arXiv:2402.08455v1 [astro-ph.CO] 13 Feb 2024

The SRG/eROSITA All-Sky Survey

Dark Energy Survey Year 3 Weak Gravitational Lensing by eRASS1 selected Galaxy Clusters

S. Grandis[1], V. Ghirardini[2], S. Bocquet[3], C. Garrel[2], J. J. Mohr[3,2], A. Liu[2], M. Kluge[2], L. Kimmig[3], T. H. Reiprich[4],
A. Alarcon[9], A. Amon[6,7], E. Artis[2], Y. E. Bahar[2], F. Balzer[2,3], K. Bechtol[8], M. R. Becker[9], G. Bernstein[10], E. Bulbul[2],
A. Campos[11], A. Carnero Rosell[12,13,14], M. Carrasco Kind[15,16], R. Cawthon[17], C. Chang[18,19], R. Chen[20], I. Chiu[21],
A. Choi[22], N. Clerc[23], J. Comparat[2], J. Cordero[24], C. Davis[25], J. Derose[26], H. T. Diehl[27], S. Dodelson[11,28],
C. Doux[14,29], A. Drlica-Wagner[18,19,27], K. Eckert[10], J. Elvin-Poole[30], S. Everett[31], A. Ferte[32], M. Gatti[10],
G. Giannini[33], P. Giles[34], D. Gruen[3], R. A. Gruendl[15,16], I. Harrison[35], W. G. Hartley[36], K. Herner[27], E. M. Huff[31],
F. Kleinebreil[1], N. Kuropatkin[27], P. F. Leget[37], N. Maccrann[38], J. Mccullough[37], A. Merloni[2], J. Myles[3,37,39],
K. Nandra[2], A. Navarro-Alsina[40], N. Okabe[41,42,43], F. Pacaud[4], S. Pandey[10], J. Prat[18,19], P. Predehl[2], M. Ramos[2],
M. Raveri[44], R. P. Rollins[24], A. Roodman[32,37], A. J. Ross[45], E. S. Rykoff[32,37], C. Sanchez[46], J. Sanders[2],
T. Schrabback[1], L. F. Secco[18,19], R. Seppi[2], I. Sevilla-Noarbe[46], E. Sheldon[47], T. Shin[48], M. Troxel[20], I. Tutusaus[49,50],
T. N. Varga[2,3], H. Wu[51], B. Yanny[31], B. Yin[52], X. Zhang[2], Y. Zhang[53]
O. Alves[53], S. Bhargava[34], D. Brooks[55], D. L. Burke[32,37], J. Carretero[33], M. Costanzi[56,57,58], L. N. da Costa[13],
M. E. S. Pereira[59], J. De Vicente[46], S. Desai[60], P. Doel[61], I. Ferrero[62] B. Flaugher[27], D. Friedel[15], J. Frieman[19,27],
J. García-Bellido[43], G. Gutierrez[27], S. R. Hinton[25], D. L. Hollowood[59], K. Honscheid[45,65], D. J. James[66], N. Jeffrey[62],
O. Lahav[62], S. Lee[31], J. L. Marshall[67], F. Menanteau[15,16], R. L. C. Ogando[68], A. Pieres[13,68],
A. A. Plazas Malagón[32,37], A. K. Romer[34], E. Sanchez[46], M. Schubnell[53], M. Smith[69], E. Suchyta[59],
M. E. C. Swanson[15], G. Tarle[53], N. Weaverdyck[26,53], and J. Weller[2,3]
(The Dark Energy Survey Collaboration and the eROSITA-DE Consortium)

Affiliations at the end of the paper, e-mail: `sebastian.grandis@uibk.ac.at`

Received—; accepted —

ABSTRACT

Context. Number counts of galaxy clusters across redshift are a powerful cosmological probe, if a precise and accurate reconstruction of the underlying mass distribution is performed – a challenge called *mass calibration*. With the advent of wide and deep photometric surveys, weak gravitational lensing by clusters has become the method of choice to perform this measurement.

Aims. We measure and validate the weak gravitational lensing (WL) signature in the shape of galaxies observed in the first 3 years of the Dark Energy Survey (DES Y3) caused by galaxy clusters and groups selected in the first all-sky survey performed by SRG (Spectrum Roentgen Gamma)/eROSITA (eRASS1). These data are then used to determine the scaling between X-ray photon count rate of the clusters and their halo mass and redshift.

Methods. We empirically determine the degree of cluster member contamination in our background source sample. The individual cluster shear profiles are then analysed with a Bayesian population model that self-consistently accounts for the lens sample selection and contamination, and includes marginalization over a host of instrumental and astrophysical systematics. To quantify the accuracy of the mass extraction of that model, we perform mass measurements on mock cluster catalogs with realistic synthetic shear profiles. This allows us to establish that hydro-dynamical modelling uncertainties at low lens redshifts ($z < 0.6$) are the dominant systematic limitation. At high lens redshift the uncertainties of the sources' photometric redshift calibration dominate.

Results. With regard to the X-ray count rate to halo mass relation, we determine its amplitude, its mass trend, the redshift evolution of the mass trend, the deviation from self-similar redshift evolution and the intrinsic scatter around this relation.

Conclusions. The mass calibration analysis performed here sets the stage for a joint analysis with the number counts of eRASS1 clusters to constrain a host of cosmological parameters. We demonstrate that WL mass calibration of galaxy clusters can be performed successfully with source galaxies whose calibration was performed primarily for cosmic shear experiments, opening the way for the cluster cosmological exploitation of future optical and NIR surveys like *Euclid* and LSST.

Key words. Gravitational lensing: weak – Cosmology: large-scale structure of Universe – X-rays: galaxies: clusters

1. Introduction

General Relativity describes the gravitational force as curvature of space-time, induced by mass and energy densities. This cur-

vature determines the so-called null geodesics on which photons travel through space time. As a consequence, the path of the light from a distant source to the observer is deflected by intervening gravitational potentials, with a deflection angle that is proportional to the path integral of the transverse component of the negative gradient of the potential. Indeed, Dyson et al. (1920) measured the predicted displacement in the positions of distant stars when observed close to the Sun during a solar eclipse. In the cosmological context, the original position of the sources is not known. In general, also the gravitational potential of the intervening lenses is unknown, because the majority of cosmic matter is believed to be invisible and its density is among the targets of our experiments.

On cosmic scales, gravitational lensing of distant galaxies thus relies on the fact that the differential deflection of neighboring light paths locally induces a magnification and a distortion of the source galaxy image (for a pedagogical introduction see Schneider 2006). In this regime, large-scale gravitational potentials create spatially coherent, small distortions in the observed shapes of background galaxies, a process called *weak gravitational lensing* (hereafter WL). A special case is the WL caused by massive gravitationally collapsed, virialised structures, termed halos, which dominate the line-of-sight-integrated gravitational potential (for a review see Umetsu 2020). These objects induce a tangential distortion in the shapes of the background galaxies which traces the density contrast of the 2-d projected density profiles. The 3-d density profiles of halos are well understood in simulations, and are found to tightly follow a parametric class of functions with two free parameters: the total halo mass and the typical halo scale (Navarro et al. 1996). Fitting predictions derived from this family of models to the tangential reduced shear profiles thus enables the direct measurement of the halo mass.

The most massive halos host galaxy clusters, whose observed properties display tight correlations among themselves (Mohr & Evrard 1997), and with the host halos mass (Angulo et al. 2012), enabling a clean halo selection. For such samples of galaxy clusters, the WL mass information is then used to reconstruct the differential number density of halos as a function of mass and redshift, the halo mass function. It traces the growth of structure in the Universe, and is therefore a sensitive probe to the total matter density, the accelerated expansion, and gravity itself (Haiman et al. 2001; Majumdar & Mohr 2004; Allen et al. 2011). Examples of such analyses are Mantz et al. (2015), who used X-ray selected galaxy clusters with pointed WL observations, and Bocquet et al. (2019), who used galaxy clusters selected at millimeter wavelengths. Both these analyses result in competitive cosmological constraints, which are largely complementary and independent of other cosmological probes (for a recent review, see Huterer 2022).

Our ability to detect larger numbers of massive halos has recently been transformed by eROSITA, which performed its first All-Sky Survey from December 2019 to June 2020 (Predehl et al. 2021; Sunyaev et al. 2021), detecting more than 1 million X-ray sources in the Western Galactic Hemisphere (Merloni et al. 2024). Among these X-ray sources, ~ 12k are confirmed as clusters of galaxies through significant extended X-ray emission and associated red galaxy members in the 12791 deg^2 footprint covered by the DESI Legacy Survey DR9 and DR10 data (Bulbul et al. A&A, subm.; Kluge et al. A&A, subm.). In this work, we will restrict ourselves to the 5263 clusters selected for the *cosmology* sample, via a higher extent likelihood cut. They have percent accurate photometric redshifts, in the range of

0.1 to 0.8, and a well-calibrated selection function (Clerc et al. A&A, subm.).

We complement the eROSITA data with the wide photometric Dark Energy Survey year 3 (DES-Y3) data (Sevilla-Noarbe et al. 2021). From these data we use more than 10^8 galaxy shape measurements measured in the *riz* bands by Gatti & Sheldon et al. 2021 in an effective survey area of 4,143 deg^2 of the southern sky. The strength of the WL signal around galaxy clusters depends not only on their mass, but also on the geometrical configuration of lens and source. To estimate statistically the distance of the latter we leverage the exquisite accuracy of the photometric redshift calibration of the DES-Y3 sources by Myles & Alarcon et al. 2021.

Besides the uncertainty on the shape and photometric redshift measurements for the source galaxies, percent accurate cluster WL measurements need to account for the cluster miscentring errors and contamination of the source sample by cluster galaxies (for a complete analysis on DES year 1 data, see McClintock & Varga et al. 2019). Even if all observational systematics are well-calibrated and accounted for, the modelling uncertainties of baryonic feedback processes alter the mass distribution of galaxy clusters. This poses a theoretical limit to our ability to calibrate cluster masses via WL (Grandis et al. 2021a). As argued in that work, setting up many realisations of realistic, synthetic cluster WL mock datasets, and analysing them with the same mass extraction method as the real data, allows one to calibrate both the absolute values, and – crucially – also the uncertainty of the *WL mass bias*. This approach has already been successfully demonstrated in work by Chiu et al. (2022) in the context of Hyper Supreme Camera (HSC) observations of the eROSITA Final Equatorial-Depth Survey (eFEDS) used for science verification.

This paper is organized as follows: in Section 2 we describe the lens and source catalogs, as well as the calibration products used in this analysis. Section 3 then describes the measurement process for the tangential reduced shear, and the suite of associated validation and calibration steps. In Section 4 we then describe the shear profile model used for the mass measurement, as well as its calibration on realistic synthetic data. These two sections follow in significant part the analysis performed by Bocquet et al. (2023) of the DES Y3 WL signal around South Pole Telescope (SPT) selected clusters. The actual mass calibration is then performed and validated in Section 5, and we discuss our results in Section 6. Halo masses in this work are reported as spherical over density masses $M_{\Delta c}$, with $\Delta = 500, 200$. This means that they are defined via the radius $R_{\Delta c}$ enclosing a sphere of average density Δ times the critical density of the Universe at that redshift, i.e. $M_{\Delta c} = \frac{4\pi}{3}\Delta\rho_{crit}(z)R_{\Delta c}^3$. We use a flat ΛCDM cosmology with parameters $\Omega_M = 0.3$, and $h = 0.7$ as reference cosmology, where not specified otherwise. Throughout this work, we will use lens redshift z_l, and cluster redshift z_{cl} interchangeably, depending on which is more applicable to the context of the discussion.

2. Data

2.1. eRASS1 Cluster and Group sample

As a lens catalog, we use the cluster catalog acquired from the first eROSITA All-Sky Survey on the Western Galactic Hemisphere, completed on June 11, 2020. The detection properties of the X-ray sources detected in this survey, including extended sources, are discussed in detail in Merloni et al. (2024); Bulbul et al. (A&A, subm.). We use the cosmology catalog de-
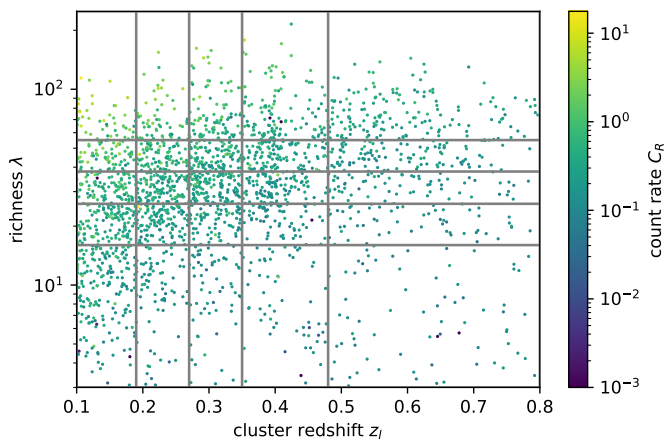
Fig. 1. Lens sample composed of the eRASS1 cosmology cluster and group sample with DES WL data, plotted in redshift against richness, with color coded the X-ray count rate. The gray lines mark the edges of the richness–redshift bins used for the calibration and validation of the WL measurement. Cosmology ready WL data products are however produced for each cluster individually.

Fig. 2. Source redshift distributions of the tomographic redshift bin of the DES Y3 data. We plot the mean of the 1000 realisations (full line), together with the 16-84 percentile (filled region), and the 2.5-97.5 percentile (faded lines). The first tomographic bin is not used, while the others are weighted so as to ensure a lens dependent background selection.

scribed in detail in Bulbul et al. (A&A, subm.), comprised of X-ray sources detected as significantly extended, with reliable photometric confirmation and redshift estimation using the DESI Legacy Survey DR10 (Dey et al. 2019). Specifically, Kluge et al. (A&A, subm.) run an adapted version of the `redMaPPer` algorithm (Rykoff et al. 2014; Rozo et al. 2015; Rykoff et al. 2016) on the X-ray cluster candidate position, thus measuring their richness and photometric redshift. Of the 5,263 clusters and groups in that catalog, 2,201 have DES Y3 shape and photo-z information (described below), and are therefore used for the WL analysis in this work (see Ghirardini et al. A&A, subm., Fig. 1 for a comparison of the two survey footprints). Their distribution in redshift, richness and X-ray photon count rate is shown in Fig. 1. The bulk of the lens sample lies at low redshift, ideally placed for WL studies. Some objects at low richness have redshifts much larger then expected for an X-ray selected sample. These are contaminants, which we account for in our mass calibration (see Section 5.1.1).

2.2. DES WL data

The Dark Energy Survey is an approximately 5,000 deg^2 photometric survey in the optical bands $grizY$, carried out at the 4m Blanco telescope at the Cerro Tololo Inter-American Observatory (CTIO), Chile, with the Dark Energy Camera (DECam Flaugher et al. 2015). In this analysis we utilize data from the first three years of observations (DES Y3), covering the full survey footprint.

2.2.1. The shape catalog

The DES Y3 shape catalog (Gatti & Sheldon et al. 2021) is built from the r, i, z-bands using the METACALIBRATION pipeline (Huff & Mandelbaum 2017; Sheldon & Huff 2017). Other DES Y3 publications contain more detailed information about the photometric dataset (Sevilla-Noarbe et al. 2021), the Point-Spread Function modeling (Jarvis et al. 2021), and image simulations (MacCrann et al. 2022). We refer the reader to these works for more information. After application of all source selection cuts, the DES Y3 shear catalog contains about 100 million galaxies
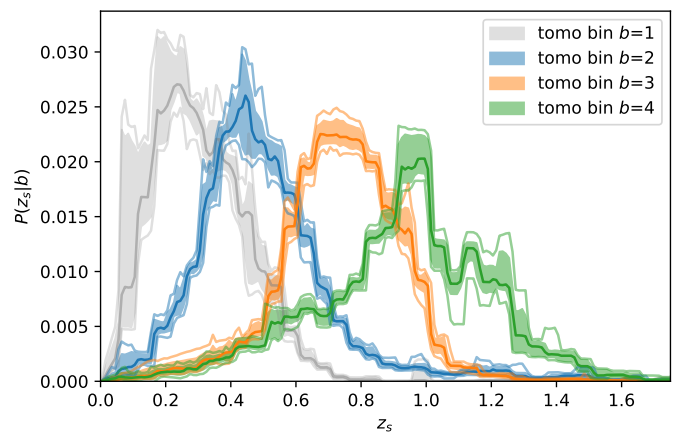
over an area of 4,143 square-degrees. Its effective source density is 5–6 arcmin^{-2}, depending on the exact definition. In Sections 3, we will derive the exact values of effective source density.

2.2.2. Source redshift distributions and shear calibration

Our analysis uses the same selection of lensing source galaxies in tomographic bins as the DES 3x2pt analysis (Dark Energy Survey Collaboration et al. 2022). This selection is defined in Myles & Alarcon et al. 2021. In that work, source redshifts are estimated with Self-Organizing Maps, and the method is thus referred to as SOMPz. The final calibration accounts for the (potentially correlated) systematic uncertainties in source redshifts and shear measurements, as determined in the image simulations by MacCrann et al. (2022). For each tomographic source bin, the estimated redshift distribution is provided, and the systematic uncertainties on this estimate are captured through 1,000 realizations (see Fig. 2). These realisations are indexed via the hyper parameter $\mathcal{H}$, which takes integer values between 0 and 999 (Cordero et al. 2022). For convenience, these redshift distributions are normalized to varying values $(1 + m)$, where the varying m spans the range of the multiplicative shear bias values. Note that the DES Y3 survey only has minor depth variations over the survey area. Survey averaged quantities, like the source redshift distributions, are thus applicable to the joint eROSITA-DE – DES Y3 footprint.

We will also use DNF (De Vicente et al. 2016) and BPZ (Benítez 2000) source redshift measurements to constrain the amount of cluster member contamination in our source sample (see Section 3.3.1). While BPZ is a Bayesian template fitting code, DNF is based on a nearest neighbor interpolation on the color magnitude space of spectroscopic reference samples. As such, BPZ is more robust against the incompleteness of the spectroscopic reference samples, while DNF is more data driven, and avoids the biases that occur from the template choice (see Sevilla-Noarbe et al. 2021, Section 6.3, for more details).

3. Measurement

In this section, we shall outline 1) the methods used to measure the WL signal around eROSITA selected clusters, 2) the

steps undertaken to correctly calibrate this measurement, and 3) the use of this measurement to determine the mass scale of the eROSITA selected clusters and groups. Note that the methods presented in this work draw significantly upon the methods presented in Bocquet et al. (2023).

3.1. WL by massive halos

Assuming that background galaxies are isotropically oriented intrinsically, the coherent tangential distortion induced by the gravitational potential on background galaxy images at a radius R from the halo center can be obtained, to linear order, by averaging the tangential component of their ellipticities,

$$g_t(R) = \mathcal{R}^{-1} \langle e_t \rangle_R, \qquad (1)$$

where $g_t(R)$ is the reduced tangential reduced shear, and $\mathcal{R}$ is the average response of the measured ellipticity to a shear (shear response). In practical applications, instrumental effects and noise make it different from 1. It is determined in the process of galaxy shape measurement and validated through image simulations. We discuss below how we include this effect in our measurements, and how we account for non-linear responses. While the ellipticity field itself is not Gaussian, our reduced shear estimator is constructed by averaging a large number of sources. The central limit theorem thus ensures that the statistical error of the shear estimate can be directly derived from the effective dispersion σ_{eff} in the ellipticity of the source galaxies, and their effective number N_{eff}. We will discuss below how we estimate these two quantities.

Given a source galaxy at redshift z_s, the lens redshift z_l, and the surface mass density $\Sigma(\mathbf{x})$ in the lens plane, the tangential reduced shear in any position on the sky $\boldsymbol{\theta} = \mathbf{x}/D_l$ is given by

$$g_t(\boldsymbol{\theta}) = \frac{\gamma_t(\mathbf{x})}{1 - \kappa(\mathbf{x})} \text{ with } \kappa(\mathbf{x}) = \Sigma_{\text{crit}}^{-1} \Sigma(\mathbf{x}), \qquad (2)$$

where $\gamma_t(\mathbf{x})$ is the tangential component of the shear, $\kappa(\mathbf{x})$ the convergence, D_l the angular diameter distance to the lens, and $\Sigma_{\text{crit}}^{-1}$ the inverse of the critical lensing surface density. The latter expresses the geometrical configuration of the source and lens, and is given by

$$\Sigma_{\text{crit,ls}}^{-1} = \frac{4\pi G}{c^2} \frac{D_l}{D_s} \max\left[0, D_{ls}\right], \qquad (3)$$

where G is the gravitational constant, c the speed of light. Furthermore, we use the angular diameter distance between observer and source (D_s), and lens and source (D_{ls}). Note that for sources in front of the lens, D_{ls} becomes negative. These sources are not lensed, which we account for by setting their inverse critical lensing surface density to zero, hence the term $\max[0, D_{ls}]$.

In General Relativity, the azimuthally averaged tangential shear at projected radius R is proportional to the density contrast,

$$\gamma_t(R) = \Sigma_{\text{crit}}^{-1}\left[\langle \Sigma(< R) \rangle - \Sigma(R)\right], \qquad (4)$$

while the orthogonal, B-mode like component γ_x, called cross-shear (see below for the exact definition of the decomposition), averages to zero when integrated over closed paths such as radial bins,

$$\gamma_x(R) = 0. \qquad (5)$$

We will use the latter as a validation test for our measurement.

3.2. Measurement

To perform the measurement, for each lens l we query the shape catalog for all sources at a projected physical distance of < 15 Mpc h^{-1} in the reference cosmology ($h = 0.7$, $\Omega_M = 0.3$). For each source lens pair (i, l), we compute the position angle φ at which the source is seen from the lens' position.

We then project the ellipticity components $e_{1,2}$ onto the tangential and cross components as

$$e_t = -e_1 \cos 2\varphi + e_2 \sin 2\varphi$$
$$e_x = e_1 \sin 2\varphi + e_2 \cos 2\varphi, \qquad (6)$$

We use here the tangential/cross decomposition for the case where the e_2-component is defined w.r.t. right ascension, which is the case in the DES Y3 shape catalog.[*] We also record the source weight w_i^s, the response $\mathcal{R}_i$, as presented in Gatti & Sheldon et al. 2021, Section 4.3, and the photometric redshift estimates in form of the SOMPz cell $c_{\text{SOM},i}$ of the source, and the photometric redshift estimates $\hat{z}_{\text{DNF},i}$ and $\hat{z}_{\text{BPZ},i}$.

To select only sources in the background of each lens, we discard the first tomographic redshift bin and apply a weight to each source depending on the tomographic redshift bin b it resides in,

$$w^b = \begin{cases} \left\langle \Sigma_{\text{crit,ls}}^{-1} \right\rangle_{s \in b} & \text{for } z_l < z_{\text{med},b} \text{ and } b > 1 \\ 0. & \text{otherwise,} \end{cases} \qquad (7)$$

where the average is taken using the mean source redshift distribution of the tomographic bin, as reported by Myles & Alarcon et al. 2021. Our background selection is thus a weighted sum of the tomographic bin selections. This is convenient, as the shear and photo-z calibration (Myles & Alarcon et al. 2021; Cordero et al. 2022; MacCrann et al. 2022) is only valid for the specific selection criteria of the tomographic redshift bins. It can be extended to our analysis by using the same weights as above.

To validate and calibrate the cluster WL measurement, we define lens richness bins (3, 16, 26, 38, 55, 250), and redshift bins (0.10, 0.19, 0.27, 0.35, 0.48, 0.8), shown as grey lines in Fig. 1, and tuned to result in approximately even occupations of the lenses. In these bins we measure the following quantities for radial bins whose upper edges are log-equally spaced between 0.2–15 h^{-1}Mpc at the lens redshift in the reference cosmology:

– the raw tangential and cross-shear, with the tangential components shown in Fig. 3,

$$g_{\alpha,\text{raw}} = \frac{\sum_{b=2,3,4} w^b \sum_{i \in b} w_i^s e_{\alpha,i}}{\sum_{b=2,3,4} w^b \sum_{i \in b} w_i^s \mathcal{R}_i} \text{ for } \alpha \in (t, x), \qquad (8)$$

where $i \in b$ stands for the source i in the tomographic bin b and in the respective lens richness, redshift and distance bins.

– the effective number of sources

$$N_{\text{eff}} = \frac{\left(\sum_{b=2,3,4} \sum_{i \in b} w^b w_i^s \mathcal{R}_i\right)^2}{\sum_{b=2,3,4} \sum_{i \in b} \left(w^b w_i^s \mathcal{R}_i\right)^2}, \qquad (9)$$

– the effective dispersion of source ellipticities

$$\sigma_{\alpha,\text{eff}}^2 = \frac{\sum_{b=2,3,4} \sum_{i \in b} \left(w^b w_i^s\right)^2 e_{\alpha,b,i}^2}{\sum_{b=2,3,4} \sum_{i \in b} \left(w^b w_i^s \mathcal{R}_i\right)^2} \text{ for } \alpha \in (t, x), \text{ and} \qquad (10)$$

---

[*]The prevalent notation (e.g. Umetsu 2020, Eq. 83) is valid for e_2 defined for a right-handed, local coordinate system ($x \propto -\text{RA}$). It can be recovered by the transformation $e_2 \mapsto -e_2$.
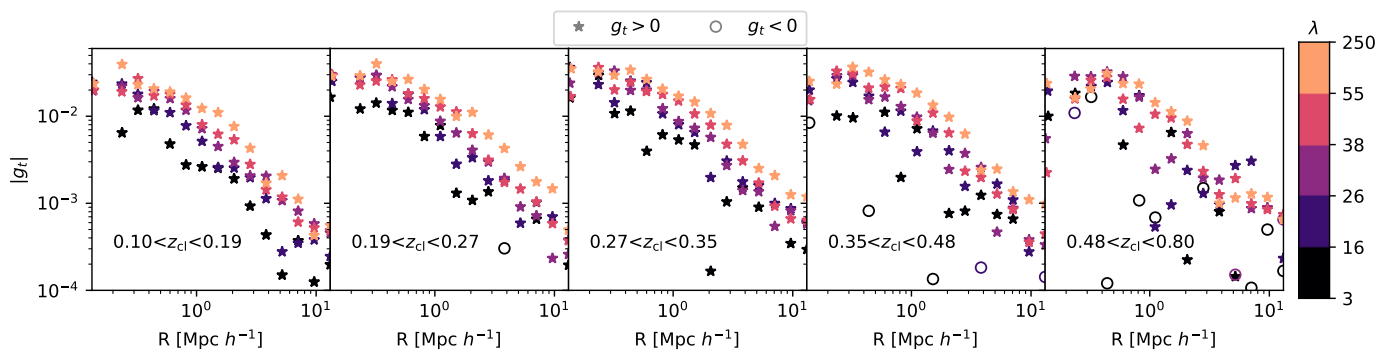
Fig. 3. Absolute value of the raw tangential reduced shear profile, Eq. 8, in bins of cluster redshift (panels: lowest redshift left, higher redshift right) and of richness (color coding), for our background selection. Stars denote positive values, while empty circles denote negative values. The profiles show a clear increase towards the cluster centers, as well as a trend with richness, in line with our qualitative expectations.

– the source redshift distribution for fine source redshift bins with edges (z_{s-}, z_{s+}), reading

$$\hat{P}_\beta(z_s) = \frac{\sum_{b=2,3,4} w^b \sum_{i \in b} w_i^s \mathcal{R}_i \, \mathrm{I}_s(\hat{z}_{\beta,i})}{\sum_{b=2,3,4} w^b \sum_{i \in b} w_i^s \mathcal{R}_i} \text{ for } \beta \in (\text{BPZ, DNF}), \quad (11)$$

where $\mathrm{I}_s(\hat{z}_{\beta,i}) = 1$ when $z_{s-} < \hat{z}_{\beta,i} < z_{s+}$. Practically speaking, this is a weighted histogram of the photo-z estimates $\hat{z}_{\beta,i}$.

It is important to note that also the estimates of the effective source density, the shape dispersion and the redshift distribution need to take the shear response $\mathcal{R}$ into account. For a derivation for Eq. 9, and Eq. 10, we refer the reader to App. A, where we present the derivation of the effective number of sources and their effective shape dispersion in the presence of a shear response and un-normalized weights.

3.3. Calibration and validation

We shall now discuss several calibration and validation steps we take to ensure the quality of our WL measurements.

3.3.1. Cluster member contamination

As suggested by their name, galaxy clusters are substantial overdensities not only in matter, but also in the galaxy field. As a result, an a priori unknown fraction of the cluster members contaminates any source background selection. These cluster member contaminants carry no shear distortion signal from the cluster potential, and therefore dilute the shear signal (Hoekstra et al. 2012; Gruen et al. 2014; Dietrich et al. 2019; Varga et al. 2019). This effect is equivalent to source-lens clustering in other WL applications (compare, for instance, with Prat et al. 2022, Section III.B). Current simulations are not accurate enough to reconstruct the colors and magnitudes of cluster member galaxies with sufficient fidelity to calibrate this effect directly in simulations. We therefore resort to empirical calibration methods. We shall outline in the following paragraphs the model for the cluster member contamination, and the two methods we use to fit for it, as well as a comparison of the fit results.

Cluster member contamination model In modeling the cluster member contamination we adopt an approach that was developed previously in Paulus (2021). We parametrize the radial

fraction of cluster contaminants as

$$f_{\mathrm{cl}}(R|\lambda, z) = \frac{A(\lambda, z, R)}{1 + A(\lambda, z, R)} \text{ with}$$

$$A(\lambda, z_j, R) = e^{A_j}\left(\frac{\lambda}{25}\right)^{B_\lambda} \Sigma_{\mathrm{NFW}}^{\mathrm{norm}}\left(R\Big|r_S = c^{-1}\left(\frac{\lambda}{20}\right)^{1/3}\right), \quad (12)$$

where A_j is a free amplitude parameter for each cluster redshift bin j we consider, B_λ describes the richness trend of the cluster member contamination, and $\Sigma_{\mathrm{NFW}}^{\mathrm{norm}}(R|r_S)$ is a 2d projected NFW profile, normalized to 1 at its scale radius r_S, which we parameterize via a concentration c and a scaling with richness.

As shown in App. B, $A(\lambda, z, R)$ is proportional to the radial number density profile of the cluster member contaminants, which we therefore model as a 2d projected NFW profile. Given that at this stage of the analysis, the cluster mass is still unknown, we resort to using the richness as a mass proxy. For ICM-selected cluster samples, the richness is known to scale approximately linearly with mass (Saro et al. 2015; Bleem et al. 2020; Grandis et al. 2020, 2021b). Furthermore, the richness provides a convenient proxy also for the total number density of cluster member galaxies. It also has contributions from correlated structures along the line of sight, that would also contribute almost unsheared contaminants to the background sample.

Regularisation Given the complex interactions of cluster redshift, cluster member colors, photo-z estimation and background selection, we opt for a non-parametric fit for the redshift evolution by fitting independent amplitudes A_j for each cluster redshift bin. To impose a data-driven smoothness on the redshift evolution, we place a Gaussian prior with unknown variance σ_{reg}^2 on the difference between the amplitudes of neighboring redshift bins, implemented via the Gaussian likelihood

$$\ln \mathcal{L}_{\mathrm{reg}} = \frac{n_{\text{z-bins}} - 1}{2} \ln \sigma_{\mathrm{reg}}^2 - \frac{1}{2\sigma_{\mathrm{reg}}^2} \sum_j \left(A_j - A_{j+1}\right)^2, \quad (13)$$

where $n_{\text{z-bins}}$ is the number of redshift bins used. This likelihood adds an extra fit parameters, σ_{reg}, when added to the primary likelihood of the cluster member contamination.

Source density fit The cluster member contamination can be fitted from the effective source density as a function of projected distance from the clusters (which are our lenses) stacked in bins of cluster redshift and richness, which we present in Fig. 4. This density is computed by dividing the effective number of sources,
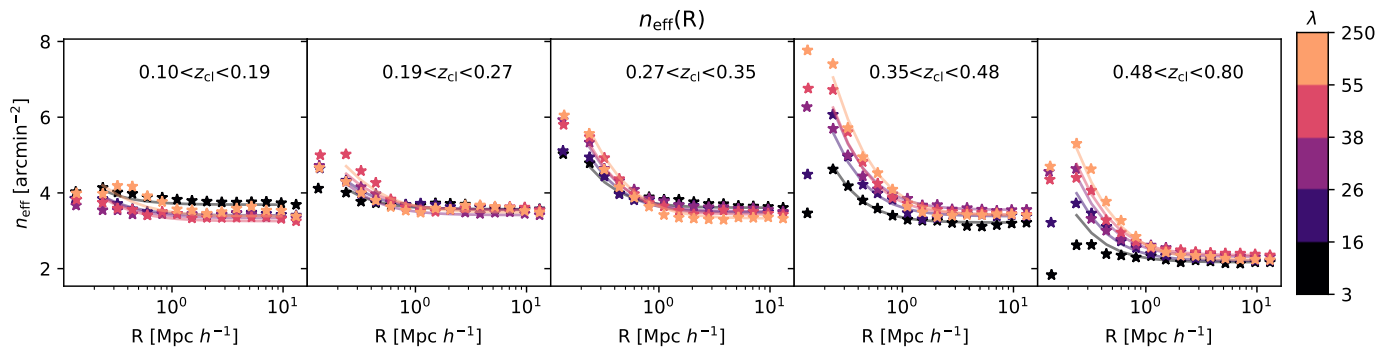
Fig. 4. Effective source density estimated from the effective number of source, Eq. 9, and the geometric area of the radial bins stacked in bins of cluster redshift (panels, lowest redshift left, higher redshift right) and of richness (color coding) as stars, for our background selection. Cluster members contaminating our background source sample lead to an increase of the effective source density towards the cluster center which depends on both cluster richness and redshift. This signal needs to be determined, as cluster member contaminants are not sheared and dilute the shear estimator. Our model (lines) captures this trend well.
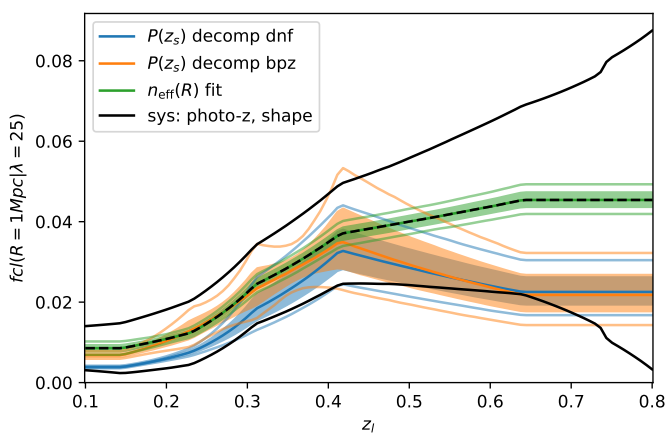
Fig. 5. Cluster redshift trend of the cluster member contamination of a richness $\lambda = 25$ object at 1 Mpc from the cluster center. In green the fit to the effective source number density, while in blue and orange the constraints from the DNF and BPZ source redshift distribution decomposition are shown. The latter method is independent of masking. While we find marginal agreement between the contamination fractions determined by the two methods, the difference is smaller than the relative error induced by shape and photo-z systematics – plotted in black extending from the source density fit for comparison.

Eq. 9, by the geometric area of the radial bin. Given the presence of cluster member galaxies obstructing the detection of background galaxies, and of detection masks, this results in a under-estimation of the actual source density, as explored in detail by Kleinebreil et al. (in prep.) in the context of Kilo Degree Surveys (KiDS) WL around SPT selected clusters. This can be seen most clearly in the reported number density in the most central radial bin $R < 0.2\ h^{-1}$ Mpc. On these radial scales, cluster lines of sight are typically dominated by the massive central galaxy and its stellar envelope. We exclude this radial bin from the following analyses.

We fit the effective number density of cluster members as a function of projected radius R for a cluster of richness λ and redshift z as

$$n_{\text{eff}}(R|\lambda, z) = \frac{n_{\text{field}}(z)}{1 - f_{\text{cl}}(R|\lambda, z)}, \tag{14}$$

where the effective number density in the field $n_{\text{field}}(z)$ is extracted around our clusters in the outer regions, $11\ h^{-1}$Mpc $< R < 15\ h^{-1}$ Mpc.

For the fit of the parameters of the cluster member contamination, $\boldsymbol{p}_{\text{cl}} = (A_j, B_\lambda, c)$, we set up a Gaussian likelihood for the effective number density measured in the above mentioned richness-redshift bins. Our data vector, the effective source density, results from the weighted sum of all tomographic redshift bins (see Eq. 9). We therefore fit one global cluster member contamination for our overall background selection. We assume a covariance matrix which combines Poisson noise from the effective number of sources and the variance we find in the outer radial bin, as a proxy for the cosmic variance in the source density. Sampling this likelihood together with the regularization likelihood provides us with constraints on the parameters of the cluster member contamination.

Source redshift distribution decomposition The estimate from the effective source density around clusters needs to be validated independently, as in the current implementation we did not consider source obstruction and masking. As demonstrated by Varga et al. (2019), and applied by McClintock & Varga et al. 2019; Paulus (2021); Shin et al. (2021); Bocquet et al. (2023), cluster member contamination can also be determined by decomposing the source redshift distribution as a function of radius into a field component and a Gaussian component of cluster members, i.e.,

$$P(z_s|R, \lambda, z) = (1 - f_{\text{cl}}(R|\lambda, z))\hat{P}_{\text{field}}(z_s|z) + f_{\text{cl}}(R|\lambda, z)P_{\text{cl}}(z_s|z), \tag{15}$$

where the field source redshift distribution $\hat{P}_{\text{field}}(z_s|z)$ is measured in the outer radial bin. The cluster member component is modelled as a Gaussian $P_{\text{cl}}(z_s|z) = \mathcal{N}(z_s|z + \mu(z), \sigma^2(z))$ with mean offset $\mu(z) = \mu_0 + \mu_z(z - z_0)$, and standard deviation $\sigma(z) = \sigma_0 + \sigma_z(z - z_0)$, with pivot $z_0 = 0.3$, close to the cluster redshift median. This technique is an adaptation of the method proposed by Gruen et al. (2014) to wide photometric surveys with readily available photometric redshift estimates. Given that we fit for the shape of a normalized source redshift distribution, the amount of masking does not impact this inference. The likelihood is

$$\ln \mathcal{L}_\beta = \sum N_{\text{eff}} \hat{P}_\beta(z_s) \ln P(z_s|R, \lambda, z) \text{ for } \beta \in (\text{BPZ}, \text{DNF}), \tag{16}$$
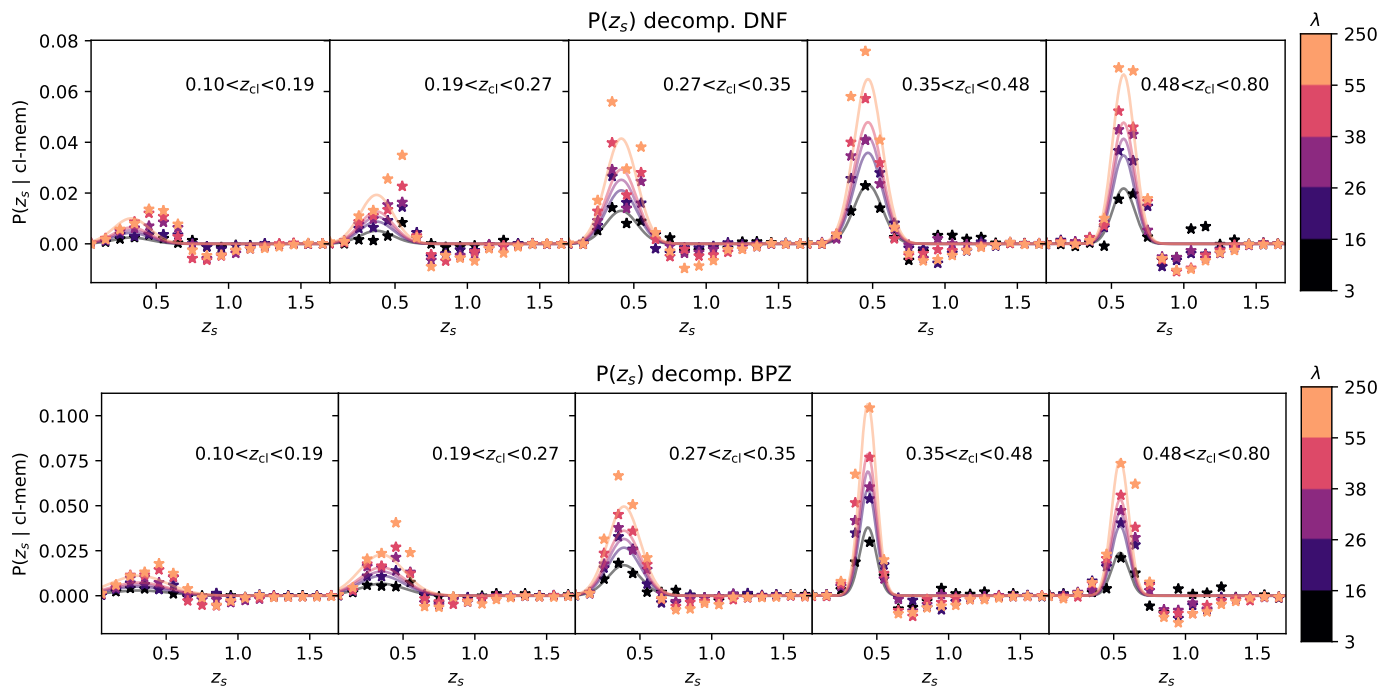
Fig. 6. Validation of the source redshift distribution decomposition for DNF (top) and BPZ (bottom) source redshift estimates. In different richness (color coded) and redshift bins (different panels), we show the difference between the source redshift distribution measured along the cluster line of sights and the field distribution extracted at a safe distance from the cluster center, shown as stars. Both are shown here for the projected cluster-centric distance of 1 Mpc in our reference cosmology. The resulting increment is well modelled by a Gaussian component (lines) caused by the cluster members contaminating the source sample. The amplitude of this component scales with cluster richness.

where the sum runs over richness, cluster redshift, cluster-centric distance and source redshift bins.[†] Sampled together with the regularization likelihood, this provides us with constraints on the parameters of the cluster member contamination. We perform this fit for redshift distributions constructed both for BPZ and DNF photo-zs.

Results and Comparison The best-fit results of the number density fit are over plotted as lines on the effective source density data in Fig. 4, demonstrating that our model is able to capture the trends in the data well. Furthermore, we find a concentration for the contaminants profile of $c = 2.5 \pm 0.1$. This matches the concentration values found for the galaxy populations of massive clusters (Hennig et al. 2017). Our results also indicate a slight richness slope of the contamination fraction of $B_\lambda = 0.47 \pm 0.1$. The redshift evolution is shown in Fig. 5 in green (with the filled band corresponding to the 1-sigma uncertainty region, and the green lines indicate the 2-sigma). For a cluster at richness pivot $\lambda = 25$ at a cluster-centric distance of $R = 1$ Mpc, we find that the contamination fraction increase from around $f_{\rm cl} \approx 0.01$ at $z_{\rm cl} = 0.1$ to $f_{\rm cl} \approx 0.05$ for $z_{\rm cl} > 0.5$, with the strongest increase happening between redshifts $0.25 < z < 0.4$. We will use this setting to gauge the impact of the different cluster member contamination fits w.r.t. other systematics.

The source redshift distribution decomposition fits are validated by plotting differences between the measured source redshift distribution and the re-scaled field distribution, as shown in Fig. 6 for 1 Mpc from the center. This subtraction clearly

highlights an approximately Gaussian component, that is well described by our model, plotted as full lines. The low amplitude negative dips at source redshifts slightly larger than the Gaussian component show that our field re-scaling, and thus our cluster member estimate, is not perfect. These residuals, however, are of the scale of the fluctuations in the data. We therefore accept the source redshift distribution decomposition as a valid fit.

Constraints on the concentration of the contaminants profile, and the richness trend of the contamination fraction, are in statistical agreement between the fit to the effective number density and to the source redshift distributions. The contamination fraction resulting from the source redshift distribution is plotted in Fig 5. It shows qualitative agreement with the contamination fraction constrained from the effective source density at low redshift. We will argue in the following, why the high redshift differences will not bias the mass calibration.

3.3.2. Shape measurement and photo-z uncertainty

We compare the difference between the cluster member contamination estimates to the shape and photo-z measurement uncertainties of the DES Y3 data as follows. The shape and photo-z uncertainties of the tomographic redshift bin b are summarized by realisations of the source redshift distributions, $P_{\mathcal{H}}(z_s|b)$, running over a hyper parameter $\mathcal{H}$. The variation in the shape of these distributions expresses the uncertainty in the photo-z calibration, while variation in the normalization expresses uncertainties in the multiplicative shear bias $1 + m$, as described in more detail in Cordero et al. (2022). Folding this together with our background selection, implemented via lens-redshift dependent weights w_b for the different tomographic redshift bins b (e.g., Eq. 7), we find a fractional uncertainty of mean lensing efficiency

---

[†]This follows directly from the likelihood of a sample x_i being drawn from a model distribution $P(x)$, $\ln \mathcal{L} = \sum_{i \in b} \ln P(x_i)$. In our case, the model distribution is $P(z_s|R, \lambda, z)$, and the number of samples in each richness, cluster-centric distance, source redshift bin is $N_{\rm eff} \hat{P}_\beta(z_s)$.
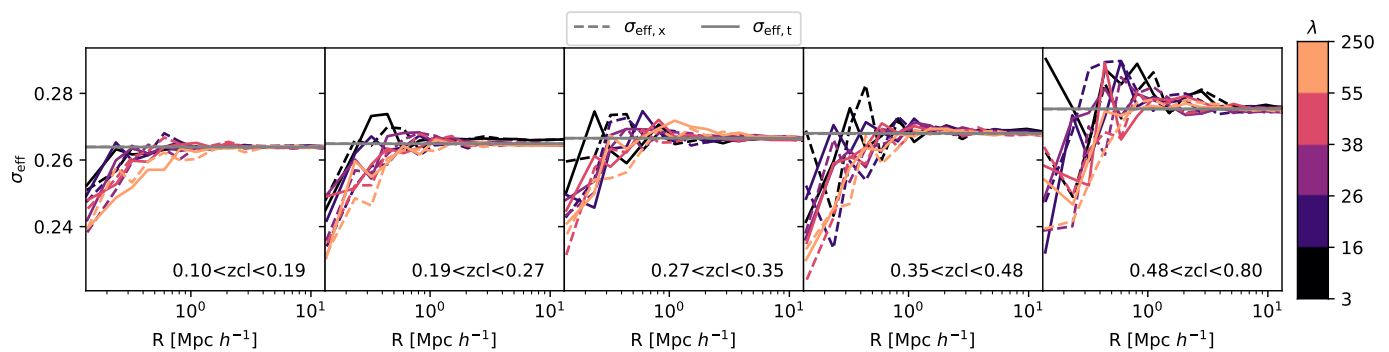
Fig. 7. Effective shape noise as a function of cluster-centric distances stacked in bins of cluster redshift (panels: lowest redshift left, higher redshift right) and of richness (color coding). Despite significant scatter, increasing towards the cluster center, a trend of lower shape noise towards the center is visible. As grey horizontal lines, the effective shape noise in the field, used for the semi-analytical covariance of the shear profiles is shown.
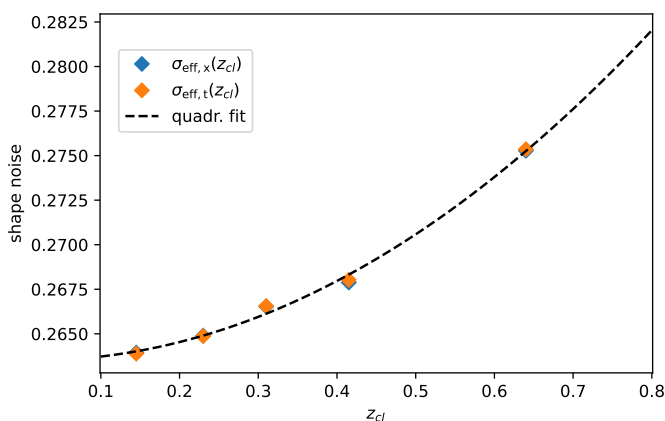

Fig. 8. Cluster redshift trend of the effective shape noise in the field as blue and orange diamonds for the cross and tangential reduced shear respectively. Given our cluster-redshift-dependent background selection, the mix of source galaxies changes as a function of cluster redshift. The trend is well fitted by a quadratic expression, shown as a black dashed line.

$$r_{g_t,s} = \frac{\sqrt{\mathrm{Var}_{\mathcal{H}}\left[\sum_b w_b \int \mathrm{d}z_s\, P_{\mathcal{H}}(z_s|b)\Sigma_{\mathrm{crit,ls}}^{-1}\right]}}{\mathrm{E}_{\mathcal{H}}\left[\sum_b w_b \int \mathrm{d}z_s\, P_{\mathcal{H}}(z_s|b)\Sigma_{\mathrm{crit,ls}}^{-1}\right]}, \tag{17}$$

which we take to express the impact of shear and photo-z measurement uncertainties on the measured shear profiles. It increases from 0.55% at $z_{cl} = 0.1$ to 4.3% at $z_{cl} = 0.8$. We plot these systematic uncertainties around the best-fit values for our cluster member contamination from the effective source density in Fig. 5. The difference between our different cluster member contamination estimates is smaller than this systematic uncertainty. We therefore conclude that we can determine the cluster member contamination to a higher accuracy than provided by the shape and photo-z measurement uncertainties.

3.3.3. Shape noise

In Fig. 7 we show the measured effective shape noise as a function of cluster redshift and richness, and as a function of cluster-centric projected distance. The effective shape noise increases with cluster redshift and richness, and decreases towards the

cluster center. In conjunction with the cluster member contamination estimate from the previous section, we can reconstruct the effective shape noise of cluster members from this decrement. To this end, we fit the measured effective shape noise with a squared sum of field shape noise determined in the outskirt of the clusters, and an unknown cluster-member shape noise. The latter is found to be between 18% (for $0.1 < z_{cl} < 0.19$) and 6% percent (for $0.48 < z_{cl} < 0.80$) lower than the field shape noise, with the trend steadily decreasing with redshift. Given that most of the signal-to-noise in cluster mass measurements comes from scales larger than $R > 1\ h^{-1}\mathrm{Mpc}$, we ignore this effect. In future work, we aim to investigate if this is due to cluster member galaxies being inherently rounder, if the survey-averaged shear response is impacted by the stronger blending in the crowded cluster lines of sight, or if cluster members are just brighter, and therefore have a smaller shape measurement uncertainty contribution to the effective shape noise.

For our shape noise modelling, we instead focus on these larger scales, where we measure that the effective shape noise increases with source redshift, as shown in Fig. 8, both for the tangential and the cross component. This trend matches the survey averaged effective shape dispersion found by Friedrich et al. (2021) for the tomographic redshift bins. For later use in our semi-analytical covariance matrix modelling, we fit this trend with the quadratic expression

$$\sigma_{\mathrm{eff}}(z_{cl}) = 0.2635 + 0.0300\, z_{cl}^2 - 0.0008\, z_{cl}. \tag{18}$$

This fit is shown as a dashed line in Fig. 8, and excellently matches the trend we measure in the data. Extrapolated to $z_{cl} \to 0$ it matches closely the value reported by Gatti & Sheldon et al. 2021. In absence of a reliable estimate for the uncertainty of the effective shape noise, we forego reporting error bars on this fit, using it as an effective smoothing to avoid noise on the error estimate of our target quantity, the reduced shear profile.

3.3.4. Cross-shear

Using our fitting formula for the effective shape noise, for each bin in richness and lens-redshift we can compute the consistency of the cross-shear signal with zero as

$$\chi_x^2 = \sum_{r\text{-bins}} \frac{g_x^2 N_{\mathrm{eff}}}{\sigma_{\mathrm{eff}}^2}. \tag{19}$$

These squared losses follow a chi-squared distribution with n_r degrees of freedom, where n_r is the number of radial bins we
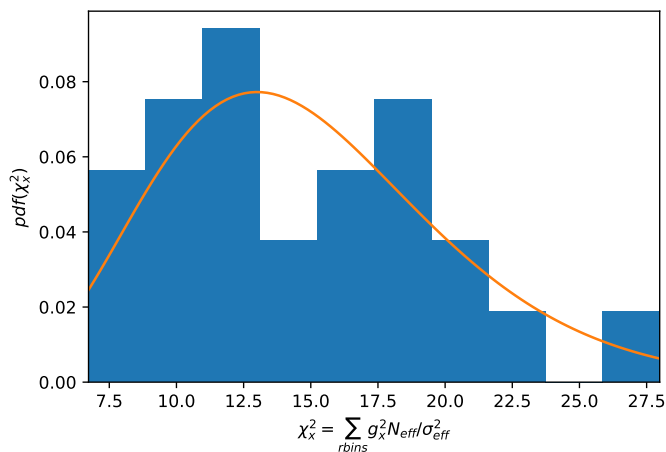
Fig. 9. Distribution of the uncertainty scaled squared residuals of the cross-shear profile w.r.t. zero signal, as a blue histogram. In orange, the chi-squared distribution for 15 degrees of freedom, matching the number of radial bins. The squared losses distribution is well described by the chi-squared distribution, indicating that the cross-shear profiles are consistent with zero given our shape noise modelling.

consider, as shown in Fig. 9. This demonstrates that our cross-shear component averages to zero, consistently with the expectation of well-calibrated shape measurements distorted by gravity. We also visually inspect the cross-shear profiles in the above defined richness, redshift bins, and find no significant deviations from zero. Computing instead the global degree of consistency of the raw tangential reduced shear profiles with zero, we find a signal-to-noise of 92.[‡]

3.3.5. Selection response

The quantities that we use for our source selection are impacted by the shear we try to measure. For instance, the flux of a sheared galaxy is different from the flux of the original galaxy if it was not sheared. This means that the shear we try to measure impacts our selection of source galaxies, and therefore biases our estimators. To correct for this bias, we have to measure how our estimator fares on catalogs selected from artificially sheared versions of the actual DES Y3 images. Four such catalogs have been created, each with an artificial shear of $\Delta\gamma = 0.01$ positive/negative (+/-) on the first/second Cartesian component. The selection response is given by

$$\mathcal{R}_{\alpha,\text{sel}} = \frac{\mathcal{R}_{\alpha,\text{sel},1} + \mathcal{R}_{\alpha,\text{sel},2}}{2} =$$
$$= \frac{1}{2\Delta\gamma}\left(\langle g_{\alpha,\text{raw}}\rangle^{1+} - \langle g_{\alpha,\text{raw}}\rangle^{1-} + \langle g_{\alpha,\text{raw}}\rangle^{2+} - \langle g_{\alpha,\text{raw}}\rangle^{2-}\right),$$
(20)

for $\alpha \in (t, x)$. Here $\langle g_{\alpha,\text{raw}}\rangle^{1+}$ denotes the tangential reduced shear estimator on the source sample extracted from the images that have been artificially sheared in positive (+) direction along the first (1) Cartesian shear component, and similarly for the negatively sheared images (-), and the second Cartesian component (2), respectively. The radial, richness and redshift trends

‡Here we use the definition of signal-to-noise S/N $= \sqrt{\sum \chi_t^2}$, summed over all bins. This quantity does not enter our analysis, and is simply used to give a rough estimate of the constraining power of our dataset.

of the selection response are shown in Fig. 10. Especially for high cluster redshift bins, the selection response is very noisy such that no evident trends can be made out. A reliable estimate can thus only be obtained after averaging over richness and radius. We find that the global selection response is positive and bounded by $\mathcal{R}_{\alpha,\text{sel}} < 0.001$ for $z_{\text{cl}} < 0.48$, and $\mathcal{R}_{\alpha,\text{sel}} \sim 0.004$ for $0.48 < z_{\text{cl}} < 0.8$.

We directly compare the selection response values to the mean shear responses given by

$$\mathcal{R} = \frac{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^s \mathcal{R}_i}{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^s} \text{ for } \alpha \in (t, x),$$
(21)

shown in Fig 11.

Notably, the shear response increases towards the cluster center, while showing also a gentle trend with richness. This signal seems to affect all lens redshift bins equally, independently of their cluster member contamination levels. Another contributing effect could be that cluster member contaminants have different sizes and signal-to-noise ratios compared to the field galaxies. We will account for changes in shear response in the cosmology-ready data products, as shown below. The shear response used here is a smooth function of the source relative size w.r.t. the PSF and its signal-to-noise, and might therefore be impacted by the magnification effect. We plan to investigate how these properties are impacted by cluster lines of sight in future work.

Direct comparison between the mean shear responses and the selection response shows that the latter is of order 0.1% of the former for $z_l < 0.4$, increasing to 0.5% for the redshift bin $0.48 < z_l < 0.8$. In both cases, this is a factor of a few less than the uncertainty induced by shape and photo-z measurements, and as such can be ignored. This assessment is in contrast to the impact of the selection response on galaxy-galaxy lensing studies (Prat et al. 2022). The main difference is that we do not consider each tomographic bin independently, but rather make a selection based on the lens redshift. For low lens redshifts, $z_l < 0.48$, we thus include a large fraction of the total source sample, reducing the impact of selection effects compared to a tomographic redshift bin selection. At high lens redshift, $0.48 < z_l < 0.8$, the close proximity of the source sample to the lens increases dramatically the effects of photo-z uncertainties on the lensing efficiency, dwarfing the effect of the selection response.

3.4. Cosmology-ready data products

Given the various distance dependencies of the WL signal, the mass calibration is cosmology dependent. To marginalise over this dependency in this work, and to enable the WL mass calibration of our sample self-consistently with the cosmological number counts experiment in Ghirardini et al. (A&A, subm.), we need to define cosmology independent data products. Furthermore, given that we pursue a Bayesian population analysis in which each individual cluster likelihood is evaluated, we construct the following data products for each cluster in the cosmology sample that has DES Y3 lensing information:

– the measured reduced shear profile

$$\hat{g}_t = \frac{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^s e_{t,b,i}}{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^s \mathcal{R}_i},$$
(22)

in radial bins with $0.5h^{-1}$ Mpc $< R < 3.2(1 + z_{\text{cl}})^{-1} h^{-1}$Mpc, in our reference cosmology. The outer bin is chosen following Grandis et al. (2021a) to limit our extraction to the 1-halo term region. We use only the shear response $\mathcal{R}_i$, and ignore the selection response, as argued in Section 3.3.5;
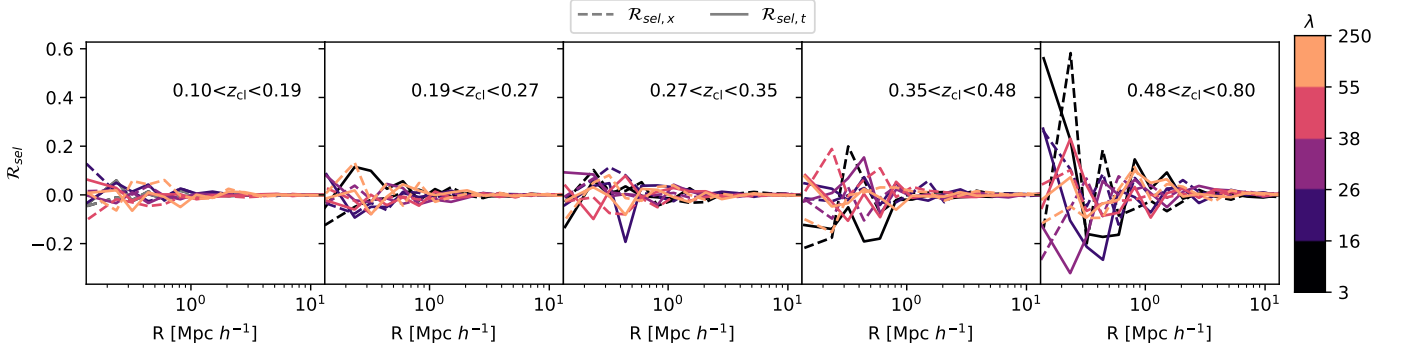
Fig. 10. Selection response of our tangential (full lines) and cross (dashed lines) shear estimator, averaged over artificial shears in the two Cartesian coordinates, as a function of cluster-centric distance, stacked in bins of cluster richness (color coded), and cluster redshift (panels: lowest redshift left, higher redshift right). The selection response is very noisy, but shows no significant trends with richness or cluster-centric distance.
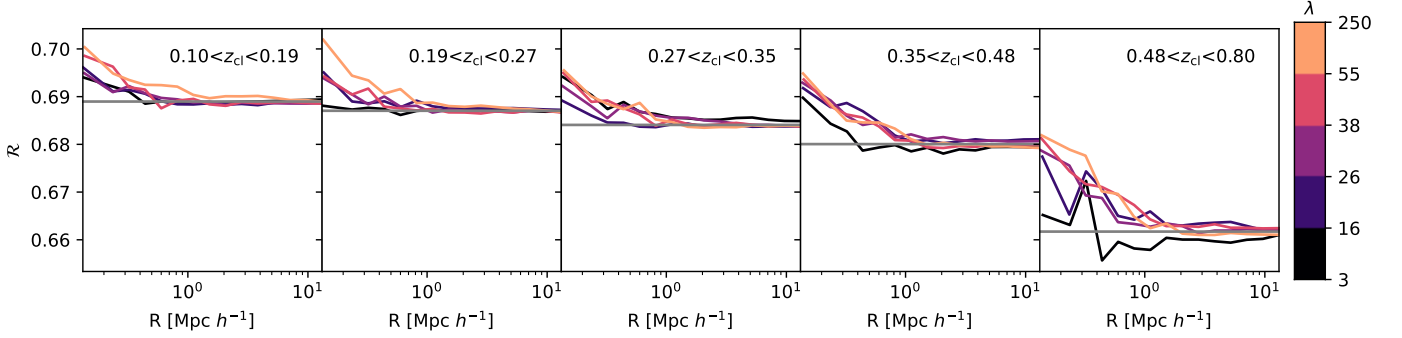

Fig. 11. Average shear response, as a function of cluster-centric distance, stacked in bins of cluster richness (color coded), and cluster redshift (panels: lowest redshift left, higher redshift right). Noticeably, the shear response increases towards the cluster center. Interestingly this signal seems to affect all lens redshift bins equally, independently on their cluster member contamination levels. We account for this effect in the cosmology ready data products.

Table 1. Overview of the simulation inputs used for the calibration of the WL bias and scatter.

name	use case	Section
TNG300	projected surface mass density of massive halos at different redshift snapshots	4.1
Magneticum Pathfinder	offsets between X-ray surface brightness peak and projected halo center	4.2.1
eROSITA all-sky survey twin	offsets between X-ray surface brightness peak and measured X-ray position	4.2.2
	mapping between halo mass and redshift, and detection likelihood and extent	4.4
Monte Carlo realisations	propagation of systematic uncertainties to error budget on the WL bias	4.5

The respective references, as well as detailed explanations, are given in the respective sections.

– the statistical uncertainty on the reduced shear,

$$\delta g_{\rm t} = \frac{\sigma_{\rm eff}(z_{\rm cl})}{\sqrt{N_{\rm eff}}} \text{ with } N_{\rm eff} = \frac{\left(\sum_{b=2,3,4} \sum_{i\in b} w^b w_i^{\rm s} \mathcal{R}_i\right)^2}{\sum_{b=2,3,4} \sum_{i\in b} \left(w^b w_i^{\rm s} \mathcal{R}_i\right)^2}, \quad (23)$$

in the same bins. Note here that we use the quadratic fit to the global shape noise given in Eq. 18 to suppress the noise in the individual shear variance estimators;

– the average angular distance of the binned sources from the cluster position in the corresponding radial bins

$$\theta = \frac{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^{\rm s} \mathcal{R}_i \theta_i}{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^{\rm s} \mathcal{R}_i}. \quad (24)$$

We report the angular scale, because in the cosmological inference, the distance–redshift relation needs to be self-consistently re-calculated, while the data binning is frozen; and

– the field source redshift distribution extracted in the outer regions of the cluster ($11h^{-1}$ Mpc $< R < 15h^{-1}$Mpc) based on the SOM-Pz redshift estimation method,

$$P(z_{\rm s}) = \frac{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^{\rm s} \mathcal{R}_i P(z_{\rm s}|\hat{c}_i)}{\sum_{b=2,3,4} w^b \sum_{i\in b} w_i^{\rm s} \mathcal{R}_i}, \quad (25)$$

where $\hat{c}_i$ is the SOM cell in which the respective source falls. This extraction ensures that our source redshift distribution is a representation of the local field sources.

Using these scale cuts, we reduce the signal-to-noise ratio in the tangential reduced shear to 65. These cuts are necessary to avoid the increased systematic uncertainty towards the cluster centers (see Grandis et al. 2021a, Section 3.2), and to limit our WL observable to the 1-halo regime. As such, we avoid the 2-halo regime with its more complicated cosmology dependence and possible effects of assembly bias.

4. Calibration of mass extraction

Proper cosmological exploitation of the data products derived above must accurately map the WL signal to the halo mass used in the number counts experiment. In this work, we follow the approach of Grandis et al. (2021a) by simulating synthetic shear profiles starting from hydro-dynamical simulations, WL survey specifications and characteristics of the lens sample. The actual mass measurement is performed on the synthetic shear profiles, resulting in the so-called *WL mass* (hereafter WL mass), which displays a bias and scatter w.r.t. the true halo mass. Priors on the WL bias and scatter are derived from Monte-Carlo realisations of the synthetic shear profiles, which sample the range of systematic uncertainties on the input specifications. The different simulation inputs used for this calibration are summarized in Table 1.

4.1. Hydro-dynamical simulations input

We base the creation of our synthetic shear profiles on the surface mass maps of halos with $M_{200c} > 3 \times 10^{13} \, h^{-1} \, M_\odot$, extracted with the technique outlined in Grandis et al. (2021a) from the cosmological hydro-dynamical TNG300 simulations (Pillepich et al. 2018; Marinacci et al. 2018; Springel et al. 2018; Nelson et al. 2018; Naiman et al. 2018; Nelson et al. 2019) at redshifts $z_{snap} = 0.24, 0.42, 0.64, 0.95$. To boost the statistical sample size, for each halo we create three maps by projecting along each Cartesian coordinate. As described in Grandis et al. (2021a), these maps are processed into scaled tangential shear $\Gamma_t(R, \phi | R_{mis})$ and scaled convergence $\Sigma(R, \phi | R_{mis})$ maps, binned in polar coordinates (R, ϕ) around isotropically mis-centered positions for a range of mis-centering distances R_{mis}. $\Sigma(R, \phi | R_{mis})$ is the surface mass density in units of $h \, M_\odot \, \mathrm{Mpc}^2$, which multiplied by the inverse critical surface density gives the convergence, $\kappa = \Sigma_{crit}^{-1}\Sigma$. The definition of Γ_t has been introduced in Grandis et al. (2021a) to signify the result of applying the inverse of Kaiser-Squires algorithm to the surface mass density $\Sigma(x)$, and then determining the tangential component w.r.t. the chosen center. As both these operations are linear, the resulting quantity is directly related to the tangential shear map as $\gamma_t(R, \phi | R_{mis}) = \Sigma_{crit}^{-1}\Gamma_t(R, \phi | R_{mis})$. Being independent from the specific source redshift, this is a convenient simulation output.

This data format conserves the azimuthal anisotropy sourced by the halo triaxiality and by the mis-centering. Possible inaccuracies induced by the correlation between halo shape and miscentring direction, recently highlighted by Sommer et al. (2023), have been shown to not significantly impact the cosmological results derived from the mass calibration presented in this work (see Ghirardini et al. A&A, subm., Section 5.1), but likely need to be accounted for in future analyses with larger statistical constraining power.

Note also that we employ in this work the strategy to mitigate hydro-dynamical modelling uncertainties proposed by Grandis et al. (2021a), and used by Chiu et al. (2022, 2023); Bocquet et al. (2023). The hydro-dynamical simulation from which we source the surface density maps have gravity-only twin runs with the same initial conditions. It is well established that both runs form exactly the same halos (Castro et al. 2021; Grandis et al. 2021a), albeit with different masses. We purposefully choose the fictitious mass in the gravity-only simulation, as the halo mass functions used in the number counts experiment are calibrated on gravity-only simulations. We can thus use these calibrations to absorb the effects of hydro-dynamical feedback on
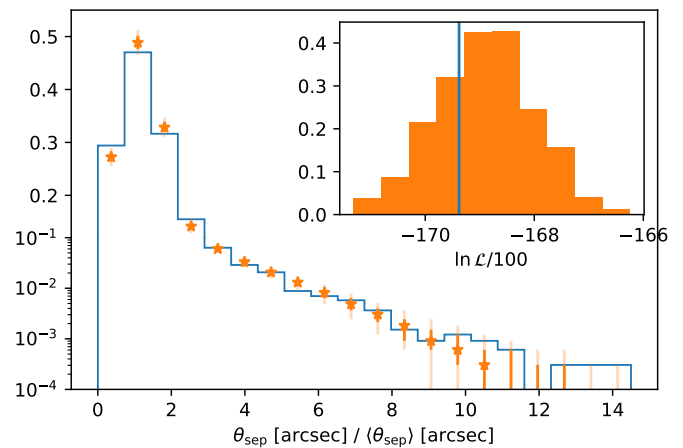

Fig. 12. Observational mis-centering between the input X-ray surface brightness peaks and the output positions of clusters in the eROSITA digital twin, θ_{sep}, scaled by the best-fit average offset, $\langle\theta_{sep}\rangle$, shown in blue as a histogram. Overlaid in orange are realizations of mock offsets drawn from the posterior of our model. The data is indistinguishable from the mock draws. Our model is also validated by comparing, in the insert, the likelihood of the data in blue with the likelihoods of the mocks, orange.

the mass distributions of massive halos in the mapping between the (gravity-only) halo mass and the WL mass, and account for the uncertainty in this mapping in the systematic error budget (see Section 4.5).

4.2. Mis-centering

A faithful simulation of synthetic shear profiles requires us to account for the fact that the observed position of the lens is mis-centered from the projected position of the true halo center, which is defined consistently through this work as the position of the most-bound particle. The surface mass maps are centered around this position (Grandis et al. 2021a).

We distinguish two sources of mis-centering: 1) the offset between the noise-free peak of the 2d X-ray surface brightness profile and the 2d position of the halo center (hereafter called *intrinsic*), and 2) the offset between the noise-free peak of the 2d X-ray surface brightness profile and the measured X-ray position (hereafter called *observational*). The latter displacement is induced by the photon shot noise and the PSF of the eROSITA cameras and is a special case of the astrometric uncertainty discussed in Merloni et al. (2024).

4.2.1. Intrinsic mis-centering

The intrinsic mis-centering is studied through the use of *Box2b/hr* of the hydro-dynamical cosmological simulation suite MAGNETICUM PATHFINDER[§] (Dolag et al. in prep.), performed assuming the WMAP-7 cosmology as $\Omega_0 = 0.272$ and $h = 0.704$ as given by Komatsu et al. (2011). The volume of the *Box2b/hr* box is $(909 \, \mathrm{cMpc})^3$, allowing for a broad range of galaxy cluster masses up to virial masses M_{vir} of several times $10^{15}M_\odot$. The galaxy cluster sample, comprised of 116 clusters with $M_{vir} \geq 10^{14}M_\odot$ at $z = 0.252$, is described in detail in Kimmig et al. (2023). We select an additional 75 galaxy clusters at $z = 0.518$ to account for redshift trends. The second sample is taken from

---

[§]www.magneticum.org

a different box of equal resolution, *Box2/hr* of size $(500\,\text{cMpc})^3$, to avoid double counting the halos at different redshifts.

Every cluster is projected from 100 random viewing angles. We then determine the projected center-of-mass via the shrinking sphere method (Power et al. 2003) and around this center the peak of the X-ray emission between $0.5 - 2\,\text{keV}$. The intrinsic mis-centering is then the projected displacement between the X-ray peak and the center of the halo, given by the most-bound particle.

We find that the mis-centering is well described by a single Rayleigh distribution with mean $\left\langle R_{\text{mis}}^{\text{intr}} \right\rangle = \sigma_{\text{intr}} R_{500c}$, with $\sigma_{\text{intr}} = 0.104 \pm 0.016$, and with no further resolved mass or redshift trends. As shown below in Section 4.5 and Fig. 14, this uncertainty is not relevant when taken together with other systematic effects, especially photometric source redshift and hydrodynamical modelling uncertainties. Valid concerns about the fidelity of the Magneticum prediction of the surface brightness in inner cluster regions are therefore unlikely to impact our results. In future analyses, we plan to validate the centering choice more carefully, for instance through the blind comparison of the results based on different centers (as done for instance by Bocquet et al. 2023).

4.2.2. Observational mis-centering

For the observational mis-centering, we use the offset θ_{sep} between input and output positions of clusters in the "eROSITA all-sky survey twin" (Comparat et al. 2019, 2020; Seppi et al. 2022), which produces and analyses simulations of the X-ray signature from a realistic active galactic nuclei and cluster population in eRASS1 observational conditions that include exposure time variations and background. Furthermore, the X-ray cluster finding algorithm, as described in Merloni et al. (2024); Bulbul et al. (A&A, subm.) is run on the event file generated from this simulation, ensuring that PSF and shot noise of the real observations are well mapped. We select the input clusters matched with simulated sources that have detection likelihood $\mathcal{L}_{\text{det}} > 3$, extent likelihood $\mathcal{L}_{\text{EXT}} > 6$, and extent $4 \leq \text{EXT} \leq 60$, to match the real cosmology sample. For each such object, we compute the separation θ_{sep} between input and output position, using the entries described in Table A.1 by Seppi et al. (2022).

We model the distribution of mis-centering with a mixture of a Rayleigh distribution and a Gamma distribution with shape parameter 2, as

$$P\left(\theta_{\text{sep}} | \left\langle \theta_{\text{sep}} \right\rangle, f, \lambda\right) = \frac{1}{1+f} \frac{\theta_{\text{sep}}}{\left\langle \theta_{\text{sep}} \right\rangle^2} \exp\left(-\frac{1}{2} \frac{\theta_{\text{sep}}^2}{\left\langle \theta_{\text{sep}} \right\rangle^2}\right)$$
$$+ \frac{f}{1+f} \frac{\lambda^2 \theta_{\text{sep}}}{\left\langle \theta_{\text{sep}} \right\rangle^2} \exp\left(-\frac{\lambda \theta_{\text{sep}}}{\left\langle \theta_{\text{sep}} \right\rangle}\right), \quad (26)$$

with the mean observational mis-centering given by

$$\left\langle \theta_{\text{sep}} \right\rangle = \left\langle \theta_{\text{sep}} | \alpha, \sigma_0, k \right\rangle = \left(\frac{\mathcal{L}_{\text{det}}}{38.7}\right)^{-\alpha} \sqrt{\sigma_0^2 + k^2 \text{EXT}^2}, \quad (27)$$

where $\mathcal{L}_{\text{det}}$ is the detection likelihood, EXT the extent, as defined by the best-fit detection beta model, and $\boldsymbol{p}_{\text{mct}} = (\alpha, \sigma_0, k, \ln f, \ln \lambda)$ are free parameters. f is the fraction of mis-centered objects in the heavy tailed component, α the detection likelihood trend of the mean observational mis-centering, σ_0 the extrapolated mean observational mis-centering for point

source (EXT = 0), k the slope of the mean mis-centering w.r.t. the sources extent, and λ the length of the mis-centering tail in units of the mean mis-centering.

We fit these by sampling the likelihood

$$\ln \mathcal{L} = \sum_v \ln P\left(\theta_{\text{sep}}^v | \left\langle \theta_{\text{sep}}^v | \alpha, \sigma_0, k \right\rangle, \ln f, \ln \lambda\right), \quad (28)$$

assuming flat priors on all parameters, with v running over the clusters selected from the digital twin simulation. All parameters are well constrained, their means and standard deviations are reported in Table 2. Noticeably, the mean dependence on the detection likelihood $\alpha = 0.43 \pm 0.02$ is quite close to the $\alpha = 0.5$ trend one would expect for matched filter extraction in the presence of Gaussian noise (Story et al. 2011; Song et al. 2012). Our constraint on $\ln f$ translates into a relative weight of 0.318 ± 0.043 for the flatter Gamma distribution component, albeit only a part of that distributions populates the high mis-centering tail. Our other fit parameters result in a typical mis-centering $\left\langle \theta_{\text{sep}} \right\rangle \sim 11$ arcsec.

We also check that once the individual mis-centerings are corrected for the typical mis-centering of sources of that detection significance and extent, the residuals show no correlation with exposure time. This ensures that our parametrisation captures the main trends of the observational mis-centering, and that it can be safely used despite the exposure time varying over the survey footprint. To further validate our model, we draw mock catalogs, by selecting a random posterior point $\boldsymbol{p}_{\text{mct}} = (\alpha, \sigma_0, k, \ln f, \ln \lambda)$. For each simulated cluster we then compute the typical mis-centering $\left\langle \theta_{\text{sep}} \right\rangle$ and draw a mock separation angle θ_{sep} from Eq. 26. With this procedure we create 1000 mock cluster samples that follow our model, by drawing 1000 independent posterior points. In Fig. 12 we plot in blue the mis-centering distribution of the actual data, while in orange data points the distribution of mock data, sampled from our posterior, are shown. The two are indistinguishable, indicating that the actual mis-centering distribution from the digital twin looks like data drawn from our mis-centering model. Comparisons of the maximum likelihood of the data and the mocks show that this holds also at the likelihood level (see insert of Fig. 12). Our observational mis-centering model is therefore a good fit.

4.3. Extraction model

When performing the mass measurement, we need to specify a mass extraction model and coherently apply it to both the real and the synthetic data. In this analysis, we choose a simplified mis-centered model, corrected for the mean cluster member contamination, as suggested by Grandis et al. (2021a). This model faithfully captures the main biases induced by mis-centering and cluster member contamination, while not adding numerical complexity compared to plain 2d-projected NFW model. This approach is well suited for the repeated calls in the Monte Carlo Markov chains used for parameter inference. Bocquet et al. (2023) make the same choices for the same reasons.

For the model, we assume that a cluster of radius R_{500c} and mean observational mis-centering $\left\langle \theta_{\text{sep}} \right\rangle$ is effectively mis-centered by a physical distance

$$R_{\text{mis}}^{\text{extr}} = \sqrt{\left(\sigma_{\text{intr}} R_{500c}\right)^2 + \left(D_l \left\langle \theta_{\text{sep}} \right\rangle\right)^2}, \quad (29)$$

evaluated at the mean of the observational mis-centering parameters determined in section 4.2.2. The radius R_{500c} is derived from

the input mass of the extraction model. The cluster surface mass distribution is then well approximated by

$$\Sigma(R|M) = \begin{cases} \Sigma_{\mathrm{NFW}}(R_{\mathrm{mis}}^{\mathrm{extr}}|M) & \text{for } R < R_{\mathrm{mis}}^{\mathrm{extr}}, \text{ and} \\ \Sigma_{\mathrm{NFW}}(R|M) & \text{otherwise,} \end{cases} \quad (30)$$

where $\Sigma_{\mathrm{NFW}}(R|M)$ is the 2d projected NFW profile, evaluated with the mean concentration mass relation reported by Ragagnin et al. (2021). Deviations of this concentration mass relation from the one in the simulations are absorbed in the WL bias, while scatter in concentration at fixed mass contributes to the WL scatter. As shown in Bocquet et al. (2023), this surface mass profile results in a radial density contrast $\Delta\Sigma(R|M) = 0$ for $R < R_{\mathrm{mis}}^{\mathrm{extr}}$, and

$$\Delta\Sigma(R|M) = \Delta\Sigma_{\mathrm{NFW}}(R|M) - \left(\frac{R_{\mathrm{mis}}^{\mathrm{extr}}}{R}\right)^2 \Delta\Sigma_{\mathrm{NFW}}(R_{\mathrm{mis}}^{\mathrm{extr}}|M), \quad (31)$$

at larger projected radius. This expression is numerically of the same complexity as evaluating the 2d density contrast of NFW profile $\Delta\Sigma_{\mathrm{NFW}}(R|M)$, but is significantly more accurate.

We then consider the average lensing efficiency

$$\Sigma_{\mathrm{crit,l}}^{-1} = \left\langle \Sigma_{\mathrm{crit,ls}}^{-1} \right\rangle_{P_l(z_s)} \quad (32)$$

of the lens l, obtained by averaging the lensing efficiency of the individual source, $\Sigma_{\mathrm{crit,ls}}^{-1}$, Eq. 3, with source redshift distribution $P_l(z_s)$. Using also the cluster member contamination fraction $f_{\mathrm{cl}}(R|\lambda, z)$, evaluated at the mean cluster member contamination parameters, for the cluster richness λ and the cluster redshift z_{cl}, we define the reduced shear model

$$g_{\mathrm{t}}^{\mathrm{mod}}(R|M) = \frac{\Sigma_{\mathrm{crit,l}}^{-1}\Delta\Sigma(R|M)}{1 - \Sigma_{\mathrm{crit,l}}^{-1}\Sigma(R|M)} \left(1 - f_{\mathrm{cl}}(R|\lambda, z_{\mathrm{cl}})\right). \quad (33)$$

This extraction model depends also on the clusters detection likelihood $\mathcal{L}_{\mathrm{det}}$ and extent EXT, and the mean parameters of the mis-centering through the mean extraction mis-centering, eq. 29. All these quantities are readily available for each cluster.

Any mis-matches between the real profiles and the extraction model are captured by considering that the best-fit mass assuming this model, the *WL mass* M_{WL}, will be biased and scattered around the true halo mass.

4.4. Synthetic shear profiles

The creation of the synthetic shear profiles proceeds as follows.

- The redshift of the synthetic clusters is fixed to the redshift of the hydro-simulation outputs.
- Assuming the mean concentration c_{200c}–mass relation by Child et al. (2018), we convert the masses M_{200c} from the simulation to M_{500c}, using the customary transformations (see for instance Ettori et al. 2011, App. A). Note that the concentration–mass relation employed here was calibrated on gravity-only simulations, which matches the fact that our halo masses are also gravity-only masses.
- We assign a synthetic source redshift distribution $P_l^{\mathrm{snth}}(z_s)$ based on the tomographic bin weights $w_b(z_l)$, as defined in Eq. 7, and the realisations of their redshift distribution $P_{\mathcal{H}}(z_s|b)$, as provided by Myles & Alarcon et al. 2021, discussed in Section 3.3.2, and shown in Fig. 2. The normalisation of these distributions takes account of the multiplicative shear bias.

- Each cluster is assigned a richness λ based on the richness–mass relation calibrated in Chiu et al. (2022). For later use, we remind the reader that this relation is defined by the parameters of the richness–mass relation $\boldsymbol{p}_{\lambda M}$. The richness is required both as an input for the extraction model, as well as to simulate the cluster member contamination, $f_{\mathrm{cl}}(R|\lambda, z_{\mathrm{cl}}, \boldsymbol{p}_{\mathrm{cl}})$, on the synthetic shear profile, where $\boldsymbol{p}_{\mathrm{cl}}$ are the parameters of the cluster member contamination.
- Querying the halos with $|\Delta \log_{10} M_{500c}/M_\odot| < 0.1$ and $|\Delta z| < 0.05$ from the eROSITA digital twin, we select the detection likelihood $\mathcal{L}_{\mathrm{det}}$ and extent EXT of a random pick. These two observables are needed to define the mean mis-centering needed by the extraction model and to assign a correct mis-centering distribution $P(R_{\mathrm{mis}}|M_{500c}, z, \mathcal{L}_{\mathrm{det}}, \mathrm{EXT}, \boldsymbol{p}_{\mathrm{mct}})$, resulting from the convolution of the observational and intrinsic mis-centering discussed in Section 4.2. Note that this step depends on the parameters of the mis-centering $\boldsymbol{p}_{\mathrm{mct}}$, determined in the same section.
- Improving on previous work, we now compute the reduced shear not only for each polar position in the map, but also for each source redshift z_s, as presented also in Bocquet et al. (2023). The azimuthal and source redshift average are computed after the computation of the reduced shear, as follows

$$g_{\mathrm{t}}^{\mathrm{snth}}(R|R_{\mathrm{mis}}) = \int dz_s P^{\mathrm{snth}}(z_s) \int \frac{d\phi}{2\pi} \frac{\Sigma_{\mathrm{crit,ls}}^{-1}\Gamma_t(R, \phi|R_{\mathrm{mis}})}{1 - \Sigma_{\mathrm{crit,ls}}^{-1}\Sigma(R, \phi|R_{\mathrm{mis}})}. \quad (34)$$

Deviating from this integration order leads to biases in the synthetic profiles of the order of 0.01, which is not acceptable given our accuracy requirements.
- We add cluster member contamination following the the fit we derived in Section 3.3.1, specifically the number density fit. We also add a non-linear shear bias and extra noise from the uncorrelated large scale structure along the line of sight, as discussed in Grandis et al. (2021a).
- The extraction of the WL mass via the extraction model and the fit for the WL bias and scatter weighted by the mis-centering distribution follow exactly the method outlined in Grandis et al. (2021a).

The output data product of such synthetic shear profile production and subsequent fit with the extraction model is shown in Fig 13. There, we plot the input halo masses versus the WL masses, color coded by the mis-centering distribution probability for the respective synthetic halo. Note that extreme outliers in WL mass are associated predominantly with highly mis-centered synthetic clusters, and are thus quite rare given the mis-centering distribution. As such, they contribute only marginally to the fit.

4.5. WL bias and scatter

At the end of the synthetic shear profile production and WL extraction, we get a bias and scatter $(b_l^{\mathrm{WL}}, \sigma_l^{\mathrm{WL}})$ for each snap shot l, as well as a mass trend for both the bias $(b_{\mathrm{M}}^{\mathrm{WL}})$, and the scatter (s_{M}). They are obtained by fitting the relation

$$\left\langle \ln \frac{M_{\mathrm{WL}}}{M_0} \middle| M, z_{\mathrm{cl}} \right\rangle = b_l^{\mathrm{WL}} + b_{\mathrm{M}}^{\mathrm{WL}} \ln\left(\frac{M}{M_0}\right)$$
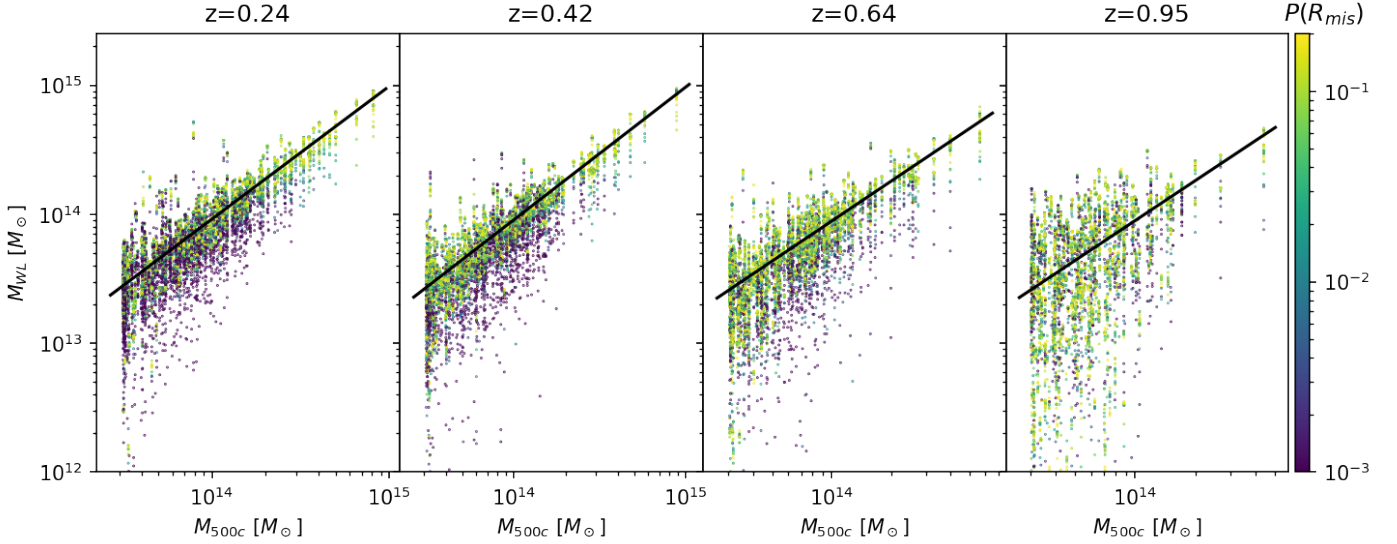$$\ln \sigma_{\mathrm{WL}}^2 = s_l + s_{\mathrm{M}} \ln\left(\frac{M}{M_0}\right), \quad (35)$$

Fig. 13. Scatter plot of the input halo mass versus the output WL mass from one realisation of the synthetic shear profiles, color coded by the value of the mis-centering distribution for that specific halos' mis-centering, and as solid black lines, the mis-centering-distribution-weighted scaling between WL and halo mass, which results in an estimate for the WL bias. The scatter around this relation results in the WL scatter. Generally speaking, extreme outliers have also highly improbable mis-centering.

Table 2. Calibration parameters with their priors, for the Monte Carlo marginalisation of the WL bias and scatter determination.

Cluster Member Contamination (Section 3.3.1)	
$\log_{10} c = 4.83 \pm 0.22$	$B_{cl} = 0.467 \pm 0.017$
$\ln \sigma_{reg} = -0.61 \pm 0.43$	$A_0 = -4.75 \pm 0.11$
$A_1 = -4.38 \pm 0.05$	$A_2 = -3.68 \pm 0.04$
$A_3 = -3.25 \pm 0.04$	$A_4 = -3.04 \pm 0.04$
Mis-centering Parameters (Section 4.2)	
$\sigma_0[\text{arcsec}] = 4.83 \pm 0.22$	$k = 0.308 \pm 0.007$
$\alpha = 0.443 \pm 0.012$	$\ln f = -0.76 \pm 0.11$
$\ln \beta = -0.18 \pm 0.04$	$\sigma_{intr} = 0.104 \pm 0.016$
$\ln \lambda = -0.18 \pm 0.04$	
Richness Mass relation (Chiu et al. 2022, app. A)	
$A_\lambda = 36.2 \pm 3.6$	$B_\lambda = 0.881 \pm 0.083$
$C_\lambda = -0.46 \pm 0.52$	$\sigma_\lambda = 0.274 \pm 0.066$
photo-z and shape calibrations	
$\mathcal{H} \sim \text{Int}(0, 999)^*$	$\alpha_{NL} = 0.6 \pm 0.4$ [†]
hydro-dynamical modelling uncertainties	
$\Delta b = 0.02, \Delta b_M = 0.018$	$\Delta s = 0.25, \Delta s_M = 0.59$

Note that the parameters for the cluster member contamination and the mis-centering are actually drawn from the respective posterior fits to conserve correlations among the parameters. Further notes: [*] Int(a, b) is a uniform distribution over the integers $a \le i \le b$ (cf. Section 3.3.2); [†] non-linear multiplicative shear bias (Grandis et al. 2021a, Section 2.1.7)

Table 3. Calibration of WL bias and scatter for DES Y3 WL on eRASS1 clusters.

z	0.24	0.42	0.68	0.95
μ_b	−0.066	−0.075	−0.101	−0.092
δ_{b1}	−0.013	−0.018	−0.034	−0.092
δ_{b2}	−0.019	−0.022	−0.023	0.015
μ_s	−3.094	−3.232	−2.925	−1.833
δ_s	0.272	0.270	0.275	0.283
$b_M = 1.030 \pm 0.021$			$s_M = -0.919 \pm 0.608$	

Numerical values for the WL bias (μ_b, δ_{b1}, δ_{b2}) and scatter (μ_s, δ_s) calibration at different redshift z, and their global mass trends b_M, and s_M, respectively.

distributions $\mathcal{H} \in (0, 999)$ (Myles & Alarcon et al. 2021, and as outlined in Section 3.3.2), the parameters of the richness–mass relation $\boldsymbol{p}_{\lambda M}$ within the posterior reported by Chiu et al. (2022), the parameters of the cluster member contamination $\boldsymbol{p}_{cl}$ within the posterior determine in Section 3.3.1, the parameters of the mis-centering distribution $\boldsymbol{p}_{mct}$ within the posterior determine in Section 4.2, as well as a wide prior on the non-linear shear biases. To these uncertainties we add a hydro-dynamical modelling uncertainty of 2% on the WL bias, and corresponding values for the other parameters, as determined by Grandis et al. (2021a, Table 2).

The resulting parameter posterior distributions are presented in the contour plots in Fig. 14. We omit parameters that do not (anti)correlate with the WL bias and scatter parameters. As can be seen, the WL bias in the highest redshift slot b_4^{WL} anticorrelates tightly with the critical surface density $\Sigma_{crit,4} = \left\langle \Sigma_{crit,s,l=z_4}^{-1} \right\rangle_{P_{z_4}^{snth}(z_s)}^{-1}$, defined as the inverse lensing efficiency, reported in units of $10^{15} \, h\text{M}_\odot \, \text{Mpc}^{-2}$. This indicates that the uncertainty on this bias parameter is dominated by the uncertainty on the lensing efficiency, as induced by the photo-z uncertainty.

Given these correlations, we compress the posterior on the WL bias into mean values $\mu_{b,l}$, and two principal components $\delta_{b1,l}$ and $\delta_{b2,l}$, scaled by their variances. This means that a realisa-
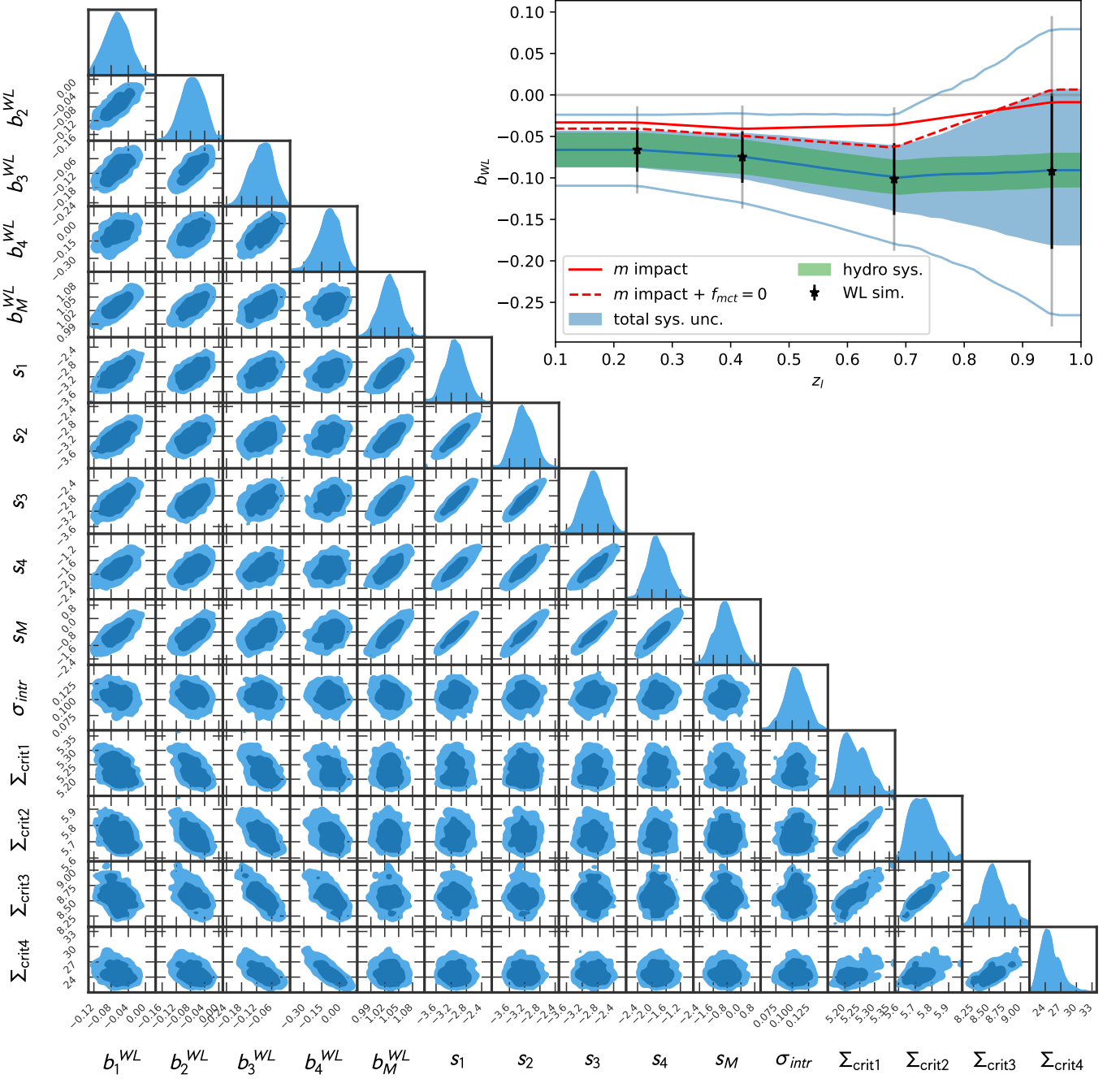
with pivot mass $M_0 = 2 \times 10^{14} \, M_\odot$, and $s_l = \ln(\sigma_l^{WL})^2$. The fit of the relation includes cluster-by-cluster weights corresponding to the probability of the mis-centering of the individual synthetic halo, as discussed in Grandis et al. (2021a). The mean relations resulting from the fit at each redshift are shown in Fig 13 as black lines.

The uncertainty on these quantities is obtained by re-running the synthetic shear profile creation and WL mass extraction $O(1000)$ times with slightly perturbed input parameters. Specifically, we vary the hyper parameter of the DES Y3 source redshift

Fig. 14. *Main panel*: Joint distribution on the WL bias and scatter parameters (b_l^{WL}, $s_l = \ln(\sigma_l^{\mathrm{WL}})^2$), and their mass slopes ($b_{\mathrm{M}}^{\mathrm{WL}}$, σ_{M}), together with the principal sources of systematic uncertainty: the lensing efficiencies $\Sigma_{\mathrm{crit},l}$ to the snapshots l, in units of $10^{15}\,h\,M_\odot\,\mathrm{Mpc}^{-2}$, and the scale of the intrinsic mis-centering σ_{intr}. The high redshift biases anticorrelate with the respective lensing efficiencies, indicating that in this regime we are limited by the DES Y3 photo-z accuracy. Conversely, the low redshift biases, and the scatter values, correlate among themselves, as they are dominated by the hydro uncertainty floor we added. *Insert*: Redshift interpolation of the WL lensing bias in cyan (filled region 1 sigma, faded lines 2 sigma), together with the 1 sigma uncertainty due to hydro modelling (green), explorations of the mean impact of the multiplicative shear bias m modelling (red), and the additional omission of the high mis-centering tail $f = 0$ (dashed red), and the posteriors of the WL bias in the four snap-shots. hydro-dynamical modelling uncertainties are dominant at low redshift. When considering our multiplicative shear bias modelling choices, the bias is consistent with zero, while the presence of the highly mis-centered tail does not affect the WL bias.

tion of the WL bias at an arbitrary redshift z_{cl} can be constructed by interpolating the mean and principal components and drawing two random standard normal variates A_{WL}, B_{WL}, such that

$$b(z_{\mathrm{cl}}) = \mathcal{I}(z_{\mathrm{cl}}|z_l, \mu_{\mathrm{b},l}) + A_{\mathrm{WL}}\mathcal{I}(z_{\mathrm{cl}}|z_l, \delta_{\mathrm{b1},l}) + B_{\mathrm{WL}}\mathcal{I}(z_{\mathrm{cl}}|z_l, \delta_{\mathrm{b2},l}),$$

$$(36)$$

where $\mathcal{I}(z_{\rm cl}|z_l, \mu_{{\rm b},l})$ is the interpolation of the data vectors $(z_l, \mu_{{\rm b},l})$ at redshift $z_{\rm cl}$. The resulting posterior on the WL bias as a function of cluster redshift is shown in the insert of Fig. 14 in blue bands for the $1\,\sigma$ region, and faded lines for the $2\,\sigma$ regions, together with the results for the 4 snapshots as black data points. For the natural logarithm of the variance, whose posterior is dominated by the uncertainty from hydro-dynamical modelling, we employ a single principal component (see below, Eq. 45).

The numerical value of our WL bias is somewhat low, ranging from -0.05 to -0.1. A contributing factor is the response of the WL bias to our treatment of the multiplicative shear bias. While we include the multiplicative shear bias in the synthetic shear profiles, we do not include this bias in the model so as to reduce the number of independent calibration product we need to communicate to the cosmological number counts analysis (Ghirardini et al. A&A, subm.). To gauge the impact of this modelling choice on the numerical value of the bias, we perform a synthetic shear and mass extraction run at the mean values of all parameters, but artificially set the multiplicative shear bias $m = 0$, that is we normalise the synthetic source redshift distributions to 1. This results in a shift, shown as the red line in the insert of Fig. 14 of the order of -0.03. It is two sigma away from our inferred bias values (black data points). This indicates that our extraction model is a reasonable representation of the synthetic data, once corrected for the above described modelling mis-match. Slightly negative WL biases can be explained by the fact that the density contrast of a halo is reduced by the projection of nearby correlated structure, resulting in a WL mass that is lower than the halo mass (Oguri & Hamana 2011; Becker & Kravtsov 2011; Bahé et al. 2012).

We also test the sensitivity of the WL bias to the presence of the highly mis-centered tail in the eROSITA digital twin (cf. Section 4.2). A similar large mis-centering tail is found by Bulbul et al. (A&A, subm.), Section 5 and Fig. 9, left panel, when comparing the centers obtained from the source detection algorithm and the X-ray post-processing. While the origin of this tail is still under investigation, we explore the hypothesis that the large mis-centering tail is completely suppressed by the X-ray post-processing. Setting the fraction $f = 0$ fully suppresses the high exponential tail, but only mildly alters the WL bias, as can be seen by the difference between the red solid and dashed line. Our WL bias numbers, and therefore the mass anchoring of the number counts experiments, are stable w.r.t. the presence of this highly mis-centered tail, which might be an artefact of the eROSITA digital twin processing, and be corrected by the X-ray post-processing.

In summary, the two main contributions to uncertainty of the WL bias are the systematic floor resulting from the comparison of different hydro-dynamical simulations established by Grandis et al. (2021a), and, at high redshift $z_l > 0.6$ also the uncertainty of the DES Y3 photometric redshift distributions, shown in Fig. 2.

5. Mass constraints

WL measurements around galaxy clusters provide a highly accurate (see Section 4) mass proxy. They have, however, a quite low precision at the individual object basis, considering that the approximately 2.2k clusters we consider collectively only reach a signal-to-noise of 65 (cf. Section 3.4). A natural approach to counter-act this is to perform the mass measurement on stacked lensing signals (e.g. McClintock & Varga et al. 2019; Bellagamba et al. 2019). While this simplifies the mass measurement task, it significantly complicates the modelling in a cosmo-

logical context by introducing scale dependent selection bias effects (Sunayama et al. 2020; Dark Energy Survey Collaboration et al. 2020; Wu et al. 2022; Sunayama 2023). In this work, we therefore follow the Bayesian Population Modelling approach pioneered by Mantz et al. (2015); Bocquet et al. (2019); Grandis et al. (2019), which predicts the distribution of expected shear profiles given the selection observable, as a function of the parameters governing the scaling between halo mass and selection observables. Complete derivations of this analysis framework can be found in Ghirardini et al. (A&A, subm.); Bocquet et al. (2023) in the context of eROSITA–selected clusters and SPT–selected clusters, respectively. Our results here are based on the implementation described in Ghirardini et al. (A&A, subm.), Section 4.

5.1. Constraints of the count rate–mass relation

In the following we discuss the likelihood setup used for the mass calibration as well as the resulting mass calibration constraints. In general, WL mass calibration relies on statistically determining the scaling relation between the observables and halo mass (Mantz et al. 2016; Dietrich et al. 2019; Schrabback et al. 2021; Zohren et al. 2022, for instance). In our case, we expand this standard treatment, to accommodate for the residual contamination of our sample by random superpositions of X-ray candidates and optical systems, and, more importantly, mis-classified active galactic nuclei (see Bulbul et al. (A&A, subm.); Kluge et al. (A&A, subm.); Ghirardini et al. (A&A, subm.)). Both mis-classified AGNs and random superpositions are expect to produce far less WL signal compared to clusters, which inhabit massive halos. We therefore employ a *mixture model*, a weighted sum of the different sub-populations, to perform the mass calibration in presence of non-negligible contamination.

5.1.1. Mixture model

For each individual cluster, we construct the probability density function (PDF) of its shear profile $g_{\rm t}$ given the cluster observables count rate $\hat{C}_{\rm R}$, redshift $z_{\rm cl}$ and sky position $\hat{\mathcal{H}}$ as

$$
\begin{aligned}
P(\hat{g}_{\rm t}|\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}}) = {} & f_{\rm C}(\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}})\, P(\hat{g}_{\rm t}|\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}}, C) \\
& + f_{\rm RS}(\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}})\, P(\hat{g}_{\rm t}|{\rm RS}) \\
& + f_{\rm AGN}(\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}})\, P(\hat{g}_{\rm t}|{\rm AGN}, z_{\rm cl}),
\end{aligned}
\tag{37}
$$

a weighted sum of the pdf of the shear profile for clusters (C), random superpositions (RS), and mis-classified active galactic nuclei (AGN), with their respective frequencies $f_\tau(\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}})$ as a function of count rate, redshift and sky position, where $\tau \in ({\rm C, RS, AGN})$. Naturally, $\sum_\tau f_\tau(\hat{C}_R, z_{\rm cl}, \hat{\mathcal{H}}) = 1$, where the sum runs over $\tau \in ({\rm C, RS, AGN})$. As outlined in Bulbul et al. (A&A, subm.); Kluge et al. (A&A, subm.); Ghirardini et al. (A&A, subm.), the shape of the AGN and the random superpositions contributions are calibrated partially on X-ray image simulations, partially with optical follow up of random positions and point sources, and have amplitudes ($f_{\rm RS}$, $f_{\rm AGN}$) that are fitted for using the richness distribution of the sample at a given count rate, redshift and sky position. They find $f_{\rm RS} = 0.0061 \pm 0.0023$ and $f_{\rm AGN} = 0.0462 \pm 0.0038$, which we adopt as priors.

For the WL modelling, we assume that the density contrast of a random source is zero, resulting in their shear profile pdf being

$$\ln P(\hat{g}_t|\text{RS}) = -\frac{1}{2} \sum_k \frac{\hat{g}_{t,k}^2}{\delta g_{t,k}^2}, \tag{38}$$

where the sum runs over the radial bins k, with the measured shear profile $\hat{g}_{t,k}$ defined in Eq. 22, and its shape noise error $\delta g_{t,k}$ defined in Eq. 23. In practice, all objects in our lens sample have a richness $\hat{\lambda} > 3$. At the lens redshift, the random source will thus have a few red galaxies along their lines of sight. Their density contrast is however expected to be small compared to cluster density contrasts.

For AGNs we assume that their density contrast $\Delta\Sigma_{\text{AGN}}(R)$ follows the profile measured by Comparat et al. (2023) for AGN in eFEDS using HSC WL data. We multiply this with the average lensing efficiency $\Sigma_{\text{crit,c}}^{-1} = \langle \Sigma_{\text{crit,cs}}^{-1} \rangle_{P_c(z_s)}$, evaluated as a function of cosmology for the source redshift distribution $P_c(z_s)$ constructed in Eq. 25. The pdf thus reads

$$\ln P(\hat{g}_t|\text{AGN}, z_{\text{cl}}) = -\frac{1}{2} \sum_k \frac{1}{\delta g_{t,k}^2} \left(\hat{g}_{t,k} - \Sigma_{\text{crit,c}}^{-1} \Delta\Sigma_{\text{AGN}}(R = D_c \theta_k) \right)^2, \tag{39}$$

where D_c is the angular diameter distance to the cluster, and θ_k the average angular distance to the sources in the respective radial bin k, as computed in Eq. 24. In practice, the density contrast of AGN is much smaller than that of galaxy clusters, given the lower host halo mass.

For galaxy clusters, the pdf of the shear profile depends on our extraction model g_t^{mod}, and the WL mass M_{WL} defined by this model (cf. Section 4.3). It therefore reads

$$\ln P(\hat{g}_t|\text{C}, z_{\text{cl}}, M_{\text{WL}}) = -\frac{1}{2} \sum_k \frac{1}{\delta g_{t,k}^2} \left(\hat{g}_{t,k} - g_t^{\text{mod}}(R = D_c \theta_k | M_{\text{WL}}) \right)^2, \tag{40}$$

with the extraction model depending on a host of cluster-specific observables, as outlined above. Note that in all the three pdfs we omitted the normalization term, which is the same for all three, and just depends on $\delta g_{t,k}$, thus being constant w.r.t. the parameters of the inference.

The pdf of the shear profile for clusters needs to be marginalised over the distribution of WL masses M_{WL} compatible with the clusters count rate $\hat{C}_R$, redshift z_{cl}, and sky position $\hat{\mathcal{H}}$, that is $P(M_{\text{WL}}|\hat{C}_R, z_{\text{cl}}, \hat{\mathcal{H}}, \boldsymbol{p}_{\text{SR}})$, whose construction and dependence on model parameters $\boldsymbol{p}_{\text{SR}}$ we shall discuss below. The pdf of the shear profile for the cluster component is thus

$$P(\hat{g}_t|\hat{C}_R, z_{\text{cl}}, \hat{\mathcal{H}}, \text{C}) = \int dM_{\text{WL}} P(\hat{g}_t|\text{C}, z_{\text{cl}}, M_{\text{WL}})$$
$$P(M_{\text{WL}}|\hat{C}_R, z_{\text{cl}}, \hat{\mathcal{H}}, \boldsymbol{p}_{\text{SR}}). \tag{41}$$

The distribution of WL masses given the count rate, redshift and sky position is proportional to

$$P(M_{\text{WL}}|\hat{C}_R, z_{\text{cl}}, \hat{\mathcal{H}}, \boldsymbol{p}_{\text{SR}}) \propto \iint dM dC_R P(M_{\text{WL}}, C_R|M, z_{\text{cl}}, \boldsymbol{p}_{\text{SR}})$$
$$P(I|C_R, z_{\text{cl}}, \hat{\mathcal{H}}_i) P(\hat{C}_R|C_R) P(M, z_{\text{cl}}), \tag{42}$$

with $P(\hat{C}_R|C_R)$ modelling the observational uncertainty on the count rate, $P(M, z_{\text{cl}})$ being proportional to the halo mass function, and $P(I|C_R, z_{\text{cl}}, \hat{\mathcal{H}}_i)$ the X-ray incompleteness due to the extent selection. This ensures that we correctly model Eddington bias, which in cluster studies manifests as more low mass objects scattering up to a given observable value, than high mass objects scattering down to same observable values. This happens even when the observable–mass scatter is modelled as symmetric on account of the fact that there are just many more low mass than high mass objects. This fact is quantitatively expressed by the halo mass function. The expression given above is normalized to be a pdf in M_{WL} via integration in M_{WL} and subsequent re-scaling. See Ghirardini et al. (A&A, subm.) for an alternative, but equivalent, derivation of the same likelihood.

5.1.2. Scaling relation

Given their deep potential wells, cluster observables scale tightly with the host halos mass. Leveraging this strong physical prior, we model the distribution of WL mass M_{WL} and intrinsic (that is noise-free) count rate C_R at given halo mass M and redshift z as a bi-variate log-normal, reading

$$P(M_{\text{WL}}, C_R|M, z_{\text{cl}}, \boldsymbol{p}_{\text{SR}}) = \mathcal{N}(\bar{\mu}, \bar{\Sigma}), \tag{43}$$

with mean and variance

$$\bar{\mu} = [\langle \ln C_R|M, z_{\text{cl}} \rangle, \langle \ln M_{\text{WL}}|M, z_{\text{cl}} \rangle]$$

$$\bar{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho_{\text{WL,X}} \sigma_X \sigma_{\text{WL}} \\ \rho_{\text{WL,X}} \sigma_X \sigma_{\text{WL}} & \sigma_{\text{WL}}^2 \end{bmatrix}$$

where $\langle \ln C_R|M, z_{\text{cl}} \rangle$ and $\langle \ln M_{\text{WL}}|M, z_{\text{cl}} \rangle$ are the scaling relations between true mass and count rate and WL mass, respectively, σ_X and σ_{WL} their intrinsic scatters, and $\rho_{\text{WL,X}}$ the correlation coefficient between the two intrinsic scatters. Inclusion of the latter is crucial to marginalise over possible selection effects, like more concentrated halos *simultaneously* having a higher count rate and WL mass compared to other objects at their mass and redshift, or any other effect that might similarly correlate the intrinsic scatters.

As calibrated on dedicated simulations in Section 4, the mean WL mass – halo mass relation is given by

$$\left\langle \ln \frac{M_{\text{WL}}}{M_p} \middle| M, z_{\text{cl}} \right\rangle = b(z_{\text{cl}}) + b_M \ln \left(\frac{M}{M_p} \right), \tag{44}$$

with the pivot value for the mass $M_p = 2 \cdot 10^{14} M_\odot$, the redshift dependent amplitude of the WL bias given by equation 36, and the mass trend $b_M = \mu_{b_M} + C_{\text{WL}} \cdot \sigma_{b_M}$, with mean and standard deviation derived in Section 4.5 and reported in Table 3.

The scatter is similarly modelled as

$$\sigma_{\text{WL}}^2 = \exp \left\{ \mathcal{I}(z_{\text{cl}}|z_l, \mu_{s,l}) + D_{\text{WL}} \mathcal{I}(z_{\text{cl}}|z_l, \delta_{s,l}) \right\}, \tag{45}$$

where $\mu_{s,l}$ and $\delta_{s,l}$ are the mean and standard deviations of the WL scatter in the simulation snapshot l, derived in Section 4.5 and reported in Table 3, and $\mathcal{I}(z_{\text{cl}}|z_l, \mu_{s,l})$ is the interpolation of the data vectors $(z_l, \mu_{s,l})$ to the redshift z_{cl}. Given the tight correlations between the systematic uncertainties in the WL scatters of the different redshift seen in Fig. 14, we opt to model the WL scatter with a single principal component. We also set the mass trend of the scatter to $s_M = 0$ as in this work we are not interested to explore mass trends in the intrinsic scatter of the other observables.

The count rate–mass relation is modelled as

$$\left\langle \ln \frac{C_R}{C_{R,p}} \middle| M, z_{cl} \right\rangle = \ln A_X + b_X(M, z_{cl}) \cdot \ln \frac{M}{M_p} + e_X(z_{cl}) \qquad (46)$$

where $C_{R,p} = 0.1 \text{ s}^{-1}$ is the pivot value for the count rate, $M_p = 2 \cdot 10^{14} M_\odot$ is the pivot value for the mass. $b_X(M, z)$ expresses the mass-redshift dependent slope of the scaling relation, given by

$$b_X(M, z_{cl}) = \left(B_X + C_X \cdot \ln \frac{M}{M_p} + F_X \cdot \ln \frac{1 + z_{cl}}{1 + z_p} \right) \qquad (47)$$

where B_X is the classic standard single value mass slope, C_X allows the slope to be mass dependent, and F_X allows the slope to be redshift dependent, and $z_p = 0.35$ is the pivot value for the redshift. As shown in Grandis et al. (2019), Appendix B, F_X needs to be sampled to retrieve unbiased cosmological results from the number counts. Chiu et al. (2023) explored the impact of allowing for a mass dependence in the slope in a cosmological analysis of the eFEDS sample, finding that it is not necessary to give the scaling relation that freedom. We thus set $C_X = 0$.

The redshift evolution $e_X(z_{cl})$ of the X-ray scaling relation is

$$e_X(z_{cl}) = D_X \cdot \ln \frac{d_L(z_{cl})}{d_L(z_p)} + E_X \cdot \ln \frac{E(z_{cl})}{E(z_p)} + G_X \cdot \ln \frac{1 + z_{cl}}{1 + z_p} \qquad (48)$$

where $E_X = 2$ and $D_X = -2$ are the values predicted by the self-similar model. The last term, with G_X, quantifies the deviation from the self-similar model. $E(z_{cl}) = H(z_{cl})/H_0$ is the dimensionless Hubble factor. The intrinsic scatter of our scaling relation is represented by a single value, σ_X, independent of mass and redshift.

5.1.3. Likelihood setup and priors

As outlined in Ghirardini et al. (A&A, subm.), the total likelihood of our shear profiles $\hat{g}_{t,\varkappa}$ of several clusters, given their count rate $\hat{C}_{R,\varkappa}$, redshift $z_{cl,\varkappa}$ and sky position $\hat{\mathcal{H}}_\varkappa$, where $\varkappa$ runs over the cluster catalog, is given by

$$\ln \mathcal{L}_{WL} = \sum_\varkappa \ln P(\hat{g}_{t,\varkappa} | \hat{C}_{R,\varkappa}, z_{cl,\varkappa}, \hat{\mathcal{H}}_\varkappa). \qquad (49)$$

In light of the derivations above, this is a function of the parameters of the WL scaling relation $\boldsymbol{p}_{WL} = (A_{WL}, B_{WL}, C_{WL}, D_{WL})$. To represent the WL anchoring performed in Section 4, we place independent Gaussian priors on these. The parameters of the count rate–mass relation $\boldsymbol{p}_X = (A_X, B_X, F_X, G_X, \sigma_X)$, instead, are the true target of our analysis. These are sampled within wide flat priors, reported in Table 4. Also the correlation coefficient between the WL and count rate intrinsic scatters is let free, but within the bounds $\rho_{WL,X} \in (-0.9, 0.9)$. We purposefully do not sample all the way to ± 1, as the covariance of intrinsic scatters becomes singular in that limit.

We place priors on the random source fractions and AGN (f_{RS}, f_{AGN}) from our optical follow-up observations. We do not expect the WL to significantly inform these fractions, as the individual shear profiles are hardly significantly different from zero themselves. As such, at an individual object basis, WL just lacks the statistical power to tell clusters apart from AGN or random sources.

We fix all cosmological parameters to their Planck values values, except $\Omega_M = 0.331 \pm 0.038$ from the first Dark Energy

Table 4. Priors on the parameters used in the fitting of the count rate–mass relation.

Parameter	Units	Prior
Cosmology		
Ω_M	-	$\mathcal{N}(0.331, 0.038)$
$\log_{10} A_S$	-	Fixed to -8.635
H_0	$\frac{\text{km}}{\text{s} \text{Mpc}}$	Fixed to 68.06
$\Omega_{b,0}$	-	Fixed to 0.048595
n_S	-	Fixed to 0.9707
w_0	-	Fixed to -1
w_a	-	Fixed to 0
$\sum m_\nu$	eV	Fixed to 0
$\Omega_{k,0}$	-	Fixed to 0
X-ray scaling relation		
A_X	-	$\mathcal{U}(0.01, 3)$
B_X	-	$\mathcal{U}(0.1, 5)$
C_X	-	Fixed to 0
D_X	-	Fixed to -2
E_X	-	Fixed to 2
F_X	-	$\mathcal{U}(-5, 5)$
G_X	-	$\mathcal{U}(-5, 5)$
σ_X	-	$\mathcal{U}(0.05, 2)$
WL mass calibration		
$A_{WL, DES}$	-	$\mathcal{N}(0, 1)$
$B_{WL, DES}$	-	$\mathcal{N}(0, 1)$
$C_{WL, DES}$	-	$\mathcal{N}(0, 1)$
$D_{WL, DES}$	-	$\mathcal{N}(0, 1)$
$\rho_{WL,X}$	-	$\mathcal{U}(-0.9, 0.9)$
Contamination modeling		
f_{AGN}	-	$\mathcal{TN}(0.0412, 0.0040, 0, \infty)$
f_{RS}	-	$\mathcal{TN}(0.0064, 0.0024, 0, \infty)$

Notes. With $\mathcal{U}(\text{min}, \text{max})$ we indicate a uniform distribution between 'min' and 'max'. With $\mathcal{N}(\mu, \sigma)$ we indicate a normal distribution centered on μ and with standard deviation σ. With $\mathcal{TN}(\mu, \sigma, \text{min}, \text{max})$ we indicate a truncated normal distribution centered on μ and with standard deviation σ bounded between 'min' and 'max'.

Survey Supernovae Type Ia results (Dark Energy Survey Collaboration et al. 2019). Several elements of our likelihood, like angular diameter distances and the extraction model, depend on cosmology via the background evolution. The exact cosmological dependence of the WL calibrated number counts is accounted for by re-running the mass calibration presented here together with the number counts. In this work, we wish to provide mass calibration results marginalised over a reasonable range of cosmologies. Our Ω_M-prior allows us to do this in a physically motivated way, while remaining independent of CMB results, with which the cosmological number counts analysis might be in tension.

5.1.4. Scaling relation parameters constraints

We sample the total likelihood for the measured shear profiles with the priors using `ultranest` (Buchner 2021) to create a sample of the resulting posterior. We find that all parameters of interest, the parameters of the count rate–mass relation, are well constrained by our analysis (shown as blue contours in Fig. 15), while all the calibration and other nuisance parameters have posteriors consistent with their priors. The find a constraint for the amplitude of the scaling relation $A_X = 0.88 \pm 0.20$, when considering the mean and standard deviation computed on the
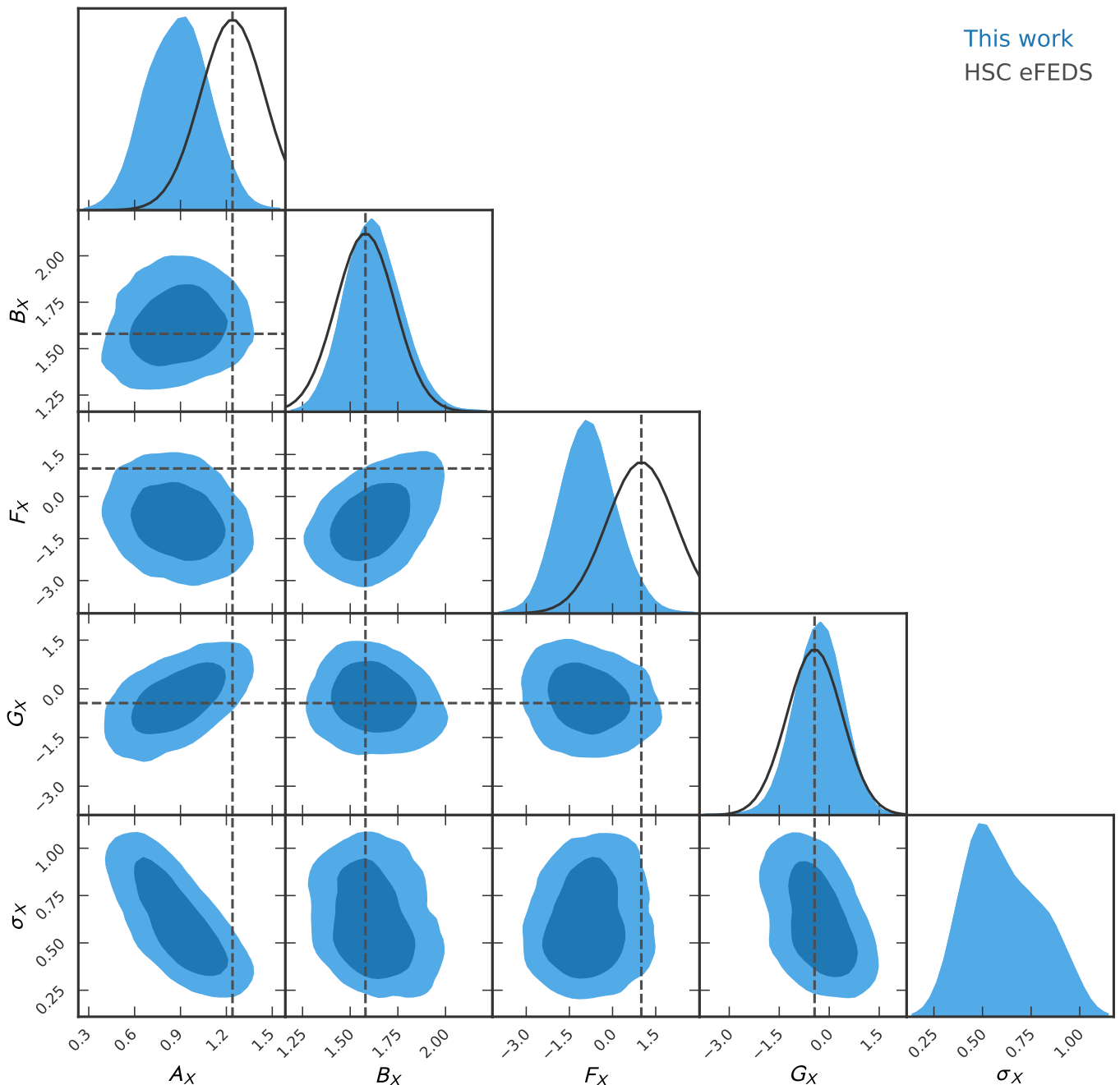
Fig. 15. 1- and 2-d marginal posterior on the parameters of the count rate–mass scaling relations from this work in blue, compared to the 1-d marginals of the analysis of HSC WL around eFEDS clusters and groups (Chiu et al. 2022) in black. We find good agreement on the mass trend B_X, and the deviation from self-similar redshift scaling G_X, while constraints on the amplitude A_X, and the redshift evolution of the mass slope F_X are shifted by 1- to 2-sigmas, consistent with statistical fluctuations. Contrary to the eFEDS-HSC analysis, we do not place an informative prior of the intrinsic scatter σ_X, and get only weak constraints on it.

posterior sample. The mass slope of the relation is constrained to $B_X = 1.62 \pm 0.14$, the redshift trend of the mass slope to $F_X = -0.85 \pm 0.93$. No significant deviation from the self-similar redshift evolution is detected, as $G_X = -0.32 \pm 0.69$ is consistent with 0. For the intrinsic scatter we find $\sigma_X = 0.61 \pm 0.19$. Noticeably, the intrinsic scatter can be constrained, despite the low precision of the individual lensing data, in line the predictions by Grandis et al. (2019). While the intrinsic scatter values is somewhat large, it matches the count rate mass scatter we find in the digital twin simulations (Comparat et al. 2020; Seppi

et al. 2022). To better visualize our scaling relation results, we create a posterior predictive distribution for the mean count rate as a function of mass and redshift. The 1 sigma region of the predicted relation at three redshifts is shown in Fig. 16 in filled bands.

5.2. Goodness of fit

To assess the goodness of fit of our mass calibration, we resort to binning the cluster sample in redshift, count
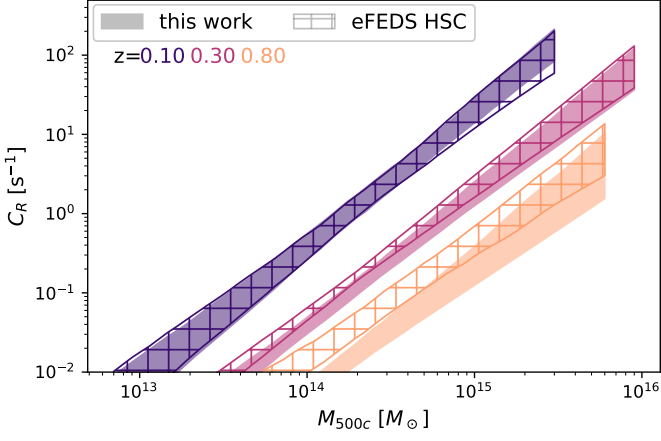
Fig. 16. Posterior predictive distribution of the mean count rate–mass relation (the band ranges from the 16th to the 84th percentile at each mass), for three different redshifts (color coded). Filled bands represent the analysis in this work, hatched bands represent the eFEDS HSC result from Chiu et al. (2022). The predictions diverge for low mass, high redshift systems.

rate bins, with edges $(0.10, 0.19, 0.27, 0.35, 0.48, 0.8)$, and $(0.010.16, 0.25, 0.30, 0.64, 20)$, in redshift and count rate respectively. For all clusters $\varkappa$ falling in each of the bins, we stack the tangential reduced shear profiles by defining the stacking weights $W_\varkappa = \sum_{b=2,3,4} w^b \sum_{i \in b_\varkappa} w_i^s \mathcal{R}_i$, computing

$$\hat{g}_t^{\text{stack}} = \frac{\sum_\varkappa W_\varkappa \hat{g}_{t,\varkappa}}{\sum_\varkappa W_\varkappa}, \tag{50}$$

where $i \in b_\varkappa$ symbolizes the sources i associated to the cluster $\varkappa$ in the tomographic bin b. Combining this expression with Eq. 22 shows that it is equivalent to the raw tangential reduced shear estimator in Eq. 8. Using the same weighting we also derive the error on the stacked reduced shear

$$\left(\delta g_t^{\text{stack}}\right)^2 = \frac{\sum_\varkappa W_\varkappa^2 \delta g_{t,\varkappa}^2}{\sum_\varkappa W_\varkappa^2}, \tag{51}$$

and the mean angular separation of the source galaxies

$$\theta^{\text{stack}} = \frac{\sum_\varkappa W_\varkappa \theta_\varkappa}{\sum_\varkappa W_\varkappa}. \tag{52}$$

The resulting stacked tangential reduced shear profiles are shown as black points with their error-bars in Fig. 17 given by the square root of the variance calculated above.

For the model prediction, we draw $N_{\text{MC}} = 1000$ points $\boldsymbol{p}_{\text{SR}}^v$ from the posterior sample. For each cluster $\varkappa$, and for each posterior point v, we compute its best-guess WL mass

$$M_{\text{WL},\varkappa}^v = \underset{M_{\text{WL}}}{\text{argmax}} \, P(M_{\text{WL}}|\hat{C}_{R,\varkappa}, z_\varkappa, \hat{\mathcal{H}}_\varkappa, \boldsymbol{p}_{\text{SR}}^v), \tag{53}$$

as the maximum of the pdf of WL masses, given the count rate, redshift and sky position, evaluated for our specific posterior draw $\boldsymbol{p}_{\text{SR}}^v$. We report the median mass of the clusters in each stacking bin in Fig. 17. For each WL mass $M_{\text{WL},\varkappa}^v$ we evaluate our extraction model $g_{t,\varkappa}^{\text{mod},v}$, given in Eq. 33 using the source redshift distribution computed on the real data, Eq. 25, and the mean positions of the sources θ_α from Eq. 24. For each cluster and posterior point, we thus construct the predicted reduced shear profile as

$$g_{t,\varkappa}^{\text{pred},v} = (1 - f_{\text{RS},\varkappa}^v - f_{\text{AGN},\varkappa}^v) g_{t,\varkappa}^{\text{mod},v} + f_{\text{AGN},\varkappa}^v g_{t,\varkappa}^{\text{AGN}}, \tag{54}$$

with $f_{\text{RS},\varkappa}^v$ and $f_{\text{AGN},\varkappa}^v$ the random source and AGN fractions at the count rate, redshift and sky position of the cluster $\varkappa$ given the model parameters $\boldsymbol{p}_{\text{SR}}^v$.

The average reduced shear model in the bin is then given by stacking the predicted reduced shear profile with the same weights as the data, and taking a mean over the posterior points,

$$g_t^{\text{pred}} = \frac{1}{N_{\text{MC}}} \sum_v \frac{\sum_\varkappa W_\varkappa g_{t,\varkappa}^{\text{pred},v}}{\sum_\varkappa W_\varkappa}, \tag{55}$$

and variance

$$\left(\delta g_t^{\text{pred}}\right)^2 = \frac{1}{N_{\text{MC}}} \sum_v \frac{\sum_\varkappa W_\varkappa \left(g_{t,\varkappa}^{\text{pred},v} - g_{t,\varkappa}^{\text{pred}}\right)^2}{\sum_\varkappa W_\varkappa}, \tag{56}$$

propagating the uncertainty in the model parameters expressed by their posterior to the reduced shear profile prediction. The mean shear profile and its systematic uncertainty are shown as green lines in Fig. 17. The visual impression of a good fit is corroborated by considering that the chi-squared between stacked data and predicted model is $\chi^2 = 180.0_{-30.4}^{+45.8}$ for 150 data points and 5 parameters that are effectively constrained. The upper and lower errors for the χ^2 result from the difference induced when evaluating it with the model $g_t^{\text{pred}} \pm \delta g_t^{\text{pred}}$, instead of the mean model. Within the errors propagated from the posterior, we attain a very good chi-squared.

6. Discussion

6.1. Comparison to previous work

The only previous work that calibrated the eROSITA count rate – halo mass relation was performed by Chiu et al. (2022), measuring, calibrating and analysing the WL signal around 313 eFEDS selected clusters and groups with HSC data. Their results are shown as black lines in Fig. 15. We agree on all parameters within 2-sigma, with larger than 1-sigma deviations on the amplitude, the redshift trend of the slope and the scatter. When considering the predicted mean count rate as a function of mass and redshift, their results predict higher count rates for low mass, high redshift objects (hatched bands in Fig. 16). The difference is however not statistically significant. The constraining power of this work is only marginally better than the calibration derived by Chiu et al. (2022).

The similar constraining power of eFEDS-HSC analysis to this work, in spite of the significantly smaller number of objects (313 compared to 2201) is explained by the higher source density in HSC, as well as the cleaner background selection. Indeed, the signal-to-noise ratio of an individual cluster is proportional to

$$\frac{S}{N} \propto \beta \sqrt{n_{\text{eff}}} \sqrt{1 - f_{\text{cl}}}, \text{ with } \beta = \left\langle \frac{\min(0, D_{\text{ls}})}{D_s} \right\rangle_{P(z_s)}, \tag{57}$$

where β is the lensing efficiency, which is 0 for sources at the lens redshift, and converges to 1 for infinitely distant sources. n_{eff} is the effective source density, and f_{cl} the cluster member contamination. When considering these elements, each cluster in the HSC analysis has a WL signal-to-noise approximately a factor of 2.3 larger then the clusters considered in this work, assuming they have the same mass, and ignoring cluster member contamination. Considering the overall signal-to-noise gain stemming from the total number of clusters, $\sqrt{2201/313} = 2.6$, we find that it just about compensates the much lower individual clusters signal-to-noise. Note also that in the eFEDS-HSC analysis,
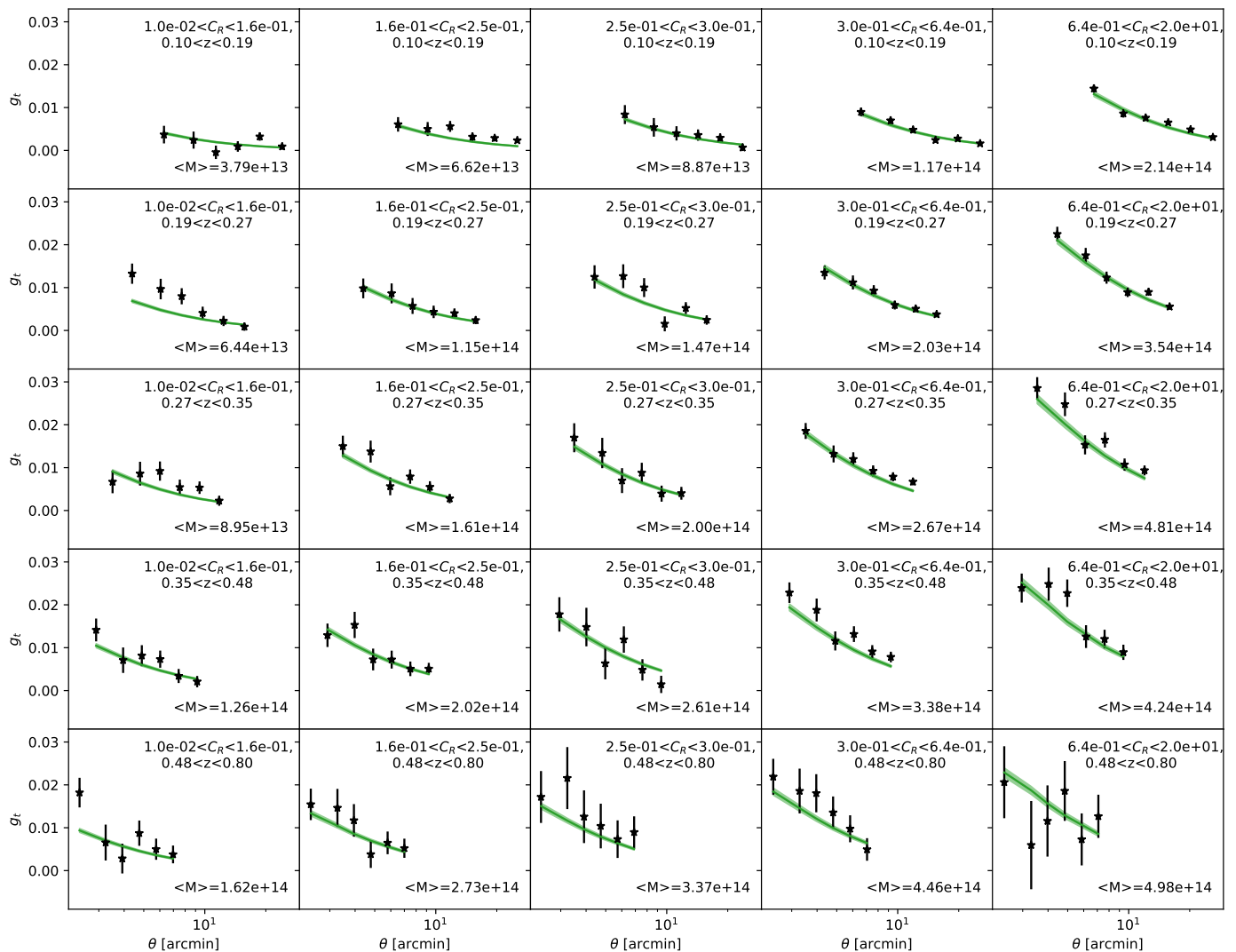
Fig. 17. Stacked measured reduced tangential shear profiles with measurement uncertainty in black, for different count rate (rows) and redshift (columns) bins. As a green line the best-fit shear profile model and with filled green regions the propagation of the posterior uncertainty on the mean model. The chosen model fits the data well at all count rates and redshifts. We report also the median WL mass of each bin, in units of $M_\odot$.

a tighter Ω_M, and an informative σ_X prior where used. Taken together, this explains the similar constraining power of the two works.

6.2. Cluster line of sight anomalies

We find several cluster line of sight anomalies, that we shall discuss here in a cohesive manner. Firstly, it is somewhat anomalous that the cluster member contamination estimates based on the effective number density of sources display a larger contaminant fraction than the estimate based on the decomposition of the source redshift distribution. The latter should be impervious to masking and obstruction effects, as argued by Gruen et al. (2014); Varga et al. (2019). Estimations of the cluster member contamination via the effective number density of sources are biased low because of masking and obstruction by foreground and bright cluster sources, as shown by Kleinebreil et al. (in prep.). It is therefore surprising that we find the number density estimate to be larger than the source redshift distribution estimate, despite it potentially being biased low. A possible source of error here is our very lenient background selection and the possible

performance loss of the source redshift decomposition method if the cluster redshift lies too close to the peak of source redshift distribution of the field. This might be hinted at by the imperfect subtraction signatures discussed in Section 3.3.1 and seen in Fig. 6.

Also interesting are the decrement of the effective shape noise and the increment of the smooth shear response towards the cluster center. The effect is rather constant in lens redshift, as opposed to the cluster member contamination. Indeed, we see this anomaly also in the lowest source redshift bin, $0.1 < z_l < 0.19$, where we have practically no cluster member contamination. This makes it implausible that cluster member contaminants have significantly different responses or intrinsic shapes. It is worthwhile noting in Eq. 10 that a larger shear response leads to a smaller estimated effective shape noise. Also, the shear response used in this work is a smooth function of the source's signal-to-noise and its size w.r.t. the PSF. Noticeably, both these quantities are magnified by the cluster's potential – a secondary effect of WL. Such magnification effects might thus alter the inferred smooth shear response. In general, shape measurement and photo-z calibration in cluster lines of sight might be effected to second order not only by magnification, but also by increased
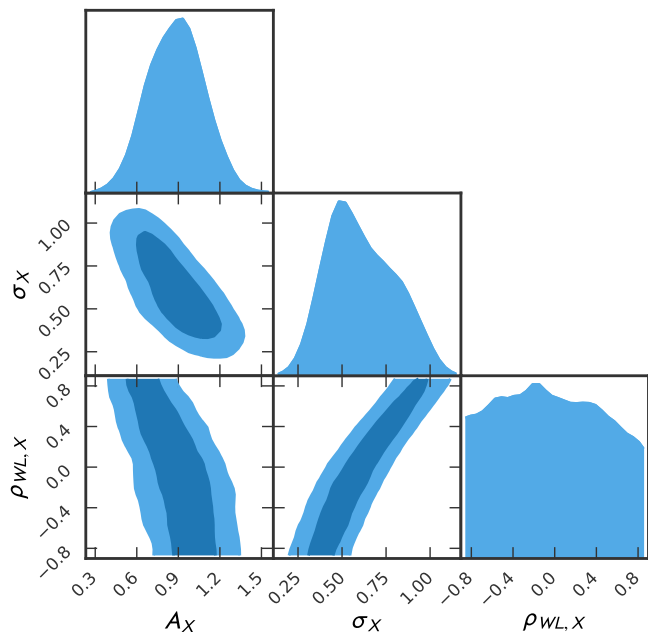
Fig. 18. 1-d and 2-d marginal contours of the amplitude of the count rate–mass scaling relation A_X, the intrinsic scatter around that relation σ_X, and the correlation coefficient between the WL and count rate scatter ρ_{M_{WL}, C_R}, labelled as $\rho_{WL,X}$ for brevity in the plot. While the correlation coefficient remains unconstrained, it leads to significant degeneracies with the amplitude and scatter. Astrophysical processes like cluster age, merger state, environment, orientation, triaxiality, etc., might impact both observables and lead to selection biases. We agnostically marginalize over the entire range of such effects, thereby weakening our amplitude and scatter measurements, but making the mass calibration impervious to such astrophysical effects.

blending (Hernández-Martín et al. 2020) and the intra-cluster light (Zhang et al. 2019). Exploitation of future wide optical and NIR surveys might necessitate dedicated assessments of the performance of shape and photo-z measurement techniques in cluster lines of sight.

6.3. Accounting for cluster selection biases

Much of the astrophysically driven conversation about the viability of galaxy clusters as a cosmological probe, and the limits of our ability to effectively calibrate their masses, focus on so called *selection biases*. When taking a forward modelling approach using Bayesian population models, these concerns can be naturally accounted for. We shall discuss in the following the most (in)famous "selection bias" as an example: the "cool core bias" (Rossetti et al. 2017). The cooling time in the center of massive galaxy clusters can fall below the Hubble time, especially if they have been undisturbed for a long time. Effective cooling boosts the central luminosity, resulting in very peaked X-ray surface brightness profiles. Conversely, the SZe selection, sensitive to the pressure, is not effected by this. As discussed already above, taking a population modelling approach, it immediately results that any sample selected on an observable that scatters around the halo mass will preferentially contain up-scattered objects compared to down-scattered objects, on account of the much higher frequency of the former. If a specific physical property drives the scatter in some observable at fixed mass, then a sample selected in that observable would show more objects with

that physical property, compared to a sample selected on another observable. X-ray selected samples will preferentially contain objects that scattered up in X-rays, for whatever astrophysical reason.

For cosmological inference the complication arises if the scatter in the follow-up observable used for the mass calibration correlates with scatter in the selection observable. A glaring case which is being much investigated, is clusters selected via the number of photometric member galaxies, and the WL signal of these objects. The scatter of both depends on the concentration, orientation and triaxiality of the halo, as well as the neighboring environment, resulting in strong selection biases hampering cosmological analyses on those samples (Dark Energy Survey Collaboration et al. 2020; Wu et al. 2022). Analogously, X-ray-selected clusters have been reported to live preferentially in nodes of the cosmic web, as compared to filaments or sheets, as found by Popesso et al. (2023) when comparing eFEDS-selected halos with spectroscopically selected ones. This means that at fixed halo mass, the cluster properties may depend on secondary halo properties like environment or concentration – an effect analogous to the *assembly bias* of galaxies (Wu et al. 2008; Wechsler & Tinker 2018). Our scale cuts on the tangential reduced shear are designed to ensure that we measure the WL signal only in the 1-halo regime (Grandis et al. 2021a), that is scales dominated by the halo's own density profile, and not by neighboring structures. This already limits significantly the impact of assembly bias. Still, the eRASS1 selection might correlate with some property that impact also the WL mass, such as concentration, triaxiality, or dynamical state. To account for any such possible effect, we introduce the correlation coefficient $\rho_{WL,X}$ among the scatters in count rate and WL mass. Fig. 18 shows the posterior on the correlation coefficient, together with its degeneracies with the amplitude of the scaling relation and intrinsic scatter at fixed halo mass and redshift. The prior range for the correlation coefficient is naturally given by $(-1, 1)$, though we sample $(-0.9, 0.9)$ to avoid numerical instabilities for singular covariance matrix. Marginalising over the correlation coefficient makes our analysis impervious to selection bias effects, but deteriorates our ability to calibrate the observable–mass relation. Given that the numerical value of the amplitude of the scaling relation directly impacts our cosmological results, and crucially their agreement with external experiments, foregoing the use of strong astrophysical priors on the selection effects at the cost of a less stringent mass calibration is undoubtedly the more prudent choice.

7. Conclusions

In this work, we perform the measurement and calibration of the WL signal around eRASS1 clusters and groups in the DES Y3 data. To this end, we define a background sample of DES galaxies by weighting tomographic redshift bins defined and calibrated for the cosmic shear experiment (Myles & Alarcon et al. 2021; Gatti & Sheldon et al. 2021) as a function of the photometric redshift of our clusters. After querying sources around our lenses, out to an angular distance corresponding to $15h^{-1}$Mpc, we validate and calibrate that measurement. We fit the lens richness, redshift and source-lens distance dependence of the effective source density and the source redshift distribution to determine the contamination of our source sample. The different methods employed agree with each other within the systematic uncertainties of the source redshift and shape measurement.

We fit a lens redshift trend to the effective shape noise, allowing us to define a semi-analytical covariance metric for the shape

noise. The profiles of the B-mode-like cross component are consistent with zero, as expected for a well-calibrated WL signal sourced by gravity. Investigation of the selection and shear response of our tangential reduced shear estimators show that the former can be ignored, as it induces a correction much smaller than the systematic uncertainty induced by source redshift and shape measurements. On small cluster-centric distances we find interesting anomalies in the shear response and effective shape noise, that might point at cluster line of sight specific calibration issues, which arguably are too small to impact our subsequent analysis steps.

We derive cosmology-ready data products, used as inputs for the mass calibration of eRASS1 selected clusters, both in this work, and self-consistently also in the subsequent cosmological analysis by Ghirardini et al. (A&A, subm.). These products are:

- the tangential reduced shear profile in angular bins, corrected for the shear response,
- a semi-analytical estimate of the shape noise on that profile, and
- an estimate of the field source redshift distribution from the local background,

for each source with DES Y3 lensing data. This results in WL data for 2201 eRASS1 selected clusters, with a raw signal-to-noise ratio of 92, reduced to 65 after the our radial scale cuts are applied.

Exploitation of these data for mass calibration and number counts cosmology requires the anchoring of the mapping between halo mass and WL signal via $O(1000)$ realisations of synthetic shear profile simulations, following the scheme presented by Grandis et al. (2021a). These simulations include:

- realistic 2d projected surface mass density maps around massive halos, extracted from the cosmological hydro-dynamical TNG300 simulations (Pillepich et al. 2018; Marinacci et al. 2018; Springel et al. 2018; Nelson et al. 2018; Naiman et al. 2018; Nelson et al. 2019),
- a distribution of the off-set between the observational centers used for the tangential reduced shear profile measurements, constructed from X-ray images of the MAGNETICUM PATHFINDER simulations (Dolag et al., in prep.), and the 'eROSITA digital sky twin' (Seppi et al. 2022),
- a painting of the richness using the richness–mass relation calibrated by Chiu et al. (2022),
- dilution of the tangential reduced shear signal based on the degree of cluster member contamination we calibrated,
- a noise contribution from the uncorrelated large-scale structure, and
- realistic DES Y3 source distributions reflecting our background selection based on weighted sums of the DES Y3 tomographic redshift bins (Myles & Alarcon et al. 2021), together with a prescription for non-linear shear bias.

These synthetic shear profiles are analysed with the same shear profile model that we use in the subsequent mass calibration. It features a simplified mis-centered NFW with fixed concentration–mass relation, and accounts for the mean cluster member contamination, as advocated by Grandis et al. (2021a). It thus has one free parameter, the *WL mass*. Extracting the WL mass on our synthetic shear profile simulations allows us to establish the relation between WL and halo mass, summarized in the WL bias and scatter. Monte-Carlo realisations of the synthetic shear profile simulation, obtained by varying all the inputs within our expectation, result in an uncertainty estimate on

the WL bias and scatter, that acts as a calibration prior for the subsequent mass calibration. We find that at low redshift our mass extraction uncertainty is dominated by the effect of hydro-dynamical modelling, leading to a systematics floor of 2%. At higher redshift, DES photo-z uncertainties become incrementally more important.

We calibrate the count rate – halo mass relation by using a Bayesian population model to forward-model the probability density function of each cluster's shear profile given its count rate and redshift. Improving on previous work (Mantz et al. 2016; Dietrich et al. 2019; Bocquet et al. 2019; Chiu et al. 2022), we take explicit account of contaminants via a mixture model. We find that we are able to measure 5 parameters of the count rate–mass relation: the amplitude, the slope and its redshift dependence, the deviation from a self-similar redshift evolution, and the intrinsic scatter around the relation. We assess the quality of this population model based fit on stacked signals in count rate, redshift bins, and find excellent agreement among the data and the model. Our analysis finds comparable, and marginally tighter constraints than the calibration of the eROSITA count rate–mass relation performed by Chiu et al. (2022) with HSC data on eFEDS-selected clusters.

This analysis will serve, along with similar analyses on HSC and KiDS, as the main mass calibration for the number counts of eRASS1 selected clusters (Ghirardini et al. A&A, subm.), which constrain different cosmological parameters like the present-day matter density, the amplitude of matter fluctuations, the equation of state of Dark Energy, and the sum of the neutrino masses. Methodologically, we demonstrate how photometric data from wide surveys, calibrated specifically for cosmic shear experiments, can be effectively used the perform a WL mass calibration of cluster number counts. These two experiments are highly complementary. Cosmic shear probes the linear and mildly non-linear scale of the large-scale structure, while WL calibrated number counts probe the most non-linear regions, massive galaxy clusters. These probes have very different sensitivities to astrophysical and observational systematics, only sharing a joint dependence on the source photo-z and shape calibration. In light of upcoming deep and wide surveys, tailored mainly for cosmic shear, like *Euclid*[¶], LSST[‖], or Roman[**], this work provides a demonstration on how to leverage these data to inform cluster number counts, enabling independent and competitive cosmological constraints.

[¶] http://www.euclid-ec.org/
[‖] https://www.lsst.org/
[**] https://roman.gsfc.nasa.gov/

References

Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409

Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, MNRAS, 426, 2046

Bahé, Y. M., McCarthy, I. G., & King, L. J. 2012, MNRAS, 421, 1073

Becker, M. R. & Kravtsov, A. V. 2011, ApJ, 740, 25

Bellagamba, F., Sereno, M., Roncarelli, M., et al. 2019, MNRAS, 484, 1598

Benítez, N. 2000, ApJ, 536, 571

Bleem, L. E., Bocquet, S., Stalder, B., et al. 2020, ApJS, 247, 25

Bocquet, S., Dietrich, J. P., Schrabback, T., et al. 2019, ApJ, 878, 55

Bocquet, S., Grandis, S., Bleem, L. E., et al. 2023, arXiv e-prints, arXiv:2310.12213

Buchner, J. 2021, The Journal of Open Source Software, 6, 3001

Bulbul, E., Liu, A., Kluge, M., Zhang, X., & X. A&A, subm.

Castro, T., Borgani, S., Dolag, K., et al. 2021, MNRAS, 500, 2316

Child, H. L., Habib, S., Heitmann, K., et al. 2018, ApJ, 859, 55

Chiu, I. N., Ghirardini, V., Liu, A., et al. 2022, A&A, 661, A11

Chiu, I. N., Klein, M., Mohr, J., & Bocquet, S. 2023, MNRAS, 522, 1601

Clerc, N., Comparat, J., Seppi, R., & Artis, E. A&A, subm.

Comparat, J., Eckert, D., Finoguenov, A., et al. 2020, The Open Journal of Astrophysics, 3, 13

Comparat, J., Luo, W., Merloni, A., et al. 2023, A&A, 673, A122

Comparat, J., Merloni, A., Salvato, M., et al. 2019, MNRAS, 487, 2005

Cordero, J. P., Harrison, I., Rollins, R. P., et al. 2022, MNRAS, 511, 2170

Dark Energy Survey Collaboration, Abbott, T. M. C., Aguena, M., et al. 2020, Phys. Rev. D, 102, 023509

Dark Energy Survey Collaboration, Abbott, T. M. C., Aguena, M., et al. 2022, Phys. Rev. D, 105, 023520

Dark Energy Survey Collaboration, Abbott, T. M. C., Allam, S., et al. 2019, ApJ, 872, L30

De Vicente, J., Sánchez, E., & Sevilla-Noarbe, I. 2016, MNRAS, 459, 3078

Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168

Dietrich, J. P., Bocquet, S., Schrabback, T., et al. 2019, MNRAS, 483, 2871

Dolag, K., X., ., & et al. in prep.

Dyson, F. W., Eddington, A. S., & Davidson, C. 1920, Philosophical Transactions of the Royal Society of London Series A, 220, 291

Ettori, S., Gastaldello, F., Leccardi, A., et al. 2011, Mass profiles and c - M_{DM} relation in X-ray luminous galaxy clusters (Corrigendum), Astronomy & Astrophysics, Volume 526, id.C1, 1 pp.

Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, AJ, 150, 150

Friedrich, O., Andrade-Oliveira, F., Camacho, H., et al. 2021, MNRAS, 508, 3125

Gatti, M., Sheldon, E., Amon, A., et al. 2021, MNRAS, 504, 4312

Ghirardini, V., Esra, B., Artis, A., Clerc, N., & Garrel, C. A&A, subm.

Grandis, S., Bocquet, S., Mohr, J. J., Klein, M., & Dolag, K. 2021a, MNRAS, 507, 5671

Grandis, S., Klein, M., Mohr, J. J., et al. 2020, MNRAS, 498, 771

Grandis, S., Mohr, J. J., Costanzi, M., et al. 2021b, MNRAS, 504, 1253

Grandis, S., Mohr, J. J., Dietrich, J. P., et al. 2019, MNRAS, 488, 2041

Gruen, D., Seitz, S., Brimioulle, F., et al. 2014, MNRAS, 442, 1507

Haiman, Z., Mohr, J. J., & Holder, G. P. 2001, ApJ, 553, 545

Hennig, C., Mohr, J. J., Zenteno, A., et al. 2017, MNRAS, 467, 4015

Hernández-Martín, B., Schrabback, T., Hoekstra, H., et al. 2020, A&A, 640, A117

Hoekstra, H., Mahdavi, A., Babul, A., & Bildfell, C. 2012, MNRAS, 427, 1298

Huff, E. & Mandelbaum, R. 2017, arXiv e-prints, arXiv:1702.02600

Huterer, D. 2022, arXiv e-prints, arXiv:2212.05003

Jarvis, M., Bernstein, G. M., Amon, A., et al. 2021, MNRAS, 501, 1282

Kimmig, L. C., Remus, R.-S., Dolag, K., & Biffi, V. 2023, ApJ, 949, 92

Kleinebreil, F., Grandis, S., Schrabback, T., & Ghirardini, V. in prep.

Kluge, M., Comparat, J., et Liu, A., X, & Y. A&A, subm.

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

MacCrann, N., Becker, M. R., McCullough, J., et al. 2022, MNRAS, 509, 3371

Majumdar, S. & Mohr, J. J. 2004, ApJ, 613, 41

Mantz, A. B., Allen, S. W., Morris, R. G., et al. 2016, MNRAS, 463, 3582

Mantz, A. B., von der Linden, A., Allen, S. W., et al. 2015, MNRAS, 446, 2205

Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, MNRAS, 480, 5113

McClintock, T., Varga, T. N., Gruen, D., et al. 2019, MNRAS, 482, 1352

Merloni, A., Lamer, G., Liu, T., et al. 2024, A&A, 682, A34

Mohr, J. J. & Evrard, A. E. 1997, ApJ, 491, 38

Myles, J., Alarcon, A., Amon, A., et al. 2021, MNRAS, 505, 4249

Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563

Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624

Nelson, D., Springel, V., Pillepich, A., et al. 2019, Computational Astrophysics and Cosmology, 6, 2

Oguri, M. & Hamana, T. 2011, MNRAS, 414, 1851

Paulus, M. 2021, PhD thesis, Ludwig-Maximilians University of Munich, Germany

Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, MNRAS, 475, 648

Popesso, P., Biviano, A., Bulbul, E., et al. 2023, arXiv e-prints, arXiv:2302.08405

Power, C., Navarro, J. F., Jenkins, A., et al. 2003, MNRAS, 338, 14

Prat, J., Blazek, J., Sánchez, C., et al. 2022, Phys. Rev. D, 105, 083528

Predehl, P., Andritschke, R., Arefiev, V., et al. 2021, A&A, 647, A1

Ragagnin, A., Saro, A., Singh, P., & Dolag, K. 2021, MNRAS, 500, 5056

Rossetti, M., Gastaldello, F., Eckert, D., et al. 2017, MNRAS, 468, 1917

Rozo, E., Rykoff, E. S., Becker, M., Reddick, R. M., & Wechsler, R. H. 2015, MNRAS, 453, 38

Rykoff, E. S., Rozo, E., Busha, M. T., et al. 2014, ApJ, 785, 104

Rykoff, E. S., Rozo, E., Hollowood, D., et al. 2016, ApJS, 224, 1

Saro, A., Bocquet, S., Rozo, E., et al. 2015, MNRAS, 454, 2305

Schneider, P. 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, ed. G. Meylan, P. Jetzer, P. North, P. Schneider, C. S. Kochanek, & J. Wambsganss, 269–451

Schrabback, T., Bocquet, S., Sommer, M., et al. 2021, MNRAS, 505, 3923

Seppi, R., Comparat, J., Bulbul, E., et al. 2022, A&A, 665, A78

Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, ApJS, 254, 24

Sheldon, E. S. & Huff, E. M. 2017, ApJ, 841, 24

Shin, T., Jain, B., Adhikari, S., et al. 2021, MNRAS, 507, 5758

Sommer, M. W., Schrabback, T., Ragagnin, A., & Rockenfeller, R. 2023, arXiv e-prints, arXiv:2306.13187

Song, J., Zenteno, A., Stalder, B., et al. 2012, ApJ, 761, 22

Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676

Story, K., Aird, K. A., Andersson, K., et al. 2011, ApJ, 735, L36

Sunayama, T. 2023, MNRAS, 521, 5064

Sunayama, T., Park, Y., Takada, M., et al. 2020, MNRAS, 496, 4468

Sunyaev, R., Arefiev, V., Babyshkin, V., et al. 2021, A&A, 656, A132

Umetsu, K. 2020, A&A Rev., 28, 7

Varga, T. N., DeRose, J., Gruen, D., et al. 2019, MNRAS, 489, 2511

Wechsler, R. H. & Tinker, J. L. 2018, ARA&A, 56, 435

Wu, H.-Y., Costanzi, M., To, C.-H., et al. 2022, MNRAS, 515, 4471

Wu, H.-Y., Rozo, E., & Wechsler, R. H. 2008, ApJ, 688, 729

Zhang, Y., Yanny, B., Palmese, A., et al. 2019, ApJ, 874, 165

Zohren, H., Schrabback, T., Bocquet, S., et al. 2022, A&A, 668, A18

Appendix A: Effective shape noise and number density in presence of shear responses

The aim of this appendix is to present a derivation for the expression of the effective shape variance, Eq. 10 and the effective number density, Eq. 9, in presence of a shear response $\mathcal{R}$.

Friedrich et al. (2021) demonstrate that the variance of a reduced shear profile $\hat{g}_t$ is given by

$$\text{Var}\left[\hat{g}_t\right] = \frac{\sigma_{\text{eff}}^2}{N_{\text{pair}}} \text{ with } \sigma_{\text{eff}}^2 = \frac{1}{2N_s} \sum |e|^2 w_n^2, \tag{A.1}$$

where N_{pair} is the number of source lens pairs, N_s the number of sources, $|e|^2 = e_t^2 + e_x^2$ the squared modulus of the ellipticity when interpreted as a complex number, and w_n normalized weights.

To generalize this expression to un-normalized weight w, just consider that such un-normalized weights can always be rescaled to normalized weights with the ansatz $w_n = Cw$. The normalisation condition then readily leads to an expression for C, reading

$$\frac{1}{N_s} \sum w_n = 1 = \frac{C}{N_s} \sum w \Leftrightarrow C^{-1} = \frac{1}{N_s} \sum w. \tag{A.2}$$

Moving into the regime of individual cluster lensing, where $N_{\text{pair}} = N_s$, combining all the above equations, we find

$$\text{Var}\left[\hat{g}_t\right] = \frac{1}{2} \frac{1}{\left(\sum w\right)^2} \sum |e|^2 w^2, \tag{A.3}$$

for generic, un-normalized weights.

When considering shear data with shear responses $\mathcal{R}$, Friedrich et al. (2021), Appendix D, suggest to transform the above equations by re-scaling both the ellipticities e, as well as the weight w, as follows

$$\tilde{e} = \frac{e}{\mathcal{R}} \text{ and } \tilde{w} = \mathcal{R}w, \tag{A.4}$$

thus correcting the ellipticity for the shear response, and adjusting the weights for the necessary re-scaling. Note that the summand in the numerator of Eq. A.3 is invariant under this rescaling, $|\tilde{e}|^2 \tilde{w}^2 = |e|^2 w^2$, and we thus only have to modify the summand in the denominator. Given that the above expression is valid for un-normalized weights, no further corrections are required.

Dropping the tildes, we can readily write the variance of a tangential reduced shear profile in presence of shear response as

$$\text{Var}\left[\hat{g}_t\right] = \frac{1}{2} \frac{1}{\left(\sum w\mathcal{R}\right)^2} \sum |e|^2 w^2. \tag{A.5}$$

Similarly, this justifies also why the source redshift distribution needs to be shear response weighted.

Interpreting this expression as the ratio between an effective shape dispersion $\sigma_{\text{eff},\alpha}$ for $\alpha \in (t, x)$, and an effective number of source N_{eff} we can recover the above expression by using

$$N_{\text{eff}} = \frac{\left(\sum w\mathcal{R}\right)^2}{\sum w^2\mathcal{R}^2} \text{ and } \sigma_{\text{eff},\alpha}^2 = \frac{\sum e_\alpha^2 w^2}{\sum w^2\mathcal{R}^2} \text{ for } \alpha \in (t, x), \tag{A.6}$$

which are the expressions that we also use for our shape noise modelling.

Appendix B: Parametrisation of cluster member contamination

We shall justify in the following the parameterisation used for the cluster member contamination in Section 3.3.1. As shown by Hennig et al. (2017), cluster member galaxies follow an NFW profile, with the total number of galaxies enclosed in the virial region correlating tightly with mass, while the concentration varies depending on the galaxy type. Red-sequence galaxies have a larger concentration, more similar to the concentration of the dark matter, than non-red-sequence galaxies. We therefore assume that the cluster member galaxies that contaminated our background sample also follow an NFW profile.

The fraction of red-sequence galaxies in the cluster member contamination is unknown, though it stands to argue that their more precise photometric redshifts make it easier to exclude them with a background selection. We therefore let the concentration c of the cluster member contaminants profile as a free parameter.

Similarly, we choose an agnostic approach to the redshift trend of the cluster member contamination. We simply assume that the ratio between the number density of field galaxies $n_{\text{field}}(z)$ and cluster member contaminants is a smooth function of redshift. We therefore employ a non-parameteric fit of the amplitudes A_{z_i} for pivot redshifts z_i.

To account for possible mass trends we use the richness λ as a mass proxy, given that we have to fit for the cluster member contamination prior to performing a mass calibration, and therefore do not have WL calibrated mass estimates at our disposal. For ICM-selected cluster samples, richness is known to scale approximately linearly with halo mass, $\lambda \sim M$ (Saro et al. 2015; Bleem et al. 2020; Grandis et al. 2020, 2021b). We therefore use the richness to define an approximate virial radius for the NFW profile of the cluster member contaminants as $R_{\text{vir}} \propto (\lambda/20)^{1/3}$. We also allow the amplitude of the cluster member contamination to vary with richness.

Taken together, the above consideration lead to the ansatz

$$n_{\text{fcl}}(\lambda, z, R) = A(\lambda, z, R)\, n_{\text{field}}(z) \text{ with}$$
$$A(\lambda, z, R) = A_j \left(\frac{\lambda}{25}\right)^{B_\lambda} \Sigma_{\text{NFW}}^{\text{norm}}\left(R\middle| r_S = c^{-1}\left(\frac{\lambda}{20}\right)^{1/3}\right). \tag{B.1}$$

The total number density results from the sum of the cluster member contaminants number density and the field number density, as

$$n_{\text{tot}} = n_{\text{fcl}}(\lambda, z, R) + n_{\text{field}}(z) = [A(\lambda, z, R) + 1]\, n_{\text{field}}(z), \tag{B.2}$$

while the fraction of cluster member contaminants is

$$f_{\text{cl}}(\lambda, z, R) = \frac{n_{\text{fcl}}}{n_{\text{tot}}} = \frac{A(\lambda, z, R)}{A(\lambda, z, R) + 1}, \tag{B.3}$$

thus justifying the parametrisation chosen in Eq. 12. This parametrisation has the benefit, that for $A(\lambda, z, R) > 0$, $0 < f_{\text{cl}}(\lambda, z, R) < 1$, as already noted by Grandis et al. (2021b) in the context of the richness-dependent contamination of photometrically selected cluster samples.

Finally, one can show that $1 + A = (1 - f_{\text{cl}})^{-1}$, thus also motivating Eq. 14.

Affiliations

1. Universität Innsbruck, Institut für Astro- und Teilchenphysik, Technikerstr. 25/8, 6020 Innsbruck, Austria
2. Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse 1, 85748 Garching, Germany
3. Faculty of Physics, Ludwig-Maximilians-Universität München, Scheinerstr. 1, 81679 Munich, Germany
4. Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
5. Linköpings Universitet, Institutionen för Systemteknik, Linköpings Universitet, Linköping 581 83, Sweden
6. Kavli Institute for Cosmology, University of Cambridge, Mading- ley Road, Cambridge CB3 0HA, United Kingdom
7. Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, United Kingdom
8. Physics Department, University of Wisconsin-Madison, Madison, WI 53706, USA
9. Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA
10. Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
11. Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15312, USA
12. Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain
13. Laboratório Interinstitucional de e-Astronomia - LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
14. Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain
15. Center for Astrophysical Surveys, National Center for Supercom- puting Applications, 1205 West Clark St., Urbana, IL 61801, USA
16. Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
17. Physics Department, William Jewell College, Liberty, MO, 64068, USA
18. Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA
19. Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
20. Department of Physics, Duke University Durham, NC 27708, USA
21. Department of Physics, National Cheng Kung University, 70101 Tainan, Taiwan
22. NASA Goddard Space Flight Center, 8800 Greenbelt Rd, Green- belt, MD 20771, USA
23. IRAP, Université de Toulouse, CNRS, UPS, CNES, Toulouse, France
24. Jodrell Bank Center for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom
25. School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia
26. Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berke- ley, CA 94720, USA
27. Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA
28. NSF AI Planning Institute for Physics of the Future, Carnegie Mellon University, Pittsburgh, PA 15213, USA
29. Université Grenoble Alpes, CNRS, LPSC-IN2P3, 38000 Grenoble, France
30. Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
31. Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA
32. SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
33. Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain
34. Department of Physics and Astronomy, Pevensey Building, University of Sussex, Brighton, BN1 9QH, UK
35. School of Physics and Astronomy, Cardiff University, CF24 3AA, United Kingdom
36. Department of Astronomy, University of Geneva, ch. d'Écogia 16, CH-1290 Versoix, Switzerland
37. Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA
38. Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom
39. Department of Physics, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305, USA
40. Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859, Campinas, SP, Brazil
41. Physics Program, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan
42. Hiroshima Astrophysical Science Center, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan
43. Core Research for Energetic Universe, Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan
44. Department of Physics, University of Genova and INFN, Via Dodecaneso 33, 16146, Genova, Italy
45. Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA
46. Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
47. Brookhaven National Laboratory, Bldg 510, Upton, NY 11973, USA
48. Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA
49. Département de Physique Théorique and Center for Astroparticle Physics (CAP), University of Geneva, 24 quai Ernest Ansermet, 1211 Genève, Switzerland
50. Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, CNRS, UPS, CNES, 14 Av. Edouard Belin, 31400 Toulouse, France
51. Department of Physics, Boise State University, Boise, ID 83725, USA
52. McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA
53. Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
54. Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, I-34131 Trieste, Italy
55. INAF-Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy
56. Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy
57. Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA
58. Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831
59. Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany

60. Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India
61. Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
62. Institute of Theoretical Astrophysics, University of Oslo. P.O. Box 1029 Blindern, NO-0315 Oslo, Norway
63. Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain
64. Department of Physics, The Ohio State University, Columbus, OH 43210, USA
65. Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
66. George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA
67. Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ - 20921-400, Brazil
68. School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK