in Section 5.

Unless otherwise stated, we use a flat ΛCDM cosmology with a matter density parameter $\Omega_m = 0.3089$ and a Hubble parameter $H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with $h = 0.6774$, which are values extracted from a CMB analysis (TT, TE, and EE power spectra, combined with lowP and lensing) in combination with external constraints from baryon acoustic oscillations, the JLA supernova sample, and H_0 (Planck Collaboration et al. 2016a).

[1] $M_{500,c}$ denotes the mass enclosed by a sphere (radius $r_{500,c}$) where the enclosed mean density is 500 times the critical density of the Universe. For convenience, we also refer to these quantities as r_{500} and M_{500} in the following.

2 DATA

We provide a short overview of the entire DES programme and then describe the SV observations and shear catalogues used in this work, followed by a discussion of the SPT observations and the SZE-selected lens sample for this analysis.

2.1 DES observations

The DES (DES Collaboration 2005; DES Collaboration et al. 2016) is designed to yield multiband imaging in *grizY* bands over an angular footprint of 5000 deg^2. To this end, it uses the 570 Megapixel DECam (Flaugher et al. 2015) mounted on the 4-m Blanco telescope at the Cerro Tololo Inter-American Observatory (CTIO). Each filter is observed in 10 tilings of 90s exposures (*Y* band: 45 s during SV) over the 5 yr survey period, and scheduling of individual exposures employs the programme OBSTAC (Neilsen & Annis 2014). OBSTAC automatically creates the timing of exposures based on seeing, sky brightness, and survey status. Observations in *riz* bands (used for WL) are preferentially carried out in conditions of good seeing. Additionally, deeper survey fields of 30 deg^2 in total are visited every 4–7 d with the main goal of measuring light curves of supernovae. These supernova fields do not include *Y*-band imaging as part of regular survey operations but are significantly deeper than the main survey and visited regularly to provide finer time resolution. The survey benefits from the very wide 3 deg^2 field of view of DECam with a pixel size of 0.27 arcsec. The 90 per cent completeness limit in each band approaches 24th magnitude. Therefore, DES will be deeper than previous surveys of similar solid angle like SDSS and wider than surveys of comparable depth like CFHTLS. The median seeing is expected to be below 1 arcsec over the full survey, and due to the addition of the *Y* band the DES wavelength coverage extends farther into the infrared compared to SDSS.

In this analysis, we use SV phase observations mostly obtained under regular survey conditions, and shape measurements from the *r*, *i*, and *z* bands, though the photo-*z* estimates additionally rely on the *g* band. After completion of the SV observations, the main quality cuts on the SV catalogue removed the SPT-E field south of $\delta = -61^\circ$. This is the region in which the Large Magellanic Cloud resides, which has a different stellar locus than the Galaxy (affecting star-galaxy separation and photometric calibration), as well as R Doradus (the second brightest star in the infrared), which affects the photometry inside a circle of several degrees. What is more, the large number of double stars in this region complicates PSF estimation. The science-ready release of SV called 'SVA1 Gold' consists of coadd catalogues that include all of these cuts and requires object detection in all four *griz* bands. These coadd catalogues are used for object detection, flux measurements (for photo-*z*), and quality flags.

2.2 DES shear catalogues

The shear measurements are extracted from fitted models to all available individual exposures for a given object after removing blacklisted exposures, as described in Jarvis et al. (2016). We use the standard SV masks (Jarvis et al. 2016). These exclude circular areas around 2MASS stars and additionally remove the 4 per cent of the remaining area containing a large fraction (≈25 per cent) of objects, whose shape could not be reliably measured. Shear measurements were performed down to magnitude $R = 24.5$ and span 139 deg^2 after masking in the SPT-E field.

The DES SV area is covered by shear catalogues from two shape measurement pipelines. We use NGMIX[2] (Sheldon 2014), a Gaussian mixture model fitting shear measurement code, as our main shear measurement code. NGMIX uses shape information from *riz* optical bands and requires at least one valid exposure for each band. NGMIX, however, was not run on the entire SV footprint. For a subsample of our lenses that is not covered by the NGMIX analysis, we use *r*-band catalogues from the model-fitting shear measurement code IM3SHAPE instead. This is includes five clusters from the pointed cluster fields.

Both codes have been shown to work well with DES SV data and produce reliable shape catalogues that pass the essential quality tests for a variety of WL applications. For these and an extensive description of the DES SV shear pipeline and shape measurement codes we refer the interested reader to Jarvis et al. (2016). We emphasize that the choice of NGMIX was due to higher number densities after quality cuts, which is likely a result of using multiband data.

The codes have been run semi-independently: though the algorithms significantly differ, they share all previous steps of data reduction, including PSF estimation and blacklisting of exposures, as outlined in Jarvis et al. (2016). Both simultaneously fit to a number of single-epoch exposures for each object, instead of a fit on coadded images (where less information would be used). Galaxies have been selected according to the 'Modest classifier', which uses the `SExtractor` catalogue parameter `spread_model` and its measurement uncertainty (Bertin & Arnouts 1996; Desai et al. 2012) extracted from the *i*-band image (see discussion in Chang et al. 2015). We remove blended objects because those are expected to have unreliable shape measurements by demanding `FLAGS_I= = 0`.

2.2.1 *Blinding*

Many scientific analyses are subject to the attempted reproduction of already published results that involves tuning the data cuts to confirm previous or expected findings (Klein & Roodman 2005). We refer to this (unconscious) effect as 'observer bias'. Our analysis is blinded in the following way to avoid observer bias: directly after processing and as part of the general DES shear pipeline, all shear values in the catalogues are multiplied by a hidden factor between 0.9 and 1. This acts as an effective unknown multiplicative bias that translates into an overall shift of the WL-derived masses and therefore the normalization of the M–ζ scaling relation, $A_{\rm SZ}$. The shift due to blinding is of similar order to the mass uncertainty for the full stack, but exceeds the statistical uncertainties of cosmic shear and larger stacked lens samples that use the full SPT-E area. Only after the full analysis is fixed and all quality tests are passed, are the catalogues unblinded. However, in the process of internal collaboration review some additional tests were requested and have been carried out after unblinding.

2.2.2 *NGMIX*

NGMIX is a multipurpose image-fitting code. It includes a re-implementation of LENSFIT (Miller et al. 2007, 2013). In the version used for the DES SV shape catalogues, it fits an exponential disc model to the single-exposure galaxy images. NGMIX fits simultaneously to all valid exposures over the *riz*-bands and requires at

[2]https://github/esheldon/ngmix

least one valid exposure in each band. It uses a shape prior from an analytical form fitted to the ellipticity distribution of COSMOS galaxies (Mandelbaum et al. 2014). We use only objects with the following quantities as reported by NGMIX: error flag when using the exponential model EXP_FLAGS = 0 (this includes a cut on general NGMIX failures, i.e. FLAGS = 0), signal-to-noise SNR_R > 10, signal-to-noise ratio (SNR) of NGMIX size measure T, SNR_T_R > 1.0, and 0.4 < ARATE < 0.6. The last item is the acceptance rate of the NGMIX sampler and ensures convergence of the fit. These selection parameters are relaxed from the strict cuts suggested by Jarvis et al. (2016) and are based on our experience gained during creation of the shear catalogues and expectation that due to the overall lower source number compared to the SV cosmic shear study (Abbott et al. 2016; Becker et al. 2016) systematic biases will remain subdominant to the increased statistical uncertainties in this work. We will later demonstrate this assumption to hold in Appendix A.

We use an inverse variance weight for each galaxy i that takes into account shape noise and the (e_1, e_2) covariance matrix C, given by

$$w_i = \frac{2 \times \sigma_\epsilon^2}{C_{11} + C_{22} + 2 \times \sigma_\epsilon^2}, \tag{1}$$

where $\sigma_\varepsilon = 0.22$ is the shape noise contribution per component from COSMOS. We choose to use only the diagonal elements of the covariance matrix to ensure that w_i is invariant under rotations.

Noise effects and choice of prior influence the observed shear and can be corrected by dividing the shear by a sensitivity that is calculated during the run of NGMIX. Typically, the shear is underestimated before this correction. Because sensitivities are noisy, we apply this correction on the ensemble of all sources used for our fitting. This is a way to estimate biases in the shape measurement algorithm in a more direct way than using external simulations. Thus, the resulting shear is effectively unbiased. This procedure is similar to the correction for noise bias in the case of IM3SHAPE described in the next section.

2.2.3 *IM3SHAPE*

We use shape catalogues from an implementation of IM3SHAPE[3] (Jarvis et al. 2016), which was significantly improved over the version used in the simulation study of Zuntz et al. (2013). IM3SHAPE is a model fitting algorithm, using a de Vaucouleurs (1948) bulge or exponential disc model. Each object is fitted to both models, and the best-fitting model is chosen as an adequate description. The amplitude of each component is allowed to vary and may be negative, and the fitting is done simultaneously over all exposures in one band. Galaxies are selected prior to the run of IM3SHAPE for better performance.

As in the case of NGMIX, we use relaxed selection criteria. This includes `SNR > 10` and ratio of convolved image size relative to PSF `MEAN_RGPP_RP>1.15`. We choose these cuts for IM3SHAPE because our statistical error bars allow for some systematic uncertainty on the overall calibration. Our choice of cuts gives a number density (over the full SPT-E field and all redshifts) of $n_g = 9.2$ arcmin^{-2}, whereas the more conservative cuts employed the DES-SV cosmic shear analysis (Abbott et al. 2016; Becker et al. 2016) would give $n_g = 5.4$ arcmin^{-2}. We show in Appendix A that the inclusion of these additional galaxies leads to statistically undetectable differences in our mass calibration.

IM3SHAPE, as all shape measurement codes based on a maximum likelihood approach, shows systematic noise biases (Kacprzak et al. 2012), typically expressed in terms of multiplicative bias m_n and additive bias c_n (the latter for each component separately):

$$e_{obs} = (1 + m_n) \times e_{true} + c_n, \tag{2}$$

where e_{obs} is the observed ellipticity and e_{true} is the true ellipticity of a galaxy. Working only with circularly averaged profiles, the additive bias is expected to average out when masking effects are negligible. The multiplicative bias however scales the tangential shear profile and therefore influences the derived masses. With help of simulations based on galaxies from the COSMOS survey, we can express the noise bias as a function of IM3SHAPE SNR and galaxy size `MEAN_RGPP_RP`. The resulting correction is then applied to the ensemble of galaxies in a given bin.

2.2.4 *Error estimation*

If systematic effects can be neglected, the dominant source of error for a WL shear measurement comes from the intrinsic ellipticity dispersion. Therefore, in the absence of measurement noise the precision of a binned measurement of one shear component cannot be better than $\sigma_\epsilon / \sqrt{N_{gal}}$ where N_{gal} is the number of source galaxies used in a given radial bin and $\sigma_\varepsilon = 0.22$ is the intrinsic ellipticity dispersion. Because systematic uncertainties are in general hard to quantify, we use Jackknife errors as an empirical approach to estimate our measurement uncertainty on the shear profile. We calculate the signal by iteratively removing one of the used sources in each iteration. The covariance matrix for g_+ then can be calculated via

$$\mathrm{Cov}_{ij} = \frac{N-1}{N} \sum_{k=1}^{N} \left(g^k_{+,i} - \langle g^l_{+,i} \rangle_l\right) \left(g^k_{+,j} - \langle g^l_{+,j} \rangle_l\right), \tag{3}$$

where i and j denote radial bins, and g^k_+ is the tangential shear without galaxy k. Analogous formulae are used for $g_\times$ and $\Delta\Sigma$ in the following. In each case, we neglect off-diagonal terms for our analysis. We test the impact and determine that including the full covariance increases the mass fitted to the IM3SHAPE stack by about $0.1 \times 10^{14}\,\mathrm{M}_\odot$ or $\approx 0.25\sigma$, and leaves the error bars essentially unchanged.

Jarvis et al. (2016) calculated the shape-noise for NGMIX in DES SV and found $\sigma_\varepsilon = 0.243$. Fig. 1 compares the Jackknife errors for the background g_+ with Gaussian errors assuming this value for shape-noise. Jackknife errors are larger on average by 26 per cent for NGMIX and 8 per cent for IM3SHAPE, indicating that systematic errors are subdominant.

Jackknife covariance matrices are often underestimated if there are too few independent samples available, and we therefore apply a correction that depends on both the number of bins and the galaxies per bin (Hartlap, Simon & Schneider 2007). This typically increases our errors by only a few per cent.

2.3 SPT observations

The SPT is a 10 m telescope located at the NSF South Pole research station. From 2007 to 2011, the telescope was configured to observe in three mm-wave bands (centred at 95, 150, and 220 GHz). The majority of this period was spent on a survey of a contiguous 2500 deg^2 area within the boundaries 20h $\leq$ RA $\leq$ 7h and -65°

[3] https://bitbucket.org/joezuntz/im3shape/

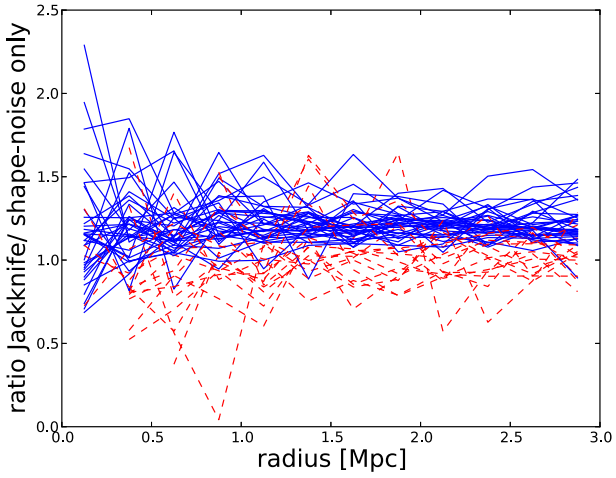

Figure 1. Ratio of Jackknife errors to intrinsic shape noise (taken to be 0.243) for the tangential shear. Each line represents an individual cluster in our sample. NGMIX is shown in solid blue lines, IM3SHAPE in red dashed lines.

$\leq$ Dec. $\leq -40^{\circ}$. In 2011 November, the observations of the whole survey area to the fiducial depth of 18 μK-arcmin in the 150 GHz band were completed. For a detailed description of the survey strategy and data processing, we refer to Staniszewski et al. (2009) (see also Vanderlinde et al. 2010; Williamson et al. 2011; Mocanu et al. 2013). Song et al. (2012) presented optical and near-infrared follow-up of a preliminary catalogue of 720 deg^2, including redshift estimates. The cluster catalogue for the full survey area appeared in Bleem et al. (2015).

Galaxy clusters are detected via their thermal SZE decrement in the 95 and 150 GHz SPT maps. These maps are created using time-ordered data processing and map-making procedures equivalent to those described in Vanderlinde et al. (2010). A multiscale matched-filter approach is used for cluster detection (Melin, Bartlett & Delabrouille 2006), where the underlying cluster model is a β model (Cavaliere & Fusco-Femiano 1976; Cavaliere & Fusco-Femiano 1978) with $\beta = 1$ and a core radius Θ_c. Twelve linearly spaced values from 0.25 to 3.0 arcmin are employed, and the observable used to quantify the cluster SZE signal is ξ, the detection significance maximized over this range of core radii.

In total, 677 cluster candidates above an SNR limit of 4.5 are detected in the full SPT-SZ survey and 516 are confirmed by optical and near-infrared imaging. This number includes 415 systems first identified with the SPT and 141 systems with spectroscopic redshift information. The median mass of this sample is $M_{500,c} \approx 3.5 \times 10^{14}$ M$_\odot$ and the median redshift 0.55. The highest redshift exceeds 1.4 (Bleem et al. 2015).

2.4 SZE-selected lens sample

The SPT-SZ catalogue has an overlap of about 100 clusters and candidates with $\xi > 4.5$ over the full DES SV area, including areas that did not survive survey quality cuts in the southern part of the SPT-E field. Shear catalogues for the SPT-W field are not available at the time of this work. Some cluster candidates have not been confirmed and hence do not have a redshift estimate and are therefore excluded from this analysis.

We restrict ourselves to clusters with redshift $0.25 < z \leq 0.8$. At DES depths, higher redshift clusters suffer from very low number densities of lensing source galaxies and small lensing efficiency, resulting in poorly measured, noise-dominated profiles even for the most massive systems like El Gordo (Melchior et al. 2015). Also, complementary work with space-based *HST* observations (e.g. Schrabback et al. 2018) is focused on providing WL-based mass estimates for systems in this redshift range. At lower redshifts the SPT selection function is not well characterized and inclusion of clusters at $z < 0.25$ could bias our estimates of the scaling relation parameters.

This leaves us with 35 clusters with $\xi > 4.5$ covered in the DES SV area. We remove SPT-CL J2242−4435 and SPT-CL J0451−4952 from our lens sample because of very low source number densities after cuts.

The remaining 33 clusters used in this analysis are listed in Table 1, including their sky position, detection significance, core radius Θ_c, and redshift. If possible we use spectroscopic redshifts (denoted by (s)). Cluster SZE-based masses $M_{500,SZ}$ are taken from Bleem et al. (2015) and have been derived assuming a flat ΛCDM cosmology with $\Omega_m = 0.3$, $\sigma_8 = 0.8$, and $h = 0.7$ and a fixed mass–observable relation with an intrinsic scatter $D_{SZ} = 0.22$. These values are informational only and are not used when deriving our scaling relation constraints.

An additional column shows the DES SV field. Most clusters are located in the SPT-E field. Several systems are in targeted cluster fields, though El Gordo is at too high a redshift to be included in our lens sample. Two systems are in one of the Supernova fields (SNE), which are deeper than the main survey. DES imaging allows optical confirmation and redshift estimates of our clusters independently of other optical follow-up observations. Hennig et al. (2017) identified the red sequences for SPT clusters in the SV footprint and derived comparable redshifts to those presented in Bleem et al. (2015) over the full redshift range. For consistency with other publications using the same SPT-SZ catalogue we use the redshift estimates from Bleem et al. (2015) whenever possible. This is the case for almost the full sample, except for three clusters at lower SNR, where we employ redshift estimates and SZE-based masses from Saro et al. (2015).

Fig. 2 shows the distribution of our sample in redshift-ξ space. The sample spans the full redshift range from 0.25 to 0.8, with the majority having significance values close to the catalogue threshold. Clusters with spectroscopic redshift information are shown as red diamonds. The most significant SPT cluster detections in our sample are in the range $0.3 < z < 0.4$, including the Bullet cluster (SPT-CL J0658–5556) and RXJ2248, which have been previously studied with DES data (Melchior et al. 2015).

Saro et al. (2015) matched SPT clusters and candidates down to $\xi = 4$ to clusters identified by the optical cluster finder redMaPPer (Rykoff et al. 2014) in the DES SV area, thereby confirming five candidates above $\xi = 4.5$ and presenting redshift estimates for these systems based on their redMaPPer counterpart. We include three systems that remain after applying the SPT point source mask into our sample. Bleem et al. (2015) have estimated the number of false detections for $\xi < 4.5$ clusters to increase from < 10 per cent at $\xi = 4.5$ to $\approx$40 per cent at $\xi = 4$. For the scaling relation analysis, we therefore use only SPT clusters above $\xi = 4.5$.

3 CLUSTER SHEAR PROFILES

In this section, we first describe how we select the background galaxy population that is needed to construct the observed shear profiles. We then explore in Section 3.2 whether the background population we have selected is contaminated by cluster galaxies. Thereafter, we describe the theoretical profile we adopt in Sec-

Table 1. Lens sample used. From left, we list the name, sky position, SZE significance, detection scale θ_c, SZE $M_{500,SZ}$, redshift (where '(s)' denotes spectroscopic redshift), DES Field (SNE = ELAIS supernova field), and whether NGMIX catalogue is available.

SPT ID	RA (deg)	Dec. (deg)	ξ	Θ_c (arcmin)	$M_{500,\,SZ}$ ($10^{14}\,M_\odot$)	Redshift	DES field	NGMIX
SPT-CL J0040–4407	10.2048	−44.1329	19.34	0.50	10.24 ± 1.56	0.350(s)	SNE	
SPT-CL J0041–4428	10.2513	−44.4785	8.84	0.50	5.83 ± 1.01	0.33 ± 0.02	SNE	
SPT-CL J0107–4855	16.8857	−48.9171	4.51	0.25	2.48 ± 0.81	0.60 ± 0.03	El Gordo	
SPT-CL J0412–5106	63.2297	−51.1098	5.15	0.25	3.42 ± 0.84	0.28 ± 0.04	SPT-E	✓
SPT-CL J0417–4748	64.3451	−47.8139	14.24	0.25	7.41 ± 1.15	0.581(s)	SPT-E	✓
SPT-CL J0422–4608	65.7490	−46.1436	5.05	0.50	2.90 ± 0.75	0.70 ± 0.03	SPT-E	✓
SPT-CL J0422–5140	65.5923	−51.6755	5.86	1.00	3.57 ± 0.77	0.59 ± 0.03	SPT-E	✓
SPT-CL J0426–5455	66.5199	−54.9197	8.85	0.50	5.17 ± 0.90	0.63 ± 0.03	SPT-E	✓
SPT-CL J0428–6049	67.0305	−60.8292	5.11	1.25	3.04 ± 0.78	0.64 ± 0.03	SPT-E	✓
SPT-CL J0429–5233	67.4315	−52.5609	4.56	0.75	2.75 ± 0.77	0.53 ± 0.03	SPT-E	✓
SPT-CL J0433–5630	68.2541	−56.5025	5.32	1.75	3.13 ± 0.76	0.692(s)	SPT-E	✓
SPT-CL J0437–5307	69.2599	−53.1206	4.52	0.25	3.20 ± 0.80†	0.29 ± 0.02[a]	SPT-E	✓
SPT-CL J0438–5419	69.5749	−54.3212	22.88	0.50	10.80 ± 1.62	0.421(s)	SPT-E	✓
SPT-CL J0439–4600	69.8087	−46.0142	8.28	0.25	5.29 ± 0.94	0.34 ± 0.04	SPT-E	✓
SPT-CL J0439–5330	69.9290	−53.5038	5.61	0.75	3.59 ± 0.80	0.43 ± 0.04	SPT-E	✓
SPT-CL J0440–4657	70.2307	−46.9654	7.13	1.25	4.63 ± 0.89	0.35 ± 0.04	SPT-E	✓
SPT-CL J0441–4855	70.4511	−48.9190	8.56	0.50	4.74 ± 0.83	0.79 ± 0.04	SPT-E	✓
SPT-CL J0444–4352	71.1683	−43.8735	5.01	1.50	3.11 ± 0.82	0.57 ± 0.03	SPT-E	✓
SPT-CL J0447–5055	71.8445	−50.9227	5.96	0.25	3.87 ± 0.82	0.39 ± 0.05	SPT-E	✓
SPT-CL J0449–4901	72.2742	−49.0246	8.91	0.50	4.90 ± 0.85	0.792(s)	SPT-E	✓
SPT-CL J0452–4806	73.0034	−48.1102	4.52	0.50	2.87 ± 0.81	0.37 ± 0.04	SPT-E	✓
SPT-CL J0456–5623	74.1753	−56.3855	4.60	0.25	2.68 ± 0.75	0.66 ± 0.03	SPT-E	✓
SPT-CL J0500–4551	75.2108	−45.8564	4.51	0.75	3.60 ± 0.91[a]	0.26 ± 0.01[a]	SPT-E	✓
SPT-CL J0502–6048	75.7240	−60.810	4.69	0.25	3.03 ± 0.76[a]	0.79 ± 0.02[a]	SPT-E	✓
SPT-CL J0509–5342	77.3374	−53.7053	8.50	0.75	5.06 ± 0.89	0.461(s)	SPT-E	✓
SPT-CL J0516–5430	79.1513	−54.5108	12.41	1.50	7.10 ± 1.14	0.295(s)	SPT-E	✓
SPT-CL J0529–6051	82.3493	−60.8578	5.58	0.50	3.39 ± 0.78	0.72 ± 0.06	SPT-E	✓
SPT-CL J0534–5937	83.6082	−59.6257	4.74	0.25	2.75 ± 0.75	0.576(s)	SPT-E	✓
SPT-CL J0540–5744	85.0043	−57.7405	6.74	0.25	3.76 ± 0.74	0.76 ± 0.03	SPT-E	✓
SPT-CL J0655–5541	103.9137	−55.6931	5.64	1.00	3.98 ± 0.88	0.29 ± 0.04	Bullet	
SPT-CL J0658–5556	104.6317	−55.9465	39.05	1.25	16.86 ± 2.49	0.296(s)	Bullet	
SPT-CL J2248–4431	342.1907	−44.5269	42.36	0.75	17.27 ± 2.54	0.351(s)	RXJ2248	
SPT-CL J2249–4442	342.4069	−44.7158	5.11	0.25	3.18 ± 0.81	0.60 ± 0.03	RXJ2248	

[a]Marks clusters presented in Saro et al. (2015).

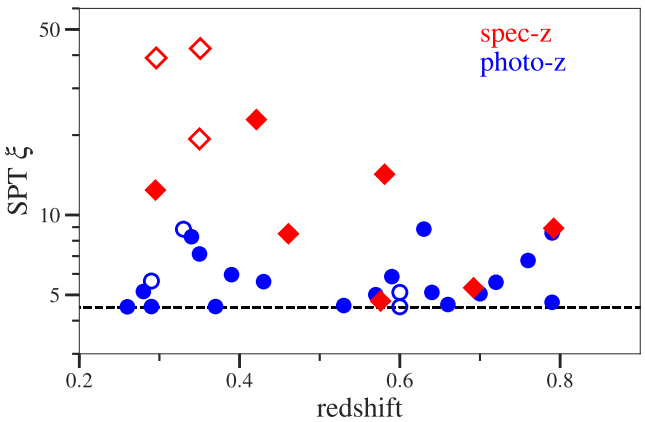

Figure 2. Our sample from the SPT-SZ catalogue (Bleem et al. 2015). Plotted is the SPT significance ξ versus redshift. Clusters with spectroscopic redshifts are shown as red diamonds, those with only photometric redshifts as blue circles. The dashed horizontal line corresponds to the $\xi = 4.5$ limit of the catalogue. Clusters covered by both shear catalogues used in this work are shown as filled symbols, those that only have shape information from IM3SHAPE catalogues as empty ones. As expected, most clusters lie near the catalogue threshold, but the full sample spans a broad range in ξ.

tion 3.3, discuss the radial ranges and binning for the shear profiles in Section 3.4, and then describe the framework we introduce to account for biases and scatter in our WL mass estimates (Section 3.5).

3.1 Background source selection

Background selection by reliable photometric redshifts has been shown to perform better than colour-cuts if enough bands are available (e.g. Applegate et al. 2014). We therefore use photometric redshifts from *griz* bands (Bonnett et al. 2016) to calculate the critical surface density

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l}D_{\rm ls}} \propto \frac{1}{D_{\rm l}\beta}, \qquad (4)$$

where c is the speed of light, G is the (Newtonian) gravitational constant, and $D_{\rm l}$, $D_{\rm s}$, and $D_{\rm ls}$ denote the angular diameter distances from the observer to the lens and the source, and from the lens to the source, respectively. $\beta = D_{\rm ls}/D_{\rm s}$ is the lensing efficiency.

We are using training-set based photo-z estimates that have been shown to perform better than template-based alternatives in the case of DES data (Sánchez et al. 2014). In particular, we match our shear catalogues to SKYNET photometric redshifts (Graff et al. 2014; Bonnett 2015; Bonnett et al. 2016). SKYNET is a training

set based photo-z code that gives both a point estimator (the mean or the peak of the distribution) and a full $P(z)$ distribution using prediction trees and random forests. The training and validation sets use 28 219 and 14 317 galaxies, respectively, with measured spectra in the DES SV footprint extending to $z = 2$. Because these galaxies typically have deeper photometry than SPT-E, they were assigned new photometric errors that were taken from objects in the SPT-E field that are closest in a 5-d colour–magnitude space. The $P(z)$ values are tabulated for 200 values from 0 to 1.8 and normalized to unity. The typical redshift error for SKYNET when applied to DES SV data is $\delta z = 0.08$ (1σ) for both point estimator and $P(z)$. We choose to select our background sample by requiring that

$$z_s > z_{cl} + 0.2 \tag{5}$$

holds simultaneously for both the mean and the peak of the $P(z)$ distribution. We use the former as a proxy for the source redshift z_s. The impact of this error for the estimation of Σ_{crit} is described below. We construct an $N(z)$ distribution for the source sample of each cluster. If contamination by cluster members can be neglected, $N(z)$ should not depend on cluster-centric distance. β is then estimated from $N(z)$ in our fitting routine for the scaling relation. This allows us to treat the dependence of β on cosmological parameters in a self-consistent way. We explore the stability of our estimation of the lensing efficiency in Section 5.4 when using a different photometric redshift catalogue.

3.2 Cluster member contamination

Because photometric redshifts are in general noisy, cluster galaxies may scatter into the background sample. Cluster galaxies would show no shear signal from the cluster, and therefore this contamination would lead to an overall dilution of the mean shear profile and a subsequent underestimation of cluster mass. This effect can be seen as an increase in the number density of sources close to the cluster centre. The radial dependence of the number density profile is also affected by magnification and the obscuration of the sky by bright foreground objects. Masking of, e.g. bright stars (including the 2MASS catalogue), image artefacts or because of survey edges also must be taken into account to derive correct number densities. Noting that magnification only contributes significantly in the very inner regions (Chiu et al. 2016), which we neglect in our shear analysis, we leave this effect uncorrected [but see Schrabback et al. (2018) for an investigation of its potentially larger impact for clusters at higher redshift].

3.2.1 *Radial trend in background density*

To estimate a correction for the contamination, we first assume that the contamination by cluster galaxies decreases with increasing distance $r/r_{500,SZ}$ from the cluster centre, where the scale radius is set by the cluster mass (as given in Bleem et al. 2015). Following Applegate et al. (2014), we model the effects of the contamination on the background number density as

$$n_{corr}(r) = n_0 \times \left(1 + f_{500} \exp\left[1 - r/r_{500,SZ}\right]\right), \tag{6}$$

where n_0 denotes the uncontaminated background number density that is a constant and f_{500} is the contamination fraction at a cluster-centric distance $r_{500,SZ}$. We perform a simultaneous fit for a global f_{500} and a different n_0 for each cluster.

Fig. 3 shows the average number density profile of our NGMIX sources as a function of cluster centric distance, including splits

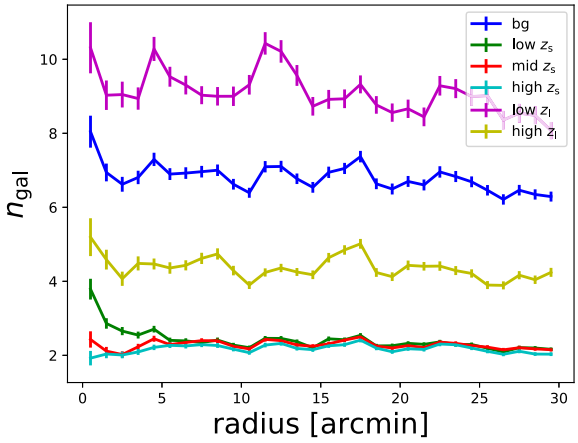

Figure 3. Number density profile of our source galaxy sample from NGMIX as a function of cluster-centric angular distance. The full sample is shown in blue, and three different slices in increasing source redshift are visible in green, red, and cyan. The full source samples for low-z and high-z lenses are shown in magenta and yellow. This figure is for illustration only, because the contamination is evaluated for individual clusters rather than the stack shown above. The error bars are the Poisson errors of the number counts.

Table 2. Cluster member contamination constraints (evaluated at $r_{500,SZ}$) extracted from the various subsamples.

Subsample	NGMIX (per cent)	IM3SHAPE (per cent)
Full bg	8.1 ± 6.9	9.3 ± 7.0
Low z_s	12.1 ± 6.9	9.5 ± 7.6
Mid z_s	2.6 ± 7.6	1.9 ± 6.6
High z_s	0.9 ± 7.9	2.7 ± 8.4
Low z_l	8.1 ± 6.9	10.7 ± 8.9
High z_l	4.6 ± 13.5	1.4 ± 15.0

in source and lens redshift. Table 2 summarizes our estimates of contamination. We find a value of $f_{500} = (8.1 \pm 6.9)$ per cent for the full sample of NGMIX sources and lenses (in blue), very close to no contamination, and (9.3 ± 7.0) per cent for the IM3SHAPE sources. Without redshift selection (equation 5), we would get (11.3 ± 2.1) per cent.

Splitting the sources for each cluster into three equally populated source redshift bins (green, red, cyan) shows a lot of fluctuation but no significant contamination for any bin. Splitting the cluster sample at the median lens redshift also gives values of f_{500} consistent with zero (magenta and yellow lines) at the 1.2σ level.

Additionally, a small f_{500} would not affect our conclusions, given the large statistical uncertainties in our current analysis. Therefore, we choose not to correct the tangential shear signal. Indeed, no significant cluster contamination is expected, because we use photometric redshifts and a background selection that corresponds to $\approx 2.5 \times \delta z$ above the cluster redshift.

3.2.2 *P(z) decomposition*

As a cross-check for our contamination correction we use an adaptation of the method described in Gruen et al. (2014) in the case of individual source redshift distributions. Because this method does not use number densities from our source catalogue, it is subject to different systematics.

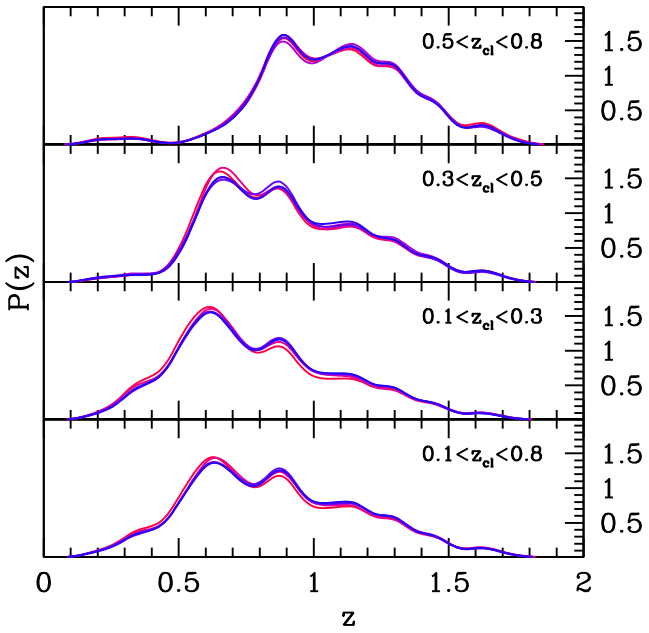

Figure 4. P(z) distribution of NGMIX sources. We split the source population into five radial bins, ranging from red (innermost) to blue (outermost). The bottom panel shows the stack of all clusters, and the three top panels show slices in lens redshift. We estimate an overall contamination of (3 ± 1) per cent in the two inner bins (see discussion in Section 3.2.2).

We summarize this method briefly and refer the interested reader to the original paper (Gruen et al. 2014) and a study of the stacked WL signal from redMaPPer clusters in DES SV data (Melchior et al. 2017) for its adaptation to DES $P(z)$'s. The source galaxy redshift distribution is modelled with two components: a spatially constant background and a radially varying contaminant of cluster galaxies. Comparing the $P(z)$'s in radial bins around the cluster centre with a local background at large separation allows one to infer the level of contamination needed to recover the observed radial change in the $P(z)$ distribution. We choose five equally populated radial bins from 0.75 to 2.5 Mpc and find an overall contamination of 3 ± 1 per cent in the two innermost bins, translating to $f_{500} = (3.8 \pm 1.3)$ per cent.[4] Fig. 4 shows the radial dependence of the $P(z)$ distribution for the full source sample and three slices in lens redshift.

Although both methods give consistent results for the scale of the contamination, the $P(z)$ decomposition approach provides higher significance due its smaller measurement errors. We find in a similar analysis (Dietrich et al. 2019) that this level of f_{500} translates to a $\approx$2 per cent shift in mass, which is about an order of magnitude smaller than our statistical error.

3.3 Assumed cluster profile

Simulations have shown that the profile of a dark matter halo is on average well approximated by a Navarro–Frenk–White (NFW) profile (Navarro, Frenk & White 1997)

$$\rho_{\rm NFW} = \frac{\rho_0}{(r/r_{\rm s})(1 + r/r_{\rm s})^2}, \tag{7}$$

which has two free parameters ρ_0 and $r_{\rm s}$, although more recent work indicate that the Einasto (1965) profile is a better fit for massive clusters (Klypin et al. 2016, and references therein). We will calibrate the impact of deviations from a spherical NFW profile using simulations (cf. Section 3.5).

Because we are interested in the mass $M_{\Delta,{\rm c}}$ residing within a sphere of radius r_Δ with an average overdensity that is Δ times the critical density of the Universe at the cluster redshift, it is convenient to rewrite the NFW profile using $M_{\Delta,{\rm c}}$ and concentration $c_{\Delta,{\rm c}} = r_{\Delta,{\rm c}}/r_{\rm s}$ as a parametrization. For the scaling relation analysis, we use $\Delta = 500$ because this will simplify comparison to previous results.

An analytic expression for the radial dependence of the tangential shear for an NFW density profile has been presented elsewhere (Bartelmann 1996; Wright & Brainerd 2000). We use this result in our weak-lensing analysis. Because our WL data barely constrain the concentration, we adopt a concentration from previously published mass–concentration relations extracted from simulations (Diemer & Kravtsov 2015). We find by comparing to another relation (Duffy et al. 2008) that our results do not depend on this choice (see Section 4.3).

3.4 Radial fitting range and binning

Masses derived from a WL analysis may show per cent level biases depending on both the inner and outer radii of the fit region (Becker & Kravtsov 2011). Excluding the central region suppresses the influence of miscentering, concentration, baryonic effects on the halo profile and a departure from the pure WL regime. On the observational side, deblending, neighbour effects and contamination by cluster galaxies degrade the reported shears for small cluster-centric distances. At large cluster-centric distances, the signal is dominated by the two-halo term and potentially by uncorrelated structures along the line of sight and the profile is not well-described by an NFW profile. To minimize the impact of these biases, we fit in the radial range from 750 kpc to 2.5 Mpc for our reference cosmology, which corresponds roughly to 0.5–2 r_{500} for a halo of mass $M_{500} = 3 \times 10^{14}\,{\rm M}_\odot$.

Because the number of sources after our cuts differs significantly from cluster to cluster (due to depth variations after cleaning and the large span in lens redshifts), we adopt an adaptive binning scheme where we have at least five bins but for background samples larger than 1000 galaxies, we divide the sample by 200 and take the truncated result to be the number of bins. We tested a variety of binning schemes and found that the choice of binning employed does not systematically influence our results.

The input data to our analysis are (i) the cosmology independent tangential shear profiles, (ii) the associated uncertainties as described in equation (3), and (iii) the source redshift distributions $N(z)$ weighted by the shear weight of our source sample. Whenever possible we use the NGMIX shear catalogue, because it has higher number-densities and larger numbers of exposures per object. For nine clusters, mainly outside of SPT-E, we rely on the IM3SHAPE shear catalogue. Table 3 shows the number of galaxies used for our fit and the derived number of bins for both catalogues.

3.5 Calibration of WL mass bias and scatter

In our analysis, we use the cluster centre derived during the SZE detection process as the shear profile centre. The SZE centre scatters about the BCG location (Song et al. 2012) in a manner consistent with the scatter of the X-ray centre about the BCG location (Lin &

[4]We note that a direct decomposition was not possible because the $P(z)$ distribution depends only very weakly on the radius. Instead we looked at differences in the cumulative redshift distribution between radial bins.

Table 3. WL Information for each cluster, where $N_{\rm gal}$ denotes the number of background galaxies used for fitting, and $N_{\rm bin}$ is the number of radial bins. These quantities are shown both for the NGMIX and the IM3SHAPE catalogues. The last column contains the median r-band seeing $\theta_{\rm psf}$ within a 10 arcmin aperture centred on each cluster.

SPT ID	$N_{\rm gal}^{\rm NG}$	$N_{\rm bin}$	$N_{\rm gal}^{\rm im3}$	$N_{\rm bin}$	$\theta_{\rm psf}$ (arcsec)
SPT-CL J0040–4407	–	–	634	5	1.25
SPT-CL J0041–4428	–	–	351	5	1.26
SPT-CL J0107–4855	–	–	200	5	1.15
SPT-CL J0412–5106	2074	10	–	–	1.23
SPT -CL J0417–4748	385	5	–	–	1.18
SPT-CL J0422–4608	266	5	–	–	1.11
SPT-CL J0422–5140	429	5	–	–	1.18
SPT-CL J0426–5455	238	5	–	–	1.30
SPT-CL J0428–6049	518	5	–	–	1.04
SPT-CL J0429–5233	550	5	–	–	1.14
SPT-CL J0433–5630	239	5	–	–	1.24
SPT-CL J0437–5307	2276	11	–	–	1.18
SPT-CL J0438–5419	961	5	–	–	1.29
SPT-CL J0439–4600	1608	8	–	–	1.18
SPT-CL J0439–5330	987	5	–	–	1.22
SPT-CL J0440–4657	2168	10	–	–	1.16
SPT-CL J0441–4855	362	5	–	–	1.14
SPT-CL J0444–4352	408	5	–	–	1.24
SPT-CL J0447–5055	1547	7	–	–	1.19
SPT-CL J0449–4901	420	5	–	–	1.05
SPT-CL J0452–4806	1914	9	–	–	1.10
SPT-CL J0456–5623	420	5	–	–	1.24
SPT-CL J0500–4551	2500	12	–	–	1.20
SPT-CL J0502–6048	336	5	–	–	1.10
SPT-CL J0509–5342	702	5	–	–	1.23
SPT-CL J0516–5430	1541	7	–	–	1.21
SPT-CL J0529–6051	169	5	–	–	1.23
SPT-CL J0534–5937	414	5	–	–	1.28
SPT-CL J0540–5744	174	5	–	–	1.24
SPT-CL J0655–5541	–	–	519	5	1.06
SPT-CL J0658–5556	–	–	691	5	1.06
SPT-CL J2248–4431	–	–	593	5	1.22
SPT-CL J2249–4442	–	–	194	5	1.17

Mohr 2004), once the additional positional uncertainties from the SPT beam are taken into account. Similar results are found in the scatter of the SZE position around the cluster optical centres (Saro et al. 2014). Studies of simulated cluster ensembles show that the offset distribution between the true centre of the cluster potential and the SZE centre behaves similarly to these observations involving the BCG positions (Gupta et al. 2017). Measuring shear profiles around a position that is offset from the true centre of the cluster potential will tend to decrease the shear signal at small radii and hence result in an underestimate of the WL mass. This effect has to be accounted for to obtain accurate cluster masses.

In addition, other effects such as our choice of the projected NFW model and the radial range we use to carry out the fitting also impact the accuracy and precision with which we can estimate the underlying halo mass from the WL mass. In addition, large-scale structure surrounding the cluster could potentially lead to biases in our WL masses, and the unassociated large-scale structure along the line of sight towards the cluster could introduce additional scatter in our measurements.

To allow for the fact that the WL masses $M_{\rm WL}$, we measure would in general be biased and noisy probes of the underlying true cluster mass $M_{\rm true}$ within $r_{500,c}$ that we seek to measure, we introduce a simple linear relationship the WL and true masses

$$M_{\rm WL} = b_{\rm WL} M_{\rm true}, \tag{8}$$

where $b_{\rm WL}$ is a bias parameter. In addition, we add a scatter parameter $\sigma_{\rm WL}$, which quantified the intrinsic scatter of the WL mass at fixed true mass. With these two additional degrees of freedom, we can then include estimates for the characteristic bias and scatter of our WL masses. As described in details in Appendix B, we then use mock observations of simulated galaxy clusters to understand the bias and scatter in the WL mass. Results of this study lead to priors on these two parameters, as justified in Appendix B, that are $b_{\rm WL} = 0.934 \pm 0.04$ and $\sigma_{\rm WL} = 0.25 \pm 0.12$. The uncertainty on the mean bias could be further reduced through studies of larger samples of mock observations, but the level of this 'theoretical' uncertainty on the bias is already much smaller than the uncertainties associated with the shear multiplicative bias, the photometric redshift bias, and the cluster contamination. These biases are listed separately in Table 4 and sum in quadrature to a total uncertainty of 0.18 that is adopted for the uncertainty on the WL bias parameter in Table 5.

Table 4. Systematic mass error budget broken down into contributions from the source redshift distribution β, the multiplicative shear bias m, and the cluster contamination f_{500}. We additionally consider errors due to miscentering, deviations from an NFW profile, as calibrated by simulations and parametrized by $b_{\rm WL}$. References are provided in column 4. The total systematic uncertainty consists of the listed effects added in quadrature.

Systematic	Error (per cent)	ΔM_{500} (per cent)	Reference
β	6.5	9.6	Section A4; Bonnett et al. (2016)
m	10	15	Extrapolated from Jarvis et al. (2016)
f_{500}	6.9	3.4	Section 3.2
$b_{\rm WL}$	4.0	4.0	Section 3.5
Total		18.6	

4 SCALING RELATION ANALYSIS

In this section, we describe the analysis method to derive the scaling relation parameters and present the results. We present the Bayesian framework in Section 4.1, detail the priors in Section 4.2, and then present our results with a comparison to prior work in Sections 4.3 and 4.4.

4.1 Bayesian foreward modelling framework

The freedom to maximize the SPT significance ξ across three parameters (right ascension, declination, and core radius $\Theta_{\rm c}$) in the presence of a noise field will tend to raise the amplitude of the observed peak. That is, the ensemble average of ξ across many noise realizations, $\langle\xi\rangle$, will be boosted by some amount as compared to an unbiased significance ζ, which is measured without these degrees freedom (Vanderlinde et al. 2010). It can be estimated for $\zeta > 2$ by

$$\zeta = \sqrt{\langle\xi\rangle^2 - 3}\,. \tag{9}$$

The unbiased significance ζ can be related to the mass enclosed by a sphere with a mean overdensity of 500 times the critical density

Table 5. ζ–M_{500} scaling relation parameter constraints and priors for three previous SPT publications as well as this analysis (DES-SV WL Shear). Constraints are shown for the four SZE–mass relation parameters and the two WL mass–mass relation. WL results are shown when adopting the mass–concentration relation from Diemer & Kravtsov (2015). Results are shown with and without a prior on B_{SZ}.

Analysis and constraints	A_{SZ}	B_{SZ}	C_{SZ}	D_{SZ}	b_{WL}	σ_{WL}
Bleem et al. (2015) fixed parameters	4.14	1.44	0.59	0.22	–	–
Bocquet et al. (2015) $SPT_{CL} + Y_X + \sigma_v$	$4.7^{+0.8}_{-1.2}$	1.58 ± 0.12	0.91 ± 0.35	0.26 ± 0.10	–	–
+Planck+WP+BAO+SNe	3.2 ± 0.3	1.49 ± 0.11	0.49 ± 0.22	0.26 ± 0.05	–	–
de Haan et al. (2016) $SPT_{CL}+Y_X$	4.8 ± 0.9	1.67 ± 0.08	0.55 ± 0.32	0.20 ± 0.07	–	–
+Planck+WP+BAO	3.5 ± 0.3	1.66 ± 0.06	0.73 ± 0.12	0.20 ± 0.07	–	–
Dietrich et al. (2019)	$5.58^{+0.96}_{-1.46}$	$1.650^{+0.097}_{-0.096}$	$1.27^{+0.47}_{-0.51}$	$0.173^{+0.073}_{-0.052}$	–	–
DES-SV WL shear						
Priors	–	1.67 ± 0.08	0.55 ± 0.32	0.20 ± 0.07	0.94 ± 0.18	0.25 ± 0.12
With B_{SZ} prior	$12.0^{+2.6}_{-6.7}$	$1.65^{+0.08}_{-0.09}$	$0.50^{+0.31}_{-0.30}$	0.20 ± 0.07	$0.94^{+0.17}_{-0.18}$	$0.24^{+0.11}_{-0.12}$
Free B_{SZ}	$10.8^{+2.3}_{-5.2}$	$1.30^{+0.22}_{-0.44}$	0.50 ± 0.32	0.20 ± 0.07	0.94 ± 0.18	$0.24^{+0.10}_{-0.12}$

of the Universe, $M_{500,c}$, by the mass–observable relation

$$\zeta = A_{SZ}\left(\frac{M_{500,c}}{3\times10^{14}M_{\odot}h^{-1}}\right)^{B_{SZ}}\left(\frac{E(z)}{E(0.6)}\right)^{C_{SZ}}, \quad (10)$$

where A_{SZ} is the normalization, B_{SZ} is the mass slope, C_{SZ} is the redshift evolution, and $E(z) = H(z)/H_0$ (Vanderlinde et al. 2010). An additional parameter D_{SZ} describes the intrinsic scatter in ζ which is assumed to be lognormal and constant as a function of mass and redshift.

Power-law scaling relations among cluster observables that exhibit low intrinsic scatter were first discovered in the X-ray (Mohr & Evrard 1997) and immediately interpreted as evidence that observable properties of clusters scale with the underlying cluster halo mass. These scaling relations (observable to observable, observable to mass) were apparent in clusters from hydrodynamical simulations of the time but with the wrong mass trends. It was quickly apparent that the mass trend of ICM-based observables depends on the thermodynamic history of the ICM, which is impacted by feedback from star formation and active galactic nuclei. The existence of these early X-ray scaling relations (see also Mohr, Mathiesen & Evrard 1999) already implied the existence of SZE scaling relations of similar form, although direct observation at that time was not possible.

The first observations of the SZE scaling relations were enabled through the SPT sample and with input from follow-up X-ray observations with *Chandra* (Andersson et al. 2011). Detailed analysis of the expected distribution of scatter and the redshift evolution of the SZE scaling relations have been studied with simulations (see e.g. Gupta et al. 2017). Finally, within the last three cosmological analyses of the SPT-selected sample, the above scaling relation has been adopted and goodness of fit tests have been carried out. To date, starting first with a sample of 100 clusters and moving now to a sample of 400 clusters, there have been no indications of tension between the observations and this underlying scaling relation (Bocquet et al. 2015; de Haan et al. 2016; Bocquet et al. 2018). Future, larger sample will, of course, allow more stringent tests and will likely lead to the need for additional freedom in the functional form (see e.g. a similar analysis framework set-up for the much larger eROSITA cluster sample that does indeed require additional parameters; Grandis et al. 2018).

The parameter values of B_{SZ} and C_{SZ} in the relation above are therefore impacted by the thermodynamic history of the ICM and cannot be predicted with precision even with the latest generations of hydrodynamical simulations. Importantly, to obtain unbiased cosmological results, one must introduce these degrees of freedom in the astrophysical scaling relation and then constrain them with the use of WL masses. This is indeed the goal of our analysis. For reference, self-similar scaling in mass and redshift for the cluster population would correspond to values of approximately $B_{SZ} \approx 1.3$ and $C_{SZ} \approx 0.7$.

To constrain the four parameters in this model, both simulation priors and X-ray and velocity dispersion information for a subset of the SPT clusters have been used. Recent calibration studies (Bocquet et al. 2015; de Haan et al. 2016) simultaneously fitted cosmological parameters to take into account the cosmological dependency of the scaling relation and the observational mass constraints.

To constrain the ζ–M scaling relation given above, we use an extension of the analysis code developed in Bocquet et al. (2015). The observational constraints include (i) the tangential shear profiles for individual clusters and (ii) the redshift distribution $N(z)$ of source galaxies. We choose these two quantities instead of a combined $\Delta\Sigma$ profile, because the latter is cosmology dependent, and we want to isolate all cosmological dependencies when pursuing either a cosmological or mass calibration analysis (e.g. Majumdar & Mohr 2003; Benson et al. 2013b; Bocquet et al. 2015). An example of the shear profile (but in this case stacked over the whole sample) appears in Fig. A3.

We use a Bayesian framework to estimate the likelihood of each cluster in our sample forward modelling from observed cluster detection significance ξ to the probability of finding the observed shear profile

$$P\left(g_{+,i}|N_i(z),\xi_i,z_i,\boldsymbol{p}\right) = \int \mathrm{d}M_{WL}\,P\left(g_{+,i}|N_i(z),M_{WL},z_i,\boldsymbol{p}\right)P\left(M_{WL}|\xi_i,z_i,\boldsymbol{p}\right), \quad (11)$$

where i runs over all clusters in our sample. To be more explicit, the model first computes how probable a given cluster WL mass is, given the observables ξ and z and the model parameters $\boldsymbol{p}$, which include the scaling relation parameters. This is the last factor on the right-hand side of equation (11). We then compute the probability of measuring the tangential shear we have in our data for a given cluster WL mass, cluster redshift, shear galaxy redshift distribution $N(z)$, and model parameters $\boldsymbol{p}$. This is the first factor in the integral in equation (11), and it is computed for each radial bin assuming that the errors on the tangential shear are the standard deviation of a Gaussian distribution. To obtain a single scalar probability distribution $P\left(g_{+,i}|N_i(z),M_{WL},z_i,\boldsymbol{p}\right)$, the probabilities of all radial bins are multiplied.

This forward modelling approach has the benefit that we can naturally deal with clusters whose measured shear profiles would be consistent with zero WL mass or that exhibit negative shear in some radial bins. It is also flexible enough to deal with cluster catalogues in which only a subset of the clusters have follow-up data, such as WL observations, in a self-consistent and unbiased way as long as the selection of follow-up observations is not correlated with SZE properties.

We correct for Eddington bias by weighting by the mass function $P(M_{\rm true}|z)$ (Tinker et al. 2008; Bocquet et al. 2016) when calculating $P\left(M|\xi_i, z_i, \boldsymbol{p}\right)$. This is necessary, because we select clusters by requiring that their SZE detection significance satisfies $\xi > 4.5$, which directly relates to ζ in our scaling relation via equation (9). The logarithm of our full likelihood is then given by

$$\ln \mathcal{L} = \sum_{i=1}^{N_{\rm cluster}} \ln P\left(g_{+,i}|\xi_i, z_i, \boldsymbol{p}\right) + \text{const.} \tag{12}$$

Using the mass function in this way lets us calibrate the mass–observable relation without self-calibration using the information present in cluster number counts. On the one hand, we can straightforwardly extend this likelihood function to include other observables, e.g. X-ray data (Dietrich et al. 2019), and scaling relations. On the other hand, we can include an additional term incorporating cluster abundance (Bocquet et al. 2018). Future cosmological analysis using the framework presented here will make use of cluster number counts to constrain cosmological parameters, self-calibrate mass-observable relations, and concurrently calibrate the normalization and evolution of these scalings.

As already mentioned, we allow for departures between the WL and true masses from either systematics or intrinsic scatter using equation (8) with an (intrinsic) scatter $\sigma_{\rm WL}$. Additionally, $\sigma_{\rm WL}$ and $D_{\rm SZ}$ may be correlated, and so we include a correlation coefficient $\rho_{\text{SZ-WL}}$. In our analysis, we are then simultaneously fitting the following seven parameters: $\boldsymbol{p} = \{A_{\rm SZ}, B_{\rm SZ}, C_{\rm SZ}, D_{\rm SZ}, b_{\rm WL}, \sigma_{\rm WL}, \rho_{\text{SZ-WL}}\}$.

We discard a burn-in phase that corresponds to five times the auto-correlation length and consider our chains converged if the Gelman & Rubin (1992) convergence diagnostic $\hat{R} < 1.1$.

4.2 Priors

In contrast to previous analyses (e.g. Bocquet et al. 2015; de Haan et al. 2016), we adopt a flat prior on $\ln A_{\rm SZ}$ instead of a flat prior on $A_{\rm SZ}$. This is motivated by the linear form of our SZE observable mass relation (equation 10) in log space. The uninformative prior on the intercept of a line is flat and transforming this back to the power-law relation (10) leads to a prior that is proportional to $1/A_{\rm SZ}$. Indeed, our experience confirms that in the limit of lower number densities, i.e. lower SNR, the prior becomes more dominant and a flat prior on $A_{\rm SZ}$ biases the results towards high values. This bias is removed by our choice of prior. Similarly, we adopt an uninformative prior on $B_{\rm SZ}$ proportional to $(1 + B_{\rm SZ}^2)^{-1.5}$. This corresponds to a flat prior on the angle of the line rather than its slope. We refer the interested reader to the original publication (Jaynes 1983) for a more detailed discussion of this choice. We use the following Gaussian priors on the other scaling relation parameters: $B_{\rm SZ} = 1.668 \pm 0.083$, $C_{\rm SZ} = 0.550 \pm 0.315$, and $D_{\rm SZ} = 0.199 \pm 0.069$, which correspond to the $\text{SPT}_{\rm CL}$ constraints presented in the latest SPT cluster cosmology analysis (de Haan et al. 2016). These constraints adopted external priors on H_0 and on $\Omega_{\rm b}$ from big bang nucleosynthesis. Additionally, we assume a flat prior on $\rho_{\text{SZ-WL}} \in [-1, 1]$. These values are listed in Table 5. We also use the prior for $b_{\rm WL}$ derived in Section 3.5.

We note that our framework is set-up to perform a full cosmological analysis. The WL observables are the shear profile (g_+ as a function of angular separation from the cluster centre) and $N(z)$, the redshift distribution of the shear source galaxies. These observables are cosmology independent, as are the SZE SNR and redshift for each cluster. However, the likelihood described in equation (11) includes cosmological dependencies of the cluster distances and the underlying halo mass function. All this is built in so that the likelihood can be employed within a full cosmological analysis context. For the current work, we choose to fix our cosmological parameters to values obtained from *Planck* and leave the full cosmological analysis to separate work that includes a larger sample of SPT-selected clusters and a subsample that have WL information (Bocquet et al. 2018).

4.3 Results for $\zeta - M_{500}$ scaling relation

Fig. 5 shows the fully marginalized and joint parameter posterior distributions from our fit using a recent mass–concentration relation (Diemer & Kravtsov 2015) and an uninformative prior on the mass slope $B_{\rm SZ}$. Parameter priors are shown as black solid lines. The corresponding mean values and the shortest 68 per cent credible region for each parameter are presented in Table 5, along with the priors and literature values from previous SPT studies.

We find $A_{\rm SZ} = 12.0^{+2.6}_{-6.7}$ when using an informative prior on the mass slope $B_{\rm SZ}$. The probability distribution of $\ln A_{\rm SZ}$ is close to Gaussian and there is a tail to high values. The mean is therefore higher than the mode (9.0) of the distribution. The correlation coefficient $\rho_{\text{SZ-WL}}$ is unconstrained by our data. For the remaining parameters we recover the prior values.

Because our sample spans a broad range in observable ($\xi = 4.5$ to $\xi = 42.4$; and therefore mass), we expect to be able to constrain the mass slope $B_{\rm SZ}$. In the next step, we therefore remove the informative prior on $B_{\rm SZ}$, and recover a value of $B_{\rm SZ} = 1.30^{+0.22}_{-0.44}$. This value is in agreement with but somewhat smaller than results from most previous studies (Bocquet et al. 2015; de Haan et al. 2016). Additionally, the normalization shifts down: $A_{\rm SZ} = 10.8^{+2.3}_{-5.2}$. This small shift may be caused by a degeneracy between the parameters, which also explains why the marginalized uncertainties of $A_{\rm SZ}$ decrease. The total posterior volume increases but rotates in a way that decreases the marginalized uncertainty on $A_{\rm SZ}$ at the expense of increased uncertainty on $B_{\rm SZ}$. The other parameters and parameter uncertainties are essentially unchanged in comparison to the run with the $B_{\rm SZ}$ prior.

We expect significantly tighter constraints on both $A_{\rm SZ}$ and $B_{\rm SZ}$ with the analysis of the full SPT cluster sample with the DES main survey data. Better knowledge of the redshift evolution of the SPT mass–observable relation requires combination with deeper (space-based) data (e.g. Schrabback et al. 2018).

4.4 Comparison to previous results

We now compare our results to SPT mass calibrations in the literature (see Table 5), simulations and abundance-matching based masses. Additionally, we compare to the shear and magnification studies of smaller SPT-selected samples presented in Gruen et al. (2014) and Chiu et al. (2016), respectively.

On the simulation side, the dark matter-only simulations of Vanderlinde et al. (2010) ($A_{\rm SZ} = 6.01$, $B_{\rm SZ} = 1.31$) and the recent hydrodynamical Cosmo-OWLs simulations ($A_{\rm SZ} = 5.38$, $B_{\rm SZ} =$

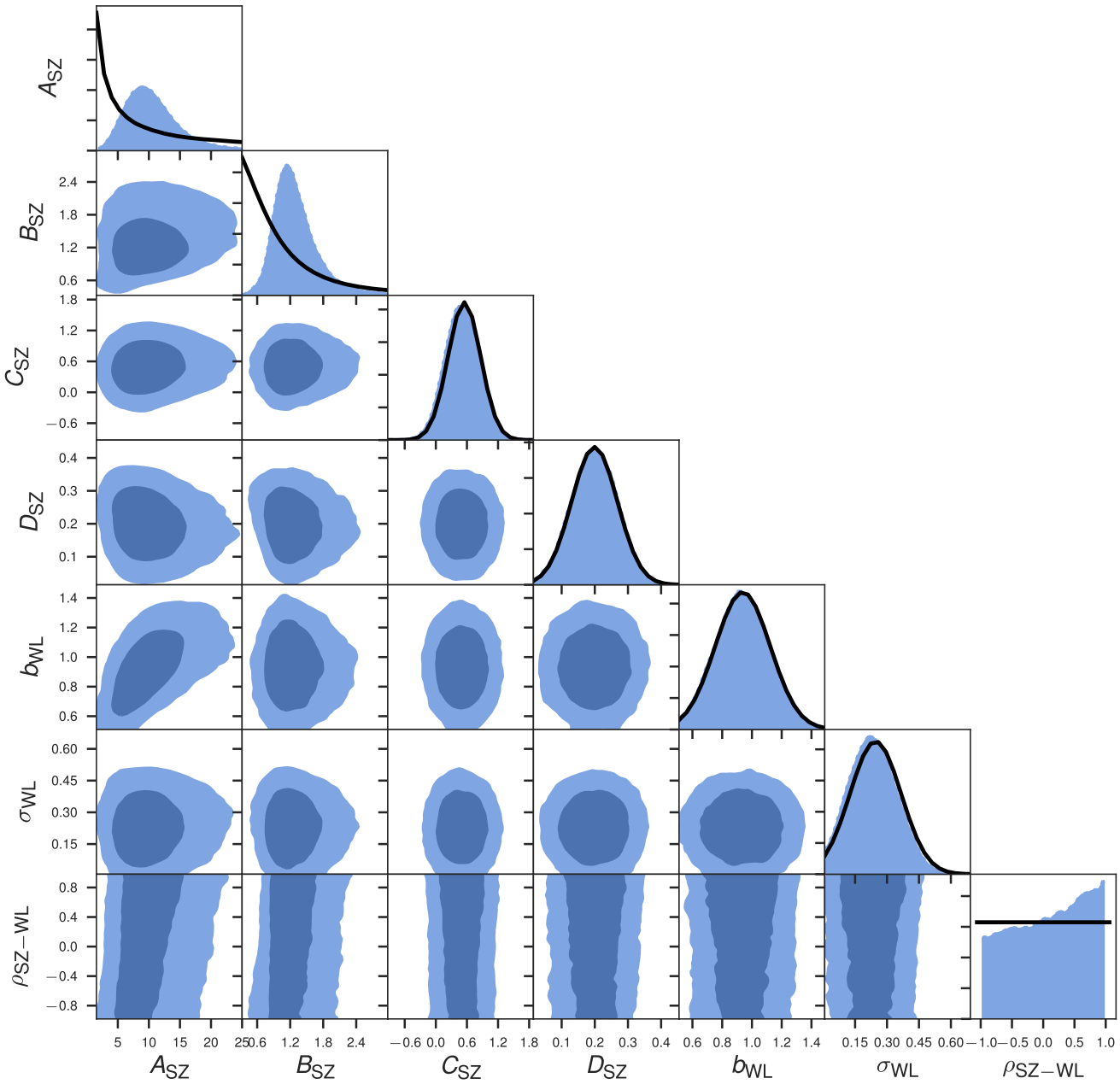

Figure 5. Scaling relation parameter constraints derived from our DES-SV WL analysis. Priors (see Section 4.2) are shown with black lines, where priors for C_{SZ} and D_{SZ} come from de Haan et al. (2016), and b_{WL} and σ_{WL} arise from our analysis of simulations. Parameters A_{SZ}, B_{SZ}, and $\rho_{SZ\text{-}WL}$ are given broad, uninformative priors and are thus constrained only by WL data. For A_{SZ}, we find higher values than expected, though still consistent with most previous analyses. Our data prefer an approximately self-similar value for B_{SZ}, although the uncertainties are large. The data provide no evidence for a correlation between the intrinsic scatter in the SZE-mass and WL-mass scaling relations.

1.34; Le Brun et al. 2014) agree with each other, and are also in agreement with our results given the larger error bars. The smaller value of B_{SZ} found when leaving this parameter free is also favoured by Le Brun et al. (2014).

Our measurement is consistent with the clusters-only constraints presented in the latest SPT cosmology analyses (Bocquet et al. 2015; de Haan et al. 2016). These studies employ a joint mass calibration and cosmology analysis using mass calibration information from velocity dispersions and X-ray mass proxies. The agreement improves when B_{SZ} is left free, though the shift in this parameter from the result of de Haan et al. (2016) used as our prior is a promising target for further investigation. Bocquet et al. (2015) and previous SPT studies recovered a slightly larger value of D_{SZ}, which is anticorrelated with A_{SZ}. We therefore attribute a part of the shift in A_{SZ} to the use of an updated prior on D_{SZ}. When including external cosmological priors in a joint mass calibration and cosmological analysis, the external priors – especially those from CMB measurements – dominate the cluster mass-scale normalization parameter A_{SZ} (Bocquet et al. 2015). This can be seen in the clear shifts of A_{SZ} to values below 4, implying masses that are significantly higher than those from this analysis.

Our normalization of the mass–ζ relation is also consistent with the recent weak-lensing calibration of the SPT cluster sample using pointed follow-up observations (Dietrich et al. 2019). This approach by the SPTteam is similar to ours; it uses the same calibration on *N*-body simulations and a more recent version of the scaling relations software employed with a more sophisticated model for the various sources of weak-lensing scatter. The data sets and the shear catalogue creation are, however, completely independent. While the direct weak-lensing mass calibration of the SPT cluster sample of Dietrich et al. (2019) is more in line with expectations from simulations, velocity-dispersion-based mass calibration, and self-calibration of the mass–ζ relation, the disagreement with our result is not significant, as we will discuss below.

In a previous WL shear analysis of five SPT-selected clusters, the WL- and SZE-based masses were compared (High et al. 2012). The mass estimates were in good agreement, with $\langle M_{\rm WL}/M_{\rm SZ}\rangle = 1.07 \pm 0.18$. We note that the employed SZE masses were from an earlier SPT cluster cosmology analysis (Reichardt et al. 2013) and that they were on average about 35 per cent lower than the masses reported in the analysis of the full SPT-SZ sample (de Haan et al. 2016).

Gruen et al. (2014) used WL shear to analyse a sample of SZE-detected clusters, five of which are selected by SPT. The authors constrain the normalization and slope, $A_{\rm SZ} = 6.0^{+1.9}_{-1.8}(7.6^{+3.0}_{-2.6})$ and $B_{\rm SZ} = 1.25^{+0.36}_{-0.28}(1.02^{+0.62}_{-0.68}$ for a single-halo fit (multi-halo fit, incl. neighbours), when fixing $C_{\rm SZ} = 0.83$. These values are in agreement with our work, with $C_{\rm SZ}$ about 1σ above the prior and the reported value of de Haan et al. (2016).

Fig. 6 illustrates the difference in mass for a hypothetical cluster with detection significance $\xi = 6.5$ at redshift $z = 0.6$. The posterior probabilities for the cluster mass are calculated by inverting the scaling relation and including the constraints on $A_{\rm SZ}$ and $B_{\rm SZ}$. Because this cluster is at the pivot redshift for our scaling relation, this comparison is insensitive to any difference in redshift evolution. We neglect the effect of the intrinsic scatter $D_{\rm SZ}$ in this plot, making no correction for the Eddington bias. The bias correction would be very similar for all sets of constraints we present and thus would have little impact on the relative differences presented here.

Following Bocquet et al. (2015, their Section 5.2), we now calculate the significance of the differences between our baseline measurement and results from previous studies. We draw samples from the $P(M)$ for our example cluster, and calculate the distances δ of pairs of sampled points. We estimate P_δ and integrate over the part of the distribution with $P_\delta < P_\delta(0)$. In the last step, we convert this probability into a significance assuming a normal distribution. Our result is consistent with Gruen et al. (2014) at the 0.8σ level, as well as with previous SPT mass calibrations (Bocquet et al. 2015; de Haan et al. 2016) when only clusters are considered (1.7σ), and the WL calibration of the SPT cluster sample (1.4σ, Dietrich et al. 2019). If one considers the results from Bocquet et al. (2015) including additional primary CMB constraints from *Planck*, there is tension at the 2.6σ level.

The use of a larger sample of SPT-selected galaxy clusters with DES data will improve constraints on both $A_{\rm SZ}$ and $B_{\rm SZ}$. At the same time, a complimentary analysis using number count magnification may allow us to further test the stability of WL-based mass estimates and our control of systematics. For example, a first magnification study of 19 SPT-selected clusters with Megacam data (three clusters overlap with this work) presented in Chiu et al. (2016) found a ratio of WL masses to SZE masses of 0.83 ± 0.24, in statistical agreement with both earlier SPT work and this analysis.

Interestingly, WL-derived mass estimates for SPT-selected clusters prefer lower values than those from non-WL calibrations, although this preference is not statistically significant given the currently large uncertainties. Given that this is true both for magnification and for shear studies from different WL observations, this likely cannot be explained by unknown systematics alone. Larger sample sizes of ongoing WL campaigns are needed to further explore this issue.

If a statistically significant tension between cluster masses calibrated with and without including *Planck* CMB, baryonic acoustic oscillations (BAO), and supernovae (SNe) data persists, it may be a hint for new physics. Tensions between CMB anisotropy constraints and constraints from growth-based probes in the context of a ΛCDM framework can be ameliorated by massive neutrinos or theories of modified gravity. At the same time, possible unknown systematics have to be controlled. We note the disagreement between Planck CMB and cluster cosmology constraints, which may be ameliorated by larger cluster masses (implying a larger bias in their hydrostatic mass estimates). Additionally, the recovered mass from WL also depends on cosmological parameters (especially h and Ω_M). Because the last effect is relatively weak and smaller than the typical precision of current and past analyses, we neglected it in this discussion.

In contrast, a calibration of the optical richness–mass relation through a stacked WL analysis shows good agreement between WL and SZE calibration (Saro et al. 2015; Melchior et al. 2017).

5 CONCLUSIONS

In this work, we use shear and photo-z catalogues obtained from SV data taken prior to the start of DES to constrain the masses of SPT SZE-selected clusters of galaxies. The DES catalogues span 139 deg^2 after masking and cuts and overlap with 33 SPT-selected galaxy clusters above an SPT SZE significance $\xi > 4.5$ and redshifts extending to $z_{\rm l} = 0.8$.

We first use photo-z's to select the background source galaxies for our WL study, and then perform a number of cluster-lensing specific tests to further validate our catalogues. These include examining the shear profiles for the cross-component, demonstrating that these profiles are consistent with the expected null signal. We also probe for contamination from cluster galaxies, using two independent methods to show that there is no measurable contamination. We examine the dependence of the implied surface mass overdensity as a function of source properties such as redshift and size, showing good consistency among all subsamples tested. We demonstrate good agreement between the two shape catalogues derived using NGMIX and IM3SHAPE, though the latter shows lower source number densities, because it was applied only to r-band images.

We then use these validated catalogues to carry out a joint fit of the SZE mass observable relation, which is described by four parameters (equation 10). In this process, we characterize systematic biases and intrinsic scatter in WL mass estimates by applying our mass profile fitting and mass estimation to simulated clusters. We incorporate these systematics and scatter in our analysis by introducing a WL mass to true mass scaling relation with a free proportionality constant and lognormal scatter (equation 8).

Due to relatively shallow data compared to deeper, pointed WL observations the uncertainties on the masses of individual clusters are relatively large. The availability of shear profiles for the sample of 33 clusters above $z_{\rm cl} > 0.25$, however, allows one to constrain the $\zeta - M$ relation. For this task, we employed an extension of the

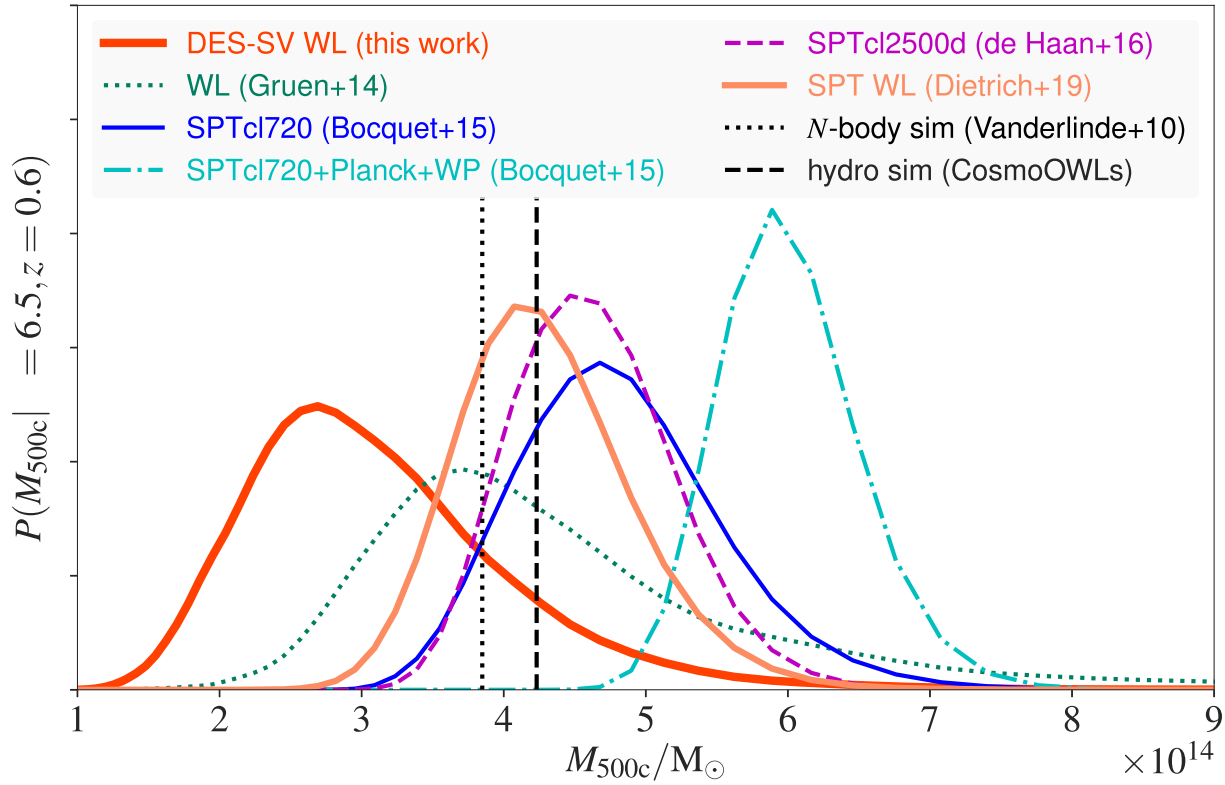

Figure 6. The implications of the best-fitting SZE mass scaling relation expressed as the posterior distribution in mass M_{500} of a typical cluster in the SPT sample ($\xi = 6.5$, $z = 0.6$). The width of the posterior distribution reflects the parameter uncertainties reported in Table 5 and does not include intrinsic scatter or measurement noise on the cluster SZE signature. Shown are comparable constraints for several different studies as detailed in the text (see also Table 5). The vertical lines correspond to the predictions from simulations in Vanderlinde et al. (2010, dotted line) and the cosmo-OWLS simulation (the dashed line Le Brun et al. 2014). As can be seen, previous SPT cosmology analyses recovered higher masses than inferred from the WL calibration in this work. When including external cosmological parameter priors from CMB anisotropy based analyses (Planck Collaboration et al. 2014), even higher masses are preferred.

code developed previously for the cosmological analysis and mass calibration of SPT-selected galaxy clusters (Bocquet et al. 2015). As inputs we use the tangential shear profiles and source redshift distributions, which are direct observables with no cosmological dependence. This approach allows us to self-consistently fit for cosmological and scaling relation parameters.

For convenience, in this initial study we adopt a flat ΛCDM cosmology with $\Omega_m = 0.3089$ and $h = 0.6774$, as motivated by the latest *Planck* cosmology analysis (Planck Collaboration et al. 2016a). We present parameter constraints on the $\zeta - M$ relation normalization and mass slope. Given the large statistical uncertainties in our shear profiles, we do not expect that marginalizing over the allowed cosmological parameter space consistent with the joint *Planck* and external data set analysis would have a significant impact on the scaling relation parameter constraints we derive.

Comparison to earlier SZE mass calibration and cosmology analyses (Bocquet et al. 2015; de Haan et al. 2016; Dietrich et al. 2019) shows that our recovered masses are lower by ≈40 per cent, but still consistent given the large error bars (see Fig. 6). A_{SZ} is insignificantly higher when adopting a prior on the mass slope B_{SZ} from the cosmological analysis of de Haan et al. (2016). When left free, the recovered mass slope is shallower than the posterior from de Haan et al. (2016), preferring values closer to the self-similar expectations for the $\zeta - M$ relation; given the large uncertainties, the two different slopes are consistent at the 1.2σ level.

Our results are in mild tension (at the 2.6σ level) with the higher cluster masses preferred by primary CMB constraints from *Planck* (Bocquet et al. 2015).

The analysis presented in this work has been blinded by multiplying the overall shear by an unknown factor to avoid observer biases. As mentioned before, however, in the process of internal collaboration review some additional tests were requested and have been carried out after unblinding.

This work differs in a number of aspects from the first study of cluster WL using SV data (Melchior et al. 2015). That analyses focused on four very massive clusters in a narrow redshift range (≈0.3–0.4) and used a $\Delta\Sigma$ profile as the only ingredient for fitting cluster masses. They used a different photo-z code, which gave only point estimates, an older implementation of IM3SHAPE was run on the coadd images and Gaussian errors were adopted. It also differs from a stacked cluster lensing analysis as presented in Melchior et al. (2017), because it uses individual shear profiles and a different treatment of systematics.

The main 5-yr DES survey will provide full coverage of the SPT footprint at depths somewhat deeper than the data we have used from the SV area. There are 433 confirmed SPT clusters below our redshift limit of $z_{cl} < 0.8$ that have been imaged by the full survey. A simple scaling with the number of lenses suggests fractional errors of 11.4 per cent and 8.7 per cent on A_{SZ} and B_{SZ} when constraining both parameters simultaneously. To make use of

the improved statistical power, further improvement on controlling systematics (see Table 4) is crucial. This will also impact cluster cosmology which at the moment is limited by our knowledge of the cluster mass scale. To this end, we are proceeding with this broader analysis using the mass calibration method developed for and presented in this paper.

ACKNOWLEDGEMENTS

We acknowledge the support by the DFG Cluster of Excellence Origin and Structure of the Universe, the Transregio programme TR33 'The Dark Universe' and the Ludwig-Maximilians-Universität. The analysis presented in this work benefited from using the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP) located at the Leibniz-Rechenzentrum (LRZ) in Munich. DA and TS acknowledge support from the German Federal Ministry of Economics and Technology (BMWi) provided through DLR under projects 50 OR 1210, 50 OR 1308, 50 OR 1407, and 50 OR 1610. 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. AS is supported by the ERC-StG 'ClustersXCosmo', grant agreement 71676. DR is supported by a NASA Postdoctoral Program Senior Fellowship at NASA's Ames Research Center, administered by the Universities Space Research Association under contract with NASA.

The SPT is supported by the National Science Foundation through grant ANT-0638937. Partial support is also provided by the NSF Physics Frontier Center grant PHY-0114422 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation and the Gordon and Betty Moore Foundation. Galaxy cluster research at Harvard is supported by NSF grant AST-1009012, and research at SAO is supported in part by NSF grants AST-1009649 and MRI-0723073. The McGill group acknowledges funding from the National Sciences and Engineering Research Council of Canada, Canada Research Chairs Program, and the Canadian Institute for Advanced Research.

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

Funding for the DES Projects has been provided by the US Department of Energy, the US 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, 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 e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the DES.

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

Facilities: South Pole Telescope, Cerro Tololo Inter-American Observatory's 4 meter Blanco Telescope

REFERENCES

Abbott T. et al., 2016, Phys. Rev. D, 94, 022001
Abell G. O., 1958, ApJS, 3, 211
Andersson K. et al., 2011, ApJ, 738, 48
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Applegate D. E. et al., 2014, MNRAS, 439, 48
Bartelmann M., 1996, A&A, 313, 697
Battaglia N. et al., 2016, J. Cosmology Astropart. Phys., 8, 013
Becker M. R. et al., 2016, Phys. Rev. D, 94, 022002
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Benson B. A. et al., 2013a, ApJ, 763, 147
Benson B. A. et al., 2013b, ApJ, 763, 147
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bleem L. E. et al., 2015, ApJS, 216, 27
Bocquet S. et al., 2015, ApJ, 799, 214
Bocquet S. et al., 2018, preprint (arXiv:1812.01679)
Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361
Bonnett C. et al., 2016, Phys. Rev. D, 94, 042005
Bonnett C., 2015, MNRAS, 449, 1043
Carlstrom J. E. et al., 2011, PASP, 123, 568
Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137
Cavaliere A., Fusco-Femiano R., 1978, A&A, 70, 677
Chang C. et al., 2015, ApJ, 801, 73
Chiu I. et al., 2016, MNRAS, 457, 3050
de Haan T. et al., 2016, ApJ, 832, 95
de Vaucouleurs G., 1948, Ann. Astrophys., 11, 247
DES Collaboration, 2005, preprint (astro-ph/0510346)
DES Collaborationet al., 2016, MNRAS, 460, 1270
Desai S. et al., 2012, ApJ, 757, 83
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Dietrich J. P. et al., 2019, MNRAS, 483, 2871
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161
Edge A. C., Stewart G. C., Fabian A. C., Arnaud K. A., 1990, MNRAS, 245, 559
Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87
Flaugher B. et al., 2015, AJ, 150, 150
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gelman A., Rubin D. B., 1992, Statist. Sci., 7, 457
Graff P., Feroz F., Hobson M. P., Lasenby A., 2014, MNRAS, 441, 1741
Grandis S., Mohr J. J., Dietrich J. P., Bocquet S., Saro A., Klein M., Paulus M., Capasso R., 2018, preprint (arXiv:1810.10553)
Gruen D. et al., 2014, MNRAS, 442, 1507
Gupta N. et al., 2017, MNRAS, 467, 3737
Haiman Z., Mohr J. J., Holder G. P., 2001, ApJ, 553, 545
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399

Hennig C. et al., 2017, MNRAS, 467, 4015
Henry J. P., Arnaud K. A., 1991, ApJ, 372, 410
High F. W. et al., 2012, ApJ, 758, 68
Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685
Israel H., Reiprich T. H., Erben T., Massey R. J., Sarazin C. L., Schneider P., Vikhlinin A., 2014, A&A, 564, A129
Jarvis M. et al., 2016, MNRAS, 460, 2245
Jaynes E. T., 1983, Papers on Probability, Statistics and Statistical Physics, Springer, Dordrecht
Kacprzak T., Zuntz J., Rowe B., Bridle S., Refregier A., Amara A., Voigt L., Hirsch M., 2012, MNRAS, 427, 2711
Klein J. R., Roodman A., 2005, Ann. Rev. Nucl. Part. Sci., 55, 141
Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340
Komatsu E. et al., 2011, ApJS, 192, 18
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Lin Y.-T., Mohr J. J., 2004, ApJ, 617, 879
Majumdar S., Mohr J. J., 2003, ApJ, 585, 603
Mandelbaum R. et al., 2014, ApJS, 212, 5
Mantz A. B. et al., 2015, MNRAS, 446, 2205
Melchior P. et al., 2015, MNRAS, 449, 2219
Melchior P. et al., 2017, MNRAS, 469, 4899
Melin J.-B., Bartlett J. G., Delabrouille J., 2006, A&A, 459, 341
Miller L. et al., 2013, MNRAS, 429, 2858
Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, MNRAS, 382, 315
Mocanu L. M. et al., 2013, ApJ, 779, 61
Mohr J., Evrard A., 1997, ApJ, 491, 38
Mohr J. J., Mathiesen B., Evrard A. E., 1999, ApJ, 517, 627
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Neilsen E., Annis J., 2014, in Manset N., Forshay P., eds, ASP Conf. Ser. Vol. 485, Astronomical Data Analysis Software and Systems XXIII. Astron. Soc. Pac., San Francisco, p. 77
Planck Collaboration et al., 2014, A&A, 571, A16
Planck Collaboration et al., 2016a, A&A, 594, A13
Planck Collaboration et al., 2016b, A&A, 594, A24
Reichardt C. L. et al., 2013, ApJ, 763, 127
Rowe B., 2010, MNRAS, 404, 350
Rykoff E. S. et al., 2014, ApJ, 785, 104
Sánchez C. et al., 2014, MNRAS, 445, 1482
Saro A. et al., 2014, MNRAS, 440, 2610
Saro A. et al., 2015, MNRAS, 454, 2305
Schrabback T. et al., 2018, MNRAS, 474, 2635
Sheldon E. S., 2014, MNRAS, 444, L25
Song J. et al., 2012, ApJ, 761, 22
Staniszewski Z. et al., 2009, ApJ, 701, 32
Sunyaev R. A., Zeldovich Y. B., 1972, Comments Astrophys. Space Phys., 4, 173
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Vanderlinde K. et al., 2010, ApJ, 722, 1180
Vikhlinin A. et al., 2009, ApJ, 692, 1060
von der Linden A. et al., 2014, MNRAS, 443, 1973
White S. D. M., Efstathiou G., Frenk C. S., 1993, MNRAS, 262, 1023
Williamson R. et al., 2011, ApJ, 738, 139
Wright C. O., Brainerd T. G., 2000, ApJ, 534, 34
Zuntz J., Kacprzak T., Voigt L., Hirsch M., Rowe B., Bridle S., 2013, MNRAS, 434, 1604

APPENDIX A: SHEAR PROFILE TESTS

Both shear catalogues we employ have been subjected to an extensive set of tests described in Jarvis et al. (2016), although as mentioned already in Section 2.2, we have adopted relaxed selection

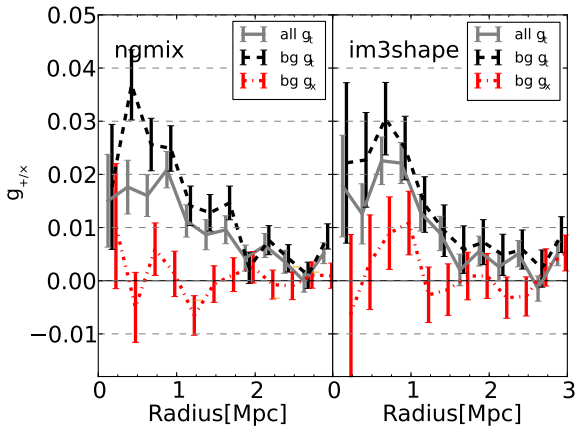

Figure A1. Tangential shear signal for the full cluster stack as a function of angular cluster-centric distance. The shear signal for all source galaxies, and the background galaxies appear in grey and black, respectively, while the cross-shear is in red. The cross-shear is consistent with zero, as expected. The error bars are the diagonal elements of the covariance matrix computed as in equation (3).

criteria that approximately double the surface density of source galaxies. These tests include PSF modelling, ρ statistics (Rowe 2010) and other tests in the context of various WL applications like galaxy-galaxy lensing and cosmic shear. To further validate our shear catalogues within the context of cluster lensing and to justify our inclusion of fainter and smaller objects, we perform a series of additional tests. For better statistics, we stack our full cluster sample in physical units. Because several of these tests involve dividing our sources in a redshift-dependent way, the tangential shear signal generally differs between different subsamples due to different values of β. In those cases, we therefore use the (source redshift independent) surface density contrast $\Delta\Sigma$ defined by

$$\Delta\Sigma = g_+ \times \Sigma_{\rm crit} \tag{A1}$$

instead of g_+ for our profile tests.

In the following subsections, we describe results of the following tests: (1) cross-shear signal, (2) dependence on source SNR, redshift and size, (3) consistency of the two shear catalogues, and (4) stability of β-distribution to choice of redshift code.

A1 Cross-shear signal

Fig. A1 contains a plot of several stacked tangential and the cross-shear profiles, where results for NGMIX are on the left and IM3SHAPE are on the right. Stacked tangential shear profiles are shown for the full and background ($z_{\rm s} < z_{\rm cl} - 0.1$) subsample within 12 linear bins between radii of 0 and 3 Mpc. For each profile, we calculate χ^2 for the null hypothesis of zero shear in the stacked profiles. We clearly detect the tangential shear signal for the background sample, obtaining a $\chi^2 = 167.59$ (73.34) for NGMIX (IM3SHAPE). The cross-shear has a $\chi^2 = 7.43$ (14.17) for NGMIX (IM3SHAPE) with the same binning, which indicates that the data are consistent with the null hypothesis. These measurements confirm the validity of our photo-z catalogue and that the physical origin of the shear signal is indeed our lens sample.

A2 Dependence on source properties

Following Melchior et al. (2015), we investigate the dependence of our shear signal on characteristics of the background source

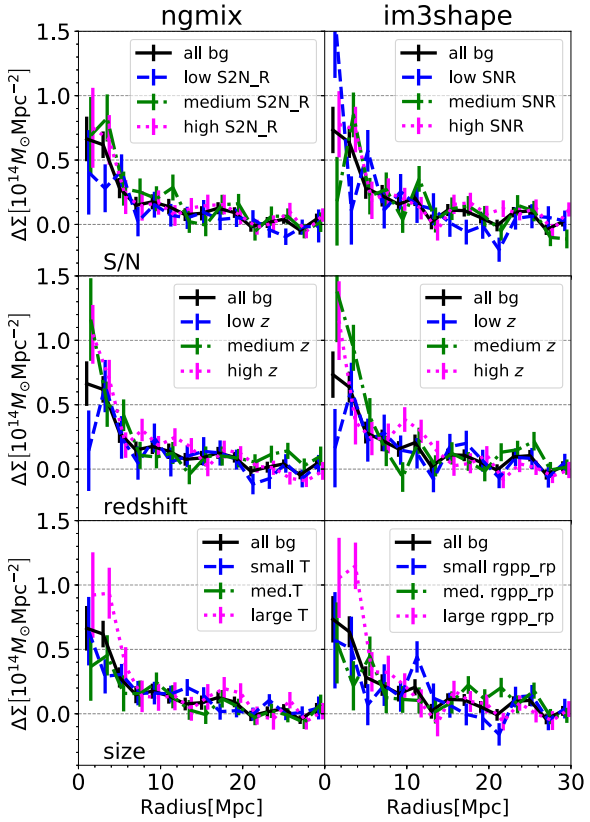

Figure A2. Dependence of measured shear signal on source properties in a stack of 28 $z < 0.8$ clusters for NGMIX and 37 for IM3SHAPE. Each row corresponds to a split in one of the following source properties: SNR, redshift, and size. We split into equally populated bins for each cluster (instead of a fixed boundary for all) to ensure that every source sample sees the same lens sample. Except for the case of the large size sources, there is no significant trend of the signal with the tested source quantities. The error bars are the standard variation of the shear in each radial bin computed via equation (3).

population. Fig. A2 is a plot of the shear profiles (in angular radius from the cluster centre) from source samples subdivided by SNR (top), redshift (centre), and size (bottom). NGMIX catalogues are on the left, and IM3SHAPE on the right. To construct the profiles, we divide the source sample for each cluster into three equally populated bins for each quantity of interest. We take this approach so that each lens contributes the same weight to each of the three source subsamples. Furthermore, this approach allows us to examine the impact of the widened selection criteria applied to the shear catalogues (see Section 2.2). Our selection criteria approximately double the number of sources by including smaller galaxies and galaxies with lower SNR. Consequently, our smallest size and lowest SNR bin contain only sources excluded in the standard cuts; the largest size and highest SNR bins contain only those galaxies included in the standard cuts, and the middle bins are approximately equally populated by both kinds of sources. Any bias caused by the additional sources should then manifest itself as a trend from small/low bins to large/high bins.

The need to adopt cluster dependent subdivisions of the background sample is most easily understood in the case of source subsamples divided by redshift, where clearly the redshift boundaries must shift with the lens redshift. We note that all three investigated quantities are correlated, with high-z sources typically being smaller and at lower SNR.

The visual impression within all panels of Fig. A2 is that all subsamples are in good agreement. To quantify this, we fit masses using the shear profiles of each subsample and then compare the consistency of the mass estimates. For this comparison we fit NFW models to the stacked $\Delta\Sigma$ profiles. We fit $M_{200,\,c}$ using the MCMC sampler `emcee`[5] for PYTHON (Foreman-Mackey et al. 2013) and adopt a flat prior on mass and a lognormal prior on c with $\sigma_{\ln c} = 0.18$. We explicitly allow for negative masses and use the absolute value of the mass in the mass–concentration relation.

We find excellent agreement within the uncertainties for the source redshift and SNR subsamples. The only subsample disagreeing with the stacked signal by more than 1σ is the IM3SHAPE large `RGPP_RP` bin at 1.7σ. Even this subsample, however, agrees at better than 1σ with the small `RGPP_RP` subsample. Furthermore, there is no consistent trend in mass from small to large size. We thus conclude that there are no statistically significant trends in inferred cluster mass with object redshift, SNR, and size, and specifically that the inclusion of additional objects with low SNR and small size does not lead to a detectable bias in cluster mass.

A3 Consistency between shear catalogues

Because there are two independent shear catalogues available for DES SV, a comparison between the tangential shear profile between those two is a valuable cross-check. We perform this test for the stacked subset of our lens sample that has shape catalogues from both pipelines, and only keep common sources after matching both catalogues.

Fig. A3 contains the stacked shear profiles of 28 clusters in the SPT-E field that are covered by both NGMIX and IM3SHAPE catalogues, using the same photometric redshifts. For this plot, we match both shape catalogues and only keep common sources that survive all quality cuts. Additionally, we use the same weights for each galaxy in both catalogues. In contrast to the previously listed tests, this procedure allows us to separate out possible redshift estimation problems and focus directly on the shear measurement. For further discussion of comparison of the NGMIX and IM3SHAPE shear catalogues, we refer the reader to Jarvis et al. (2016).

The cuts employed in this work are less strict than those for analyses that use the full SPT-E footprint. To test the dependence of our result on the details of our cut, we compare stacked tangential shear profiles for NGMIX using both our standard cuts and the most conservative cuts in Jarvis et al. (2016). The profiles are fully consistent in both cases, although the SNR is degraded with the stricter cuts due to lower number density of source galaxies. Therefore, we believe that no additional bias in incurred by the relaxed selection criteria. We adopt a Gaussian prior on the multiplicative bias with a standard deviation of 15 per cent, based on extrapolating the behaviour of m found in Jarvis et al. (2016) to the expanded selection.

A4 Stability of β estimation

We calculate β values from source redshift distributions $N(z)$ for each cluster assuming our standard cosmology (flat ΛCDM, $\Omega_{\rm m} = 0.3089$) for all four different photo-z codes available within DES. Fig. A4 shows the results for our tests. Overall all methods

[5] http://dan.iel.fm/emcee/

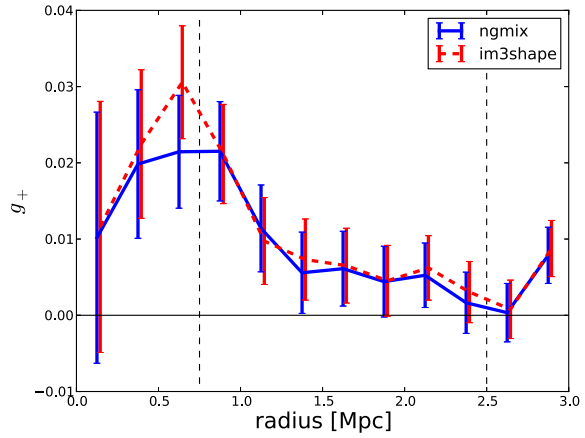

Figure A3. Comparison of the stacked shear signal using IM3SHAPE as a second independent pipeline for clusters with NGMIX coverage. This stack contains 28 clusters from the SPT-E region, and sources are selected as present in both catalogues after all quality cuts. The vertical dashed lines indicate the boundaries of the radial fitting range used for the main analysis. The error bars are the diagonal elements of the covariance matrix computed as in equation (3).

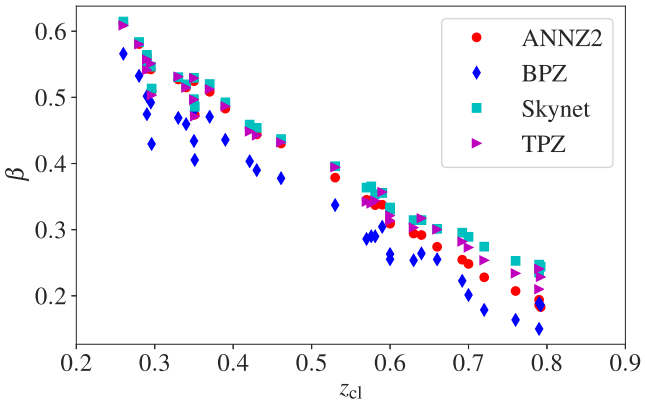

Figure A4. Estimated lensing efficiency β for our background source population for each cluster when employing different redshift codes. The different values are plotted against the cluster redshift $z_{\rm cl}$. The three training-set based methods (ANNZ2, TPZ, and SKYNET show the best agreement, while the template-based BPZ code disagrees most strongly with the fiducial SKYNET algorithm.

show reasonable agreement. Larger discrepancies exist between the template fitting code BPZ and the three training-set based methods. The mean of the β differences is 6.5 per cent when comparing SKYNET and BPZ. We consider this as an estimate of the systematic uncertainties when estimating the lensing efficiencies of the SPT clusters in DES SV data. This uncertainty in β translates into a systematic mass uncertainty of 9.6 per cent, which is what we use as the standard deviation of a Gaussian prior.

To estimate the influence of imperfect knowledge of the lens redshifts, we resample $D_{\rm l}\beta$ 100 times for each cluster from a Gaussian distribution with width equal to the redshift error of our photometric lens sample. We find a mean ratio of 1.010 ± 0.015 relative to taking the centre of the redshift distribution for our full sample. Because this is consistent with unity and negligible compared to the other sources of systematic uncertainty listed in Table 4, we can safely ignore uncertainties in the lens redshifts.

APPENDIX B: PRIORS ON WL MASS BIAS AND SCATTER

To obtain the priors on the WL mass bias $b_{\rm WL}$ and scatter $\sigma_{\rm WL}$ (equation 8), we create an ensemble of simulated observations that match the observational properties of a random subset of cluster fields and then apply the same measurement technique as we do to the real data. In general, we are aiming to reconstruct the probability distribution $P(M_{\rm meas}|M_{\rm true})$, which can then be included in forward modelling of the cluster sample. However, we simplify the relation as stated above to one lognormal distribution that is the same for all observed cluster fields. Any residuals from such an oversimplification are still insignificant compared to the obtainable statistical precision of our dataset.

To build our simulated observations for one observed cluster field, we start with the N-body simulations from Becker & Kravtsov (2011), with parameters $\Omega_{\rm m} = 0.27$, $\Omega_{\rm b} = 0.044$, $\sigma_8 = 0.79$, spectral index $n = 0.95$ and a Hubble constant of $h = 0.7$ in units of $100\,{\rm km\,s^{-1}\,Mpc^{-1}}$. We cut out $400h^{-1}$Mpc long boxes centred on the most massive 788 haloes with $M_{500,c} > 1.5 \times 10^{14} h^{-1}{\rm M}_\odot$ from the $z = 0.5$ snapshot. Particles are projected to form 2D mass maps that are then used to create shear maps via Fast Fourier transform. The observed $\langle\beta\rangle$ from a cluster observation is used to scale the shear and kappa maps appropriately. Random Gaussian noise is added to the shear map to match the observed shape noise in the observations. Because in our real observations we fit a 1D profile, we select an 'observed' cluster centre for each simulation map. We assume that the displacement between the true projected centre of the simulated cluster and the 'observed' centre is randomly oriented with respect to the underlying structure, a not unreasonable assumption given the noise sources of SPT observations and the statistical power of this sample. Centre offsets are randomly chosen following the form specified by Song et al. (2012), a Gaussian distribution with a width dependent on the SPT beam size and the core radius of the matched filter used to detect the observed cluster. The simulated 1D profiles are then fit with an NFW model as in the data analysis.

We assume that $P(M_{\rm meas}|M_{\rm true})$ follows a lognormal distribution where $\ln b_{\rm WL}$ is the mean of the distribution and $\sigma_{\rm WL}$ is its width. For the set of simulated fields, we find the maximum a posteriori location for the probability distribution

$$P(b_{\rm WL}, \sigma_{\rm WL}|{\rm mocks}) \propto \prod_i \int P(b_{\rm WL}, \sigma_{\rm WL}|M_{\rm meas}) P(M_{\rm meas}|{\rm mock}_i) {\rm d}M_{\rm meas}. \quad \text{(B1)}$$

Uninformative priors are used for the parameters of interest. Simulated observations are also created and analysed using the $z = 0.25$ snapshot from Becker & Kravtsov (2011) as well as the Millennium-XXL simulations (Angulo et al. 2012). No significant trends are seen between snapshots or simulations. We also do not see any significant trend with the observational properties of each observed field, including the amount of shape noise or different filter core size. Our final bias number is then the average of $b_{\rm WL}$ across the random subset of cluster fields targeted for mock up.

We measure $b_{\rm WL} = 0.936 \pm 0.04$ and $\sigma_{\rm WL} = 0.25 \pm 0.12$ when employing the M–c relation of Diemer & Kravtsov (2015). In practice, we add the systematic uncertainty of the weak-lensing mass to true mass relation to all other sources systematic errors (see Table 4) in quadrature and use a Gaussian prior $b_{\rm WL} = 0.936 \pm 0.185$ in our scaling relation analysis.

We can estimate the sensitivity of our analysis to the uncertainty in published mass–concentration relations by carrying out the NFW fit bias analysis for different fixed concentrations. We find that the

average mass bias at concentrations $c = 5$ and $c = 3$ is $b_{\rm WL} = 0.978$ and $b_{\rm WL} = 0.907$, respectively, implying $\mathrm{d}b_{\rm WL}/\mathrm{d}c|_{c=4} = -0.0355$. Using Gaussian error propagation on equation (8), we obtain

$$\left(\frac{\sigma_M}{M_{\rm true}}\right)^2 = \frac{1}{b_{\rm WL}^2}\left(\frac{\mathrm{d}b_{\rm WL}}{\mathrm{d}c}\right)^2 \sigma_c^2 \,. \tag{B2}$$

Because we calibrated the bias resulting from an NFW fit using the *M*–*c* relation of Diemer & Kravtsov (2015), the systematic uncertainty is not given by how well this relation describes the actual cluster sample, but by how faithfully the simulated clusters represent true clusters in the Universe. The simulations used in the previous section are Dark Matter only for a cosmology consistent with *WMAP*7 (Komatsu et al. 2011) and thus the question is how much would the concentrations for clusters of the mass and redshift in our sample be impacted by baryonic effects and the change of cosmological parameters to ones consistent with *Planck*. Duffy et al. (2010) constrain baryonic effects to an upper limit of 10 per cent, with baryons decreasing the concentration value. Klypin et al. (2016) show that concentrations are ∼10 per cent larger in *Planck* cosmologies than in the *WMAP* cosmology assumed in our calibration of the WL bias $b_{\rm WL}$. One could thus expect these effects to cancel, but we conservatively assume an uncertainty of 10 per cent on the concentration. Evaluating equation (B2), we set $\sigma_c|_{c=4} = 0.4$ and obtain a mass uncertainty due to the mass–concentration relation of 1.5 per cent. This turns out to be so much smaller than our other systematic uncertainties (cf. Table 4) that we can safely ignore it.

[1]*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany*
[2]*Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany*
[3]*Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA*
[4]*Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA*
[5]*Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121, Bonn, Germany*
[6]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, D-85748 Garching, Germany*
[7]*Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*
[8]*Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA*
[9]*National Center for Supercomputing Applications, 1205 West Clark Str, Urbana, IL 61801, USA*
[10]*Kavli Institute for Particle Astrophysics & Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA*
[11]*SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[12]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*
[13]*Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland*
[14]*INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[15]*Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA*
[16]*Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*
[17]*Department of Physics, The Ohio State University, Columbus, OH 43210, USA*
[18]*Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK*
[19]*Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[20]*Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637*
[21]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*
[22]*Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India*
[23]*Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Science, University of Colorado, Boulder, C0 80309, USA*
[24]*NASA Ames Research Center, Moffett Field, CA 94035, USA*
[25]*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*
[26]*Physics Department, Case Western Reserve University, Cleveland, Ohio 44106, USA*
[27]*Department of Physics, Yale University, New Haven, CT 06511, USA*
[28]*Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, Casilla 603, La Serena, Chile*
[29]*Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[30]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa*
[31]*Institute of Cosmology & Gravitation, University of Portsmouth, Portsmouth PO1 3FX, UK*
[32]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[33]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*
[34]*Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[35]*Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil*
[36]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*
[37]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08193 Barcelona, Spain*
[38]*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain*
[39]*Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), 28040 Madrid, Spain*
[40]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[41]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[42]*Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, E-28049 Madrid, Spain*
[43]*Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*
[44]*Australian Astronomical Observatory, North Ryde, NSW 2113, Australia*
[45]*Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP 05314-970, Brazil*
[46]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[47]*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*
[48]*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*
[49]*Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton BN1 9QH, UK*
[50]*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*
[51]*Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, Campinas, SP13083-859, Brazil*
[52]*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

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