Calibration of bias and scatter involved in cluster mass measurements using optical weak gravitational lensing

Sebastian Grandis,[1,2]★ Sebastian Bocquet,[1,2] Joseph J. Mohr,[1,2,3] Matthias Klein[1,3] and Klaus Dolag[1,2,4]

[1]*Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany*
[2]*Excellence Cluster ORIGINS, Boltzmannstr. 2, D-85748 Garching, Germany*
[3]*Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, D-85748 Garching, Germany*
[4]*Max Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85748 Garching, Germany*

ABSTRACT

Cosmological inference from cluster number counts is systematically limited by the accuracy of the mass calibration, i.e. the empirical determination of the mapping between cluster selection observables and halo mass. In this work we demonstrate a method to quantitatively determine the bias and uncertainties in weak-lensing (WL) mass calibration. To this end, we extract a library of projected matter density profiles from hydrodynamical simulations. Accounting for shear bias and noise, photometric redshift uncertainties, mis-centreing, cluster member contamination, cluster morphological diversity, and line-of-sight projections, we produce a library of shear profiles. Fitting a one-parameter model to these profiles, we extract the so-called *WL mass* M_{WL}. Relating the WL mass to the halo mass from gravity-only simulations with the same initial conditions as the hydrodynamical simulations allows us to estimate the impact of hydrodynamical effects on cluster number counts experiments. Creating new shear libraries for ∼1000 different realizations of the systematics provides a distribution of the parameters of the WL to halo mass relation, reflecting their systematic uncertainty. This result can be used as a prior for cosmological inference. We also discuss the impact of the inner fitting radius on the accuracy, and determine the outer fitting radius necessary to exclude the signal from neighbouring structures. Our method is currently being applied to different Stage III lensing surveys, and can easily be extended to Stage IV lensing surveys.

Key words: gravitational lensing: weak – galaxies: clusters: general – cosmology: large-scale structure of Universe.

1 INTRODUCTION

Measurements of the abundance of massive haloes and the galaxy clusters they host are a powerful cosmological probe (e.g. Haiman, Mohr & Holder 2001; Allen, Evrard & Mantz 2011; Kravtsov & Borgani 2012; Weinberg et al. 2013). However, the derived constraints are systematically limited by the uncertainty in the relationship between the underlying halo mass and the (multi-)observable halo properties (for a review of cluster mass measurement techniques, see Pratt et al. 2019). Of these observables, weak gravitational lensing is particularly promising because it traces the total enclosed mass and is therefore insensitive to the complex halo dynamics and to the state of the intra-cluster medium. Indeed, weak-lensing (WL) mass calibration has been established as the method of choice for a robust determination of the cluster mass scale (Hoekstra et al. 2012; Applegate et al. 2014; von der Linden et al. 2014; Schrabback et al. 2018; Dietrich et al. 2019; Murata et al. 2019; McClintock et al. 2019a). Various cluster samples have been used for cosmological studies with a joint WL mass calibration (Mantz et al. 2015; Bocquet et al. 2019; Abbott et al. 2020).

The WL shear measurement of a galaxy cluster consists of the tangential distortion of the shape of background source galaxies and an estimate of the source redshift distribution (for a didactic review of cluster WL, see Umetsu 2020). Several sources of systematic and statistical uncertainty impact the mapping between the tangential ellipticity profile and halo mass. The most prominent sources of statistical noise are

(i) the morphological heterogeneity of the cluster mass distribution (such as the different concentration, orientation, halo shape, and substructure),

(ii) the varying amount of line-of-sight projection by nearby correlated structures,

(iii) the contribution of uncorrelated large-scale structure along the line of sight,

(iv) the noise in the shape measurements of source galaxies due to the dispersion of the intrinsic ellipticities, and

(v) the lack of knowledge of the actual cluster centre and its displacement w.r.t. the observationally chosen centre.

The major sources of systematic uncertainties are

(i) the accuracy of the hydrodynamical simulation predictions of the matter distribution of haloes,

(ii) the accuracy of the shape measurements (especially in crowded environments and in the limit of large amounts of shear),

(iii) the accuracy of the photometric redshift estimates used for the source galaxy redshift distribution,

★ E-mail: s.grandis@lmu.de

(iv) the ability to constrain the distribution of offsets between the observationally chosen centres and the true cluster centre (i.e. the mis-centreing distribution), and

(v) the amount of cluster member contamination of the true lensing signal.

Typically, the modelling of the WL shear profile as a function of halo mass is anchored around the fact that the average halo mass profile is well described by a Navarro–Frenk–White density profile (NFW; Navarro, Frenk & White 1996). Conveniently, the gravitational shear caused by an NFW halo can be calculated analytically (Bartelmann 1996). However, as discussed, real haloes do not exhibit perfect spherically symmetric NFW mass profiles and they are not perfectly isolated objects. Therefore, the inferred mass from this approach – we refer to this mass as the WL mass M_{WL} – is a noisy and biased estimator of the actual halo mass M_{halo}. Numerical simulations are then used to calibrate this noise and bias as a function of halo mass and redshift (e.g. Becker & Kravtsov 2011; Oguri & Hamana 2011; Bahé, McCarthy & King 2012; Lee et al. 2018). For completeness, we also note that a parallel approach has recently emerged where the halo mass profile, parametrized as the halo–matter correlation function, is predicted using emulation techniques (Nishimichi et al. 2019).

For practical analyses of WL data, the generic calibrations of the M_{WL}–M_{halo} relation (hereafter WL to halo mass relation) discussed above are of limited use. Most importantly, these calibrations do not account for the fact that only observationally determined cluster centres are available, and these scatter around the true halo centre (e.g. Lin, Mohr & Stanford 2004; Saro et al. 2015; Zhang et al. 2017). However, the above mentioned simulation-based calibrations assume perfect centreing, and the established relations cannot be adapted to account for mis-centred shear profiles, because calculating the shear signal from a mis-centred matter density profile is not equivalent to computing the shear profile of a centred matter density profile and then estimating the effect of mis-centreing by azimuthal averaging. Instead, the effect of mis-centreing must be applied at the level of the simulations. Then, all other sources of uncertainty listed above could in principle be explicitly marginalized over in the cosmological analysis. We attempt this approach and update the cosmological pipeline used for the latest analysis of the *South Pole Telescope*-selected (SPT) cluster sample with targeted WL measurements (Bocquet et al. 2019) to explicitly marginalize over all these effects in a full forward modelling approach. We find that this procedure is computationally expensive, even for cluster samples with a few hundred WL measurements.

We therefore propose to fit a simple model for the shear profile (a projected NFW profile with an effective contribution from mis-centreing) to the data while explicitly accounting for shape noise. All other sources of statistical and systematic uncertainties are folded into the WL to halo mass relation. This treatment allows us to capture all relevant effects while reducing the computational burden on the cosmological analysis pipeline (e.g. for the analysis of SPT clusters with WL mass calibration using data from the Dark Energy Survey; Bocquet et al., in preparation). We note that a similar approach has been performed in Mantz et al. (2015), Bocquet et al. (2019), Dietrich et al. (2019), and Schrabback et al. (2021).
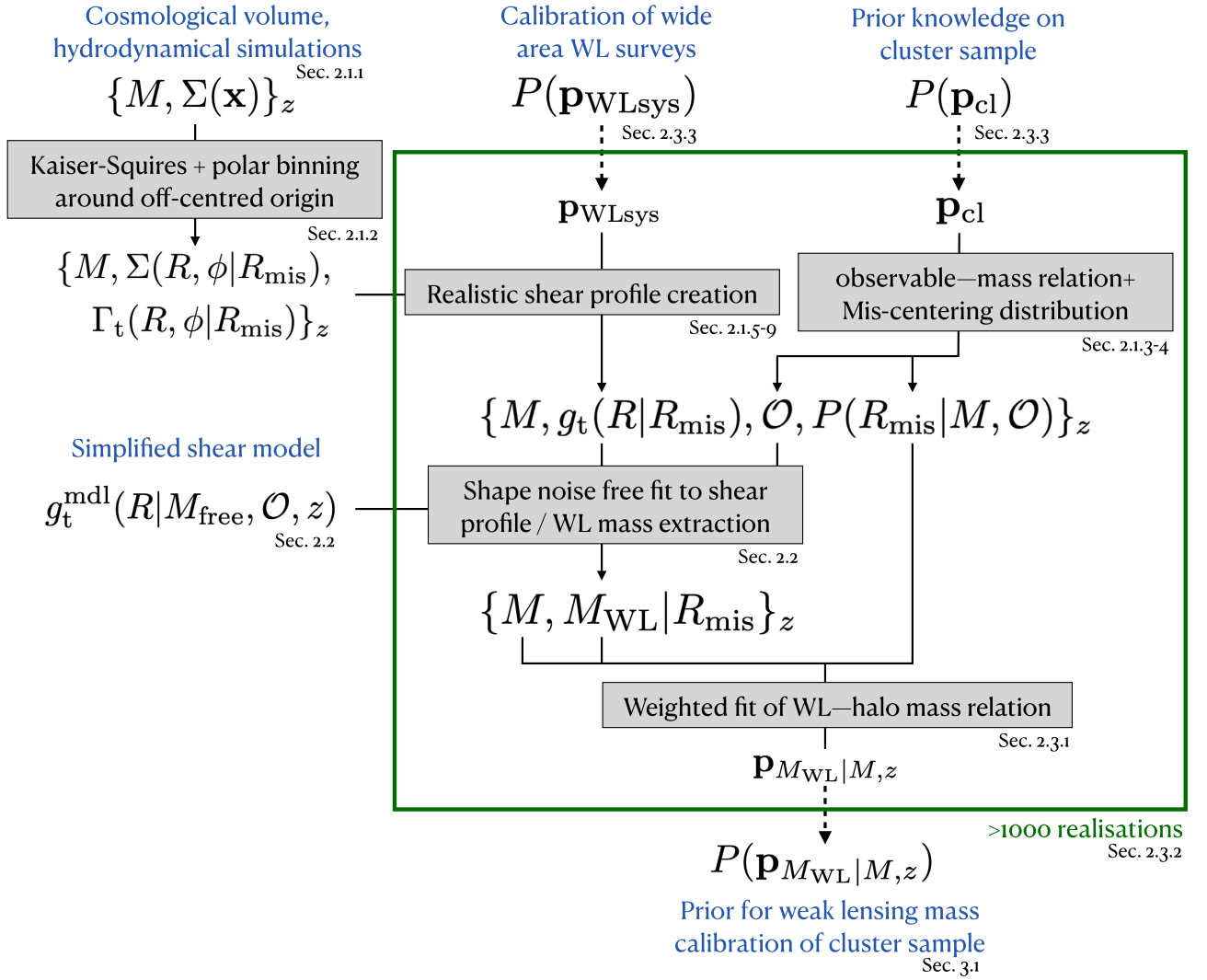
A systematic uncertainty in cluster lensing that is receiving increasing attention is caused by the impact of baryonic feedback effects on the halo matter profile (e.g. Debackere, Schaye & Hoekstra 2021). In the presence of these effects, only hydrodynamical simulations with carefully tuned sub-resolution physics models can be expected to produce realistic halo profiles. However, to perform

a cluster cosmology study, the halo profile needs to be related to a halo mass that is linked to cosmology via the halo mass function. In principle, this could be addressed by establishing both the calibration of the WL to halo mass relation (as in, e.g. Lee et al. 2018) and of the halo mass function on hydrodynamical simulations (as in, e.g. Bocquet et al. 2016). However, one reason not do so is that the advent of emulators based on suites of gravity-only simulations has dramatically increased the accuracy with which the mass function can be predicted for a given cosmological model (McClintock et al. 2019b; Nishimichi et al. 2019; Bocquet et al. 2020). To be able to benefit from these emulators, we propose a novel approach: we keep the halo mass function as defined by gravity-only halo masses. Then, we link these gravity-only masses to realistic halo profiles as obtained from hydrodynamical simulations. This is achieved by using pairs of hydrodynamical and gravity-only simulations with identical initial conditions. As a result, in our approach, the impact of hydrodynamical effects is folded into the WL to halo mass relation along with the other systematic and statistical effects already discussed. Repeating this exercise with different hydrodynamical simulations – ideally weighted by their agreement with real observations – allows us to quantify the error budget related to uncertainties in the hydrodynamical modeling.

We organize this paper as follows. In Section 2.1 we discuss how we produce realistic tangential shear profiles for halo catalogues by combining data extracted from simulations and calibration products derived from observations. We seek to emulate a typical Stage III WL survey, taking mainly inspiration from the Dark Energy Survey year 1 data (DES Y1; Drlica-Wagner et al. 2018) and the cluster catalogue selected by the SPT Sunyaev-Zel'dovich effect survey (SPT-SZ; Benson et al. 2013; Bleem et al. 2015; Bocquet et al. 2019). We then define our model for the shear profile in Section 2.2. We then fit this model to the realistic shear profiles in the library to extract WL masses. In Section 2.3 we then describe how we characterize the relation between WL and halo mass, introducing the WL bias and scatter. By varying the input values of the shear library creation, we propagate the uncertainty on the WL systematics on to the WL bias and scatter. We demonstrate how this procedure provides a posterior on the WL bias and scatter in Section 3.1, which provides a crucial prior for WL calibrated number count experiments. We then discuss the impact of the inner (Section 3.2) and outer (Section 3.3) fitting radius, as well as other systematic effects, which are harder to quantify (Section 3.4), and compare different approaches to modelling the systematic uncertainty stemming from the modelling of hydrodynamical effects (Section 3.5). Finally, we summarize our findings in the conclusions (Section 4).

2 METHOD

In this section, we first discuss how we create a realistic library of shear profiles for a halo catalogue by combining information from hydrodynamical simulations and calibrations of observed data (cf. Section 2.1). We then discuss the shear model we use to measure associated WL masses for each of the realistic shear profiles (cf. Section 2.2). Subsequently, we establish the statistical relation between WL and halo mass (cf. Section 2.3). Finally, we present a method to propagate the uncertainties on the calibration products to the parameters of the WL to halo mass relation, producing a posterior distribution of these parameters that can be used as a prior in cosmological analyses (cf. Section 2.3.2). Fig. 1 is a flow diagram of our method, supplemented with the reference to the sub-section corresponding to each module.

Cosmological volume,
hydrodynamical simulations
Sec. 2.1.1

$$\{M, \Sigma(\mathbf{x})\}_z$$

Calibration of wide
area WL surveys

$$P(\mathbf{p}_{\mathrm{WLsys}})$$
Sec. 2.3.3

Prior knowledge on
cluster sample

$$P(\mathbf{p}_{\mathrm{cl}})$$
Sec. 2.3.3

Kaiser-Squires + polar binning
around off-centred origin
Sec. 2.1.2

$$\{M, \Sigma(R, \phi|R_{\mathrm{mis}}),$$
$$\Gamma_{\mathrm{t}}(R, \phi|R_{\mathrm{mis}})\}_z$$

$$\mathbf{p}_{\mathrm{WLsys}}$$

$$\mathbf{p}_{\mathrm{cl}}$$

Realistic shear profile creation
Sec. 2.1.5-9

observable—mass relation+
Mis-centering distribution
Sec. 2.1.3-4

Simplified shear model

$$g_{\mathrm{t}}^{\mathrm{mdl}}(R|M_{\mathrm{free}}, \mathcal{O}, z)$$
Sec. 2.2

$$\{M, g_{\mathrm{t}}(R|R_{\mathrm{mis}}), \mathcal{O}, P(R_{\mathrm{mis}}|M, \mathcal{O})\}_z$$

Shape noise free fit to shear
profile / WL mass extraction
Sec. 2.2

$$\{M, M_{\mathrm{WL}}|R_{\mathrm{mis}}\}_z$$

Weighted fit of WL—halo mass relation
Sec. 2.3.1

$$\mathbf{p}_{M_{\mathrm{WL}}|M,z}$$

>1000 realisations
Sec. 2.3.2

$$P(\mathbf{p}_{M_{\mathrm{WL}}|M,z})$$

Prior for weak lensing mass
calibration of cluster sample
Sec. 3.1

Figure 1. Graphic representation of the method used to quantify the accuracy of the WL mass extraction by combining the surface matter densities $\Sigma(\mathbf{x})$ around haloes of mass M from cosmological volume, hydrodynamical simulations with calibrations of the WL data in wide-area surveys and priors on the cluster sample. By assuming a simplified shear model, these inputs can be compressed into a single posterior on the parameters of the mapping between WL and halo mass. For ease of navigation, we add the sections in which each component of the diagram is discussed.

2.1 Library of realistic shear profiles

2.1.1 Simulations

Our work is mainly based on Box1 of the Magneticum Pathfinder suite of cosmological hydrodynamical simulations (Dolag et al., in preparation).[1] Additional technical aspects of the simulations are available in the literature (Hirschmann et al. 2014; Teklu et al. 2015; Beck et al. 2016; Dolag, Mevius & Remus 2017). The box size (896 h^{-1} Mpc on a side), the resolution (2 × 1526³ particles), and the mass sampling (1.3×10¹⁰ h^{-1} M$_\odot$ for dark matter particles, 2.6×10⁹ h^{-1} M$_\odot$ for gas particles) allow for a good halo resolution down to about 10¹⁴ h^{-1} M$_\odot$ and for sufficiently large numbers of haloes. The cosmological parameters match the WMAP7 constraints for a spatially flat ΛCDM model (Komatsu et al. 2011): $\Omega_{\mathrm{m}} = 0.272$, $\Omega_{\mathrm{b}} = 0.0457$, $H_0 = 70.4$, $n_{\mathrm{s}} = 0.963$, $\sigma_8 = 0.809$. In addition to the full-physics hydrodynamical run, a gravity-only run using the same initial conditions was also produced.

We use the snapshots 52, 72, 96, 116, 144 which correspond to redshifts 1.18, 0.78, 0.47, 0.25, 0.00. Halos are identified using a modified version of the SUBFIND algorithm with a linking length $b = 0.16$ (Springel et al. 2001; Dolag et al. 2009). Spherical over-density masses are then computed out from the potential minimum of each halo. For each snapshot, we extract halo catalogues from the gravity-only simulation down to $M_{200c} > 1.56 \times 10^{14}\,h^{-1}$ M$_\odot$. With this cut, the region within r_{200c} is comfortably resolved with at least 10⁴ particles of each species on average. For the hydrodynamical run, we set a limit that is lower by a factor of 0.7. Starting from the gravity-only halo catalogues, we assign each halo a counterpart from the hydrodynamical catalogues matching by halo position. To avoid spurious matches, we require that the centre offsets are within two times the virial radius of the halo in the gravity-only catalogue. This cut removes less than 0.5 per cent of all haloes. The scatter in M_{200c} between both catalogues is within 30 per cent; this motivated the choice to extract haloes from the hydrodynamical run down to 0.7 of the mass cut in the gravity-only simulation. As an example, we illustrate the matched catalogues for $z = 0.25$ in Fig. 2. In this snapshot, the hydro

[1]http://www.magneticum.org/index.html
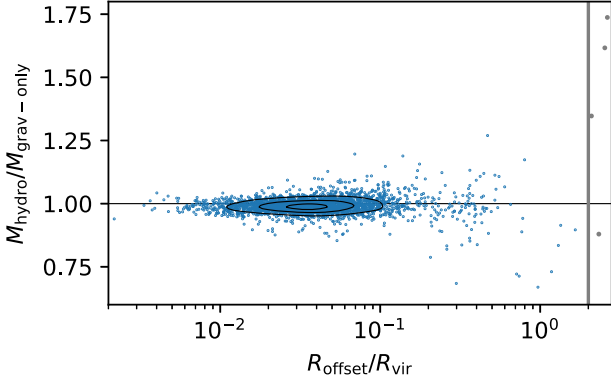
Figure 2. Mass ratio hydro/gravity-only as a function of centre offset for $z = 0.25$. Points show all haloes in the snapshot above the mass cut M_{200c}(gravity-only) $> 1.56 \times 10^{14}\, h^{-1}\, M_\odot$. Contours enclose 30 per cent, 60 per cent, and 90 per cent of all points. The vertical grey line shows the cut $R_{\mathrm{offset}} = 2R_{\mathrm{vir}}$.

masses are on average 0.99 ± 0.03 lower than their gravity-only counterparts.

To investigate how much our findings depend on the specifics of the Magneticum simulations, we use the Illustris TNG300-1 simulations as a cross-check (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019; Pillepich et al. 2018; Springel et al. 2018). These include 2×2500^3 resolution elements for a box size of 302.6 Mpc on a side. While TNG300-1 features a higher mass resolution than Magneticum Box1 its simulation volume is significantly smaller and therefore only tracks a small number of high-mass haloes. The cosmology corresponds to the Planck2015 constraints for a spatially flat ΛCDM cosmology (Planck Collaboration 2016): $\Omega_{\mathrm{m}} = 0.3089$, $\Omega_{\mathrm{b}} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$, and $h = 0.6774$. The TNG simulations feature a full-physics hydrodynamical run (TNG300-1) and a gravity-only run (TNG300-1-Dark) that uses the same initial conditions. Because we are only interested in spot-checking our results obtained from the Magneticum simulations, we restrict ourselves to using the TNG300-1 snapshots 51 and 98, which correspond to redshifts 0.95 and 0.01. Halos are identified via a FOF algorithm and spherical overdensity masses are built out from the halo potential minimum. We extract haloes and match the hydrodynamical and gravity-only halo catalogues following the procedures detailed above with the exception that the mass limit for TNG300-1-Dark is set to $M_{200c} > 5 \times 10^{13}\, h^{-1}\, M_\odot$; the limit for the hydrodynamical TNG300-1 is 0.7 times that value. The cross-check between results obtained from the Magneticum simulations and TNG are presented in Section 3.4 and Appendix A.

2.1.2 Convergence and shear maps

We now extract cylindrically projected (total) matter density maps centred on the halo potential minimum for the haloes in the simulations. Due to the steepness of the halo mass function, the halo catalogues have far more low mass haloes. In our analysis we seek to explore halo mass trends with approximately equal fidelity over a broad range of masses. Therefore, at low halo masses, we randomly sub-sample our halo catalogues such that each mass bin of width $\Delta \log_{10} M_{200c} = 0.1$ has at most about 100 haloes. Because extracting the projected matter maps from the simulations is expensive, this sub-sampling significantly reduces the cost of our analysis, while still ensuring enough statistical constraining power (cf. Section 2.3.2).

The cylindrical projection depth is set to $l_{\mathrm{proj}} = \pm 20\, h^{-1}\mathrm{Mpc}$, which is sufficient to capture the effects of correlated large-scale structure (Becker & Kravtsov 2011). We bin the map in 20 kpc $h^{-1} \times$ 20 kpc h^{-1} quadratic bins, providing a good balance between sufficient sampling for the $0.2\, h^{-1}$ Mpc minimal scale we are interested in (cf. Section 3.2), and memory usage. For each halo, we produce three (almost) independent matter density maps by projecting along the three Cartesian coordinate axes. The left-hand panel of Fig. 3 shows one of the matter density maps of an exemplary halo.

For each projection, we then compute the scaled shear components by applying the Kaiser–Squires algorithm (Kaiser & Squires 1993),

$$\Gamma_i = \mathcal{F}^{-1}\left[\chi_i(\boldsymbol{k})\,\mathcal{F}[\Sigma]\right], \ \text{with} \ \boldsymbol{\chi}(\boldsymbol{k}) = \frac{-1}{|\boldsymbol{k}|^2}\begin{pmatrix} k_1^2 - k_2^2 \\ k_1 k_2 \end{pmatrix}, \tag{1}$$

where $\mathcal{F}[\cdot]$ stands for the 2D Fourier transform, $\mathcal{F}^{-1}[\cdot]$ for the inverse 2D Fourier transform, and $\boldsymbol{k}$ is the associated wave vector. To obtain the actual shear map, the scaled shear needs to be divided by the critical density Σ_{crit} (cf. Section 2.1.5). Given that we wish to perturb the critical density to account for photometric redshift uncertainties, we save the scaled shear at this stage of the analysis.

We then define 15 offset centres for each projection direction, for each halo, at 15 different mis-centreing radii, equally log-spaced between $(0.01, 1)\, h^{-1}$ Mpc. These centres are positioned isotropically around the true centre. For each offset centre, as well as for the true centre, we compute the scaled tangential shear,

$$\Gamma_t(\boldsymbol{x}) = -\Gamma_1(\boldsymbol{x})\cos(2\phi) - \Gamma_2(\boldsymbol{x})\cos(2\phi), \tag{2}$$

where ϕ is the polar angle of $\boldsymbol{x}$. An example of such a mis-centred scaled tangential map is shown in the right-hand panel of Fig. 3. Note the pronounced azimuthal anisotropy around the offset centre (black star). In directions perpendicular to the mis-centreing direction, the tangential shear is negative. This anisotropy needs to be traced to compute the correct reduced shear profile (cf. Section 2.1.5).

To efficiently handle the large amount of data – several 100 haloes, times 3 projections, times $15 + 1$ centres – we need to compress the map information in a way that conserves the azimuthal anisotropy. To this end we re-bin the surface mass density maps and the scaled tangential shear maps in radial bins, equally log-spaced between $(0.01, 10)\, h^{-1}$ Mpc, and with 24 bins in azimuthal angle ϕ. The resulting re-binned maps are shown in Fig. 4. These maps are significantly compressed in comparison to the Cartesian maps, while capturing the relevant information (radial trend, azimuthal anisotropy) needed to produce realistic reduced shear profiles.

To transform the matter profile library into a shear profile library, we must specify several characteristics of the WL survey. These include: photometric redshift bias, shear measurement bias, impact of uncorrelated large-scale structure, cluster member contamination together with the halo catalogue cluster observables, and mis-centreing distribution. The values of these properties are subject to systematic uncertainties, which in turn contribute to the final systematic halo mass uncertainty. We parametrize these additional elements, allowing us to marginalize over these uncertainties. Note also, that the WL signal is not well defined for $z = 0$. We therefore omit that snapshot when considering results based on shear profiles.

We choose the numerical values for these parameters by drawing from the Stage III DES Y1 WL survey (McClintock et al. 2019a) and the *South Pole Telescope* cluster survey (Bleem et al. 2015). The values need to be adjusted to the specific WL survey and cluster catalogue before application in cosmological analyses.
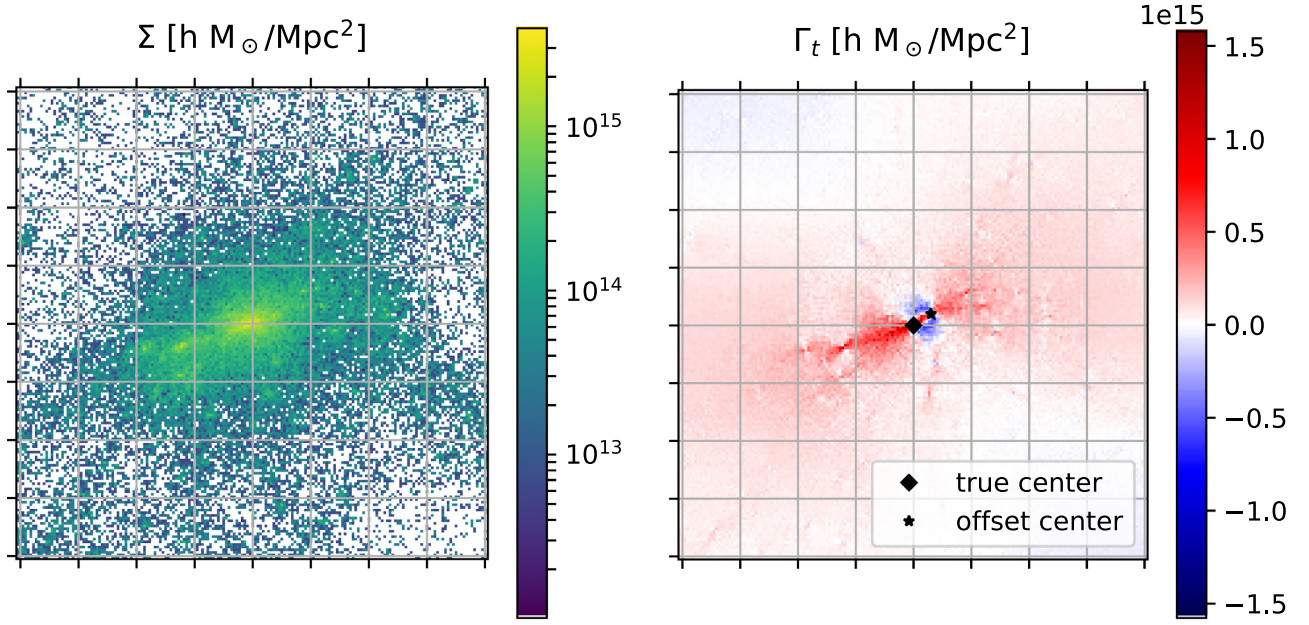
Figure 3. Example surface matter density map (left-hand panel) around a halo in the hydrodynamical simulation. The grid denotes $1\,h^{-1}\,\mathrm{Mpc} \times 1\,h^{-1}\,\mathrm{Mpc}$ squares. To capture the impact of correlated structure along the line of sight, we project along $\pm 20\,h^{-1}\,\mathrm{Mpc}$ in the perpendicular direction. After applying the Kaiser–Squires algorithm to obtain scaled versions of the shear components $\Gamma_{1,2}$ along the Cartesian coordinates, we choose isotropically oriented, offset centres (star in right-hand panel), around which we compute the scaled tangential shear Γ_t in the presence of mis-centreing.
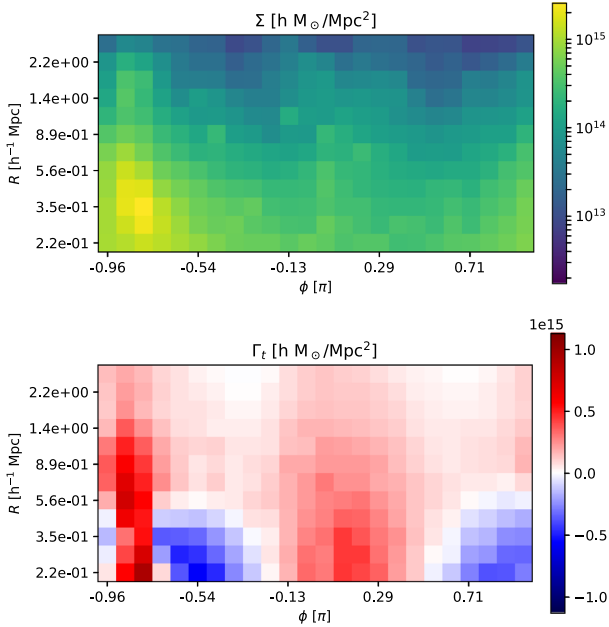
Figure 4. To compress the surface density map (upper panel) and the scaled tangetial shear map (lower panel), while conserving information on the azimuthal anisotropy of the maps, we define bins in polar coordinates, equally spaced in log-radius R and azimuthal angle ϕ around the offset centre. Combined with the critical surface density Σ_crit (cf. Section 2.1.5), we can compute the reduced shear profile while accounting for the azimuthal anisotropy of surface matter density and the tangential shear.

2.1.3 Cluster observables

Many WL systematics (such as the cluster member contamination and the mis-centreing distributions) are empirically calibrated as functions of cluster observables. We thus need to assign such observables if we want to realistically model the WL systematics. Following Saro et al. (2015), Bleem et al. (2020), and Grandis et al. (2020), we assign a richness λ by drawing

$$\ln \lambda \sim \mathcal{N}\left(\langle \ln \lambda\rangle(M_{200\mathrm{c}}, z); \sigma_\mathrm{tot}^2\right), \tag{3}$$

with

$$\langle \ln \lambda\rangle(M_{200\mathrm{c}}, z) = \ln A_\lambda + B_\lambda \ln \left(\frac{M_{200\mathrm{c}}}{3e14\,h^{-1}\,\mathrm{M}_\odot}\right) + C_\lambda \ln \left(\frac{E(z_\mathrm{cl})}{E(0.6)}\right), \tag{4}$$

where $E(z)$ is the critical density of the Universe in units of the present day critical density, and with

$$\sigma_\mathrm{tot}^2 = \exp\left(2 \ln \sigma_\lambda\right) + \frac{\exp\langle \ln \lambda\rangle(M_{200\mathrm{c}}, z) - 1}{\exp(2\langle \ln \lambda\rangle(M_{200\mathrm{c}}, z))}. \tag{5}$$

The parameters $(A_\lambda,\ B_\lambda,\ C_\lambda,\ \ln \sigma_\lambda)$ parametrize the normalization, mass trend, redshift trend, and logarithmic scatter in the λ–mass relation, whose systematic uncertainties are reflected as parameter uncertainties.

2.1.4 Mis-centreing distribution

When extracting shear profiles in real data, the chosen centre does not coincide with the halo centre. To properly assess how probable each of our mis-centreing radii is, we adopt the following mis-centreing distribution:

$$P(R_\mathrm{mis}|\lambda) = \rho\,\mathrm{Rayl}\left(\frac{R_\mathrm{mis}}{R_\lambda}; \sigma_0\right) + (1-\rho)\,\mathrm{Rayl}\left(\frac{R_\mathrm{mis}}{R_\lambda}; \sigma_1\right), \tag{6}$$

with $R_\lambda = (\lambda/100)^{0.2}\,h^{-1}\,\mathrm{Mpc}$. This is a two-component Rayleigh distribution that provides a good description of the mis-centreing of optical centres such as the central brightest cluster galaxy (Saro et al. 2015; Bleem et al. 2020) with respect to true halo centres. A large fraction ρ of clusters contain well centred objects (typically $\sigma_0 < 0.1$), and a smaller sub-population of disturbed and thus strongly mis-centred clusters (typically $\sigma_1 > 0.1$) show large mis-centreing

effects (see also optical–X-ray studies such as Lin et al. 2004). In general, the strength of the mis-centreing depends on some cluster observables, such as richness. The parameters of the mis-centreing distribution in this case are $(\rho, \sigma_0, \sigma_1)$.

2.1.5 Photo-z uncertainty in source redshifts

We assume that our generic survey has a source redshift distribution $P(z_s) = 0.5\, z_{s,0}^{-3} z_s^2 \exp(-z_s/z_{s,0})$, with $z_{s,0} = 0.2$ (following the parametrisation suggested by Smail, Ellis & Fitchett 1994). For a cluster at redshift z_{cl} we model the background selection by imposing the cut $z_s > z_{cl} + 0.1$. In a realistic case, the real source redshift distribution of a specific survey would have to be used. Also the background selection method could be different.

Given a source redshift distribution and a background selection we compute the lensing efficiency

$$\Sigma_{crit}^{-1}(z_{cl}) = \frac{4\pi G d_A(z_{cl})}{c^2} \left\langle \frac{d_A(z_{cl}, z_s)}{d_A(z_s)} \right\rangle_{z_s > z_{cl}+0.1}$$
$$\left(1 + \delta_\beta(z_{cl}) + \alpha_\beta \sigma_\beta(z_{cl})\right), \tag{7}$$

where $\delta_\beta(z_{cl}) \pm \sigma_\beta(z_{cl})$ is an estimate of the bias on the lensing efficiency due to photometric redshift measurements, together with the uncertainty on this bias. For a deep photometric survey aiming at measuring cosmic shear, this quantity is one of the most relevant systematics, and thus a natural calibration product (e.g. Hoyle et al. 2018). For our generic survey we assume

$$\delta_\beta(z_{cl}) = \begin{cases} 0.025 & \text{for } z_{cl} < 0.7 \\ 0.05 - 0.025\left(\frac{1+z_{cl}}{1.7}\right)^7 & \text{for } z_{cl} > 0.7, \end{cases} \tag{8}$$

and

$$\sigma_\beta(z_{cl}) = 0.02 \left(\frac{1+z_{cl}^2}{1.49}\right)^2. \tag{9}$$

This prescription qualitatively follows the photometric redshift bias and uncertainty in the DES Y1 data which at $z_{cl} \gtrsim 0.7$ grow quickly (McClintock et al. 2019a). The parameter α_β is introduced to vary the strength of the photo-z bias within its errors. By construction, its mean value is 0, while its variance is 1.

The lensing efficiency allows us to compute the convergence map, and the tangential shear map,

$$\kappa(R, \phi|R_{mis}) = \frac{\Sigma(R, \phi|R_{mis})}{\Sigma_{crit}(z_{cl})}, \text{ and } \gamma_t(R, \phi|R_{mis}) = \frac{\Gamma_t(R, \phi|R_{mis})}{\Sigma_{crit}(z_{cl})}, \tag{10}$$

while accounting for the azimuthal anisotropy of both maps. Besides the natural anisotropy introduced by the halo ellipticity and morphology even for well-centred cases, assuming offset centres strongly contributes to the anisotropy, as can be seen in Figs 3 and 4.

The reduced shear profile is then obtained by first computing the reduced shear map, accounting for azimuthal anisotropy, and then averaging over azimuthal angles,

$$g_t(R|R_{mis}) = \int d\phi \frac{\gamma_t(R, \phi|R_{mis})}{1 - \kappa(R, \phi|R_{mis})}. \tag{11}$$

Inverting this precise order – reduced shear map first, azimuthal average second – leads to errors in the reduced shear profile of a few per cent (especially around the position of the true centre) and to ≤ 1 per cent level shifts in the inferred WL bias.

2.1.6 Uncorrelated large-scale structure

The next step in the creation of a realistic shear measurement is to add uncorrelated large-scale structure projections to the shear profile. Using the prescription presented by Hoekstra (2003), we compute the covariance of the large-scale structure $C_{LSS}(R, R')$. We model this effect by drawing a Gaussian deviate around the reduced shear profile $\tilde{g}(R|R_{mis}) \sim \mathcal{N}(g_t; C_{LSS})$. Importantly, this approach accounts for the line-of-sight projections due to structures that are not correlated with the target halo. This effect depends on the cosmological model through the matter power spectrum and the distance sensitivity of the lensing kernel. It is, however, independent on the cluster mass. Furthermore, in current Stage III lensing surveys, this effect in much smaller than the shape noise. The correlated structure effects are already captured within the projected matter density profiles extracted from the simulations (cf. Section 2.1.2).

2.1.7 Multiplicative shear bias

The measured reduced shear results from the predicted reduced shear after accounting for the multiplicative shear bias introduced by the shape measurements. Following Sheldon & Huff (2017), we model this as

$$g_{t,\, meas}(R|R_{mis}) = \left(1 + m + \alpha_\gamma \sigma_m + \alpha_{NL} g_t^2(R|R_{mis})\right)$$
$$g_t(R|R_{mis}), \tag{12}$$

where m is the multiplicative shear bias, σ_m its error. We set $m = \sigma_m = 0.01$, which are typical values for Stage III lensing surveys. We introduce the parameter α_γ to modulate the strength of the multiplicative shear bias. We also introduce a parameter that accounts for the non-linear shear response α_{NL}. To date, image simulations used to calibrate the shape measurements focus most attention on accurate calibration of the multiplicative shear bias m and its uncertainty σ_m while little attention is given to the performance of shape measurements in the large shear regime.

2.1.8 Cluster member contamination

The contamination of the background sample by cluster member galaxies leads to a suppression of the measured shear signal, because the cluster members are not sheared by the cluster potential. Several techniques have been proposed to calibrate this effect empirically, and they all result in a cluster member contamination profile f_{cl}, which naturally has uncertainties on its parameters (see Varga et al. 2019 and references therein). We choose here the parametrization

$$f_{cl}(R|\lambda) = A_{fcl}\left(\frac{\lambda}{67}\right)^{B_{fcl}} \frac{\tilde{f}(x)}{\tilde{f}(1)}, \text{ with } x = \frac{c_{fcl} R}{R_\lambda}, \tag{13}$$

where $\tilde{f}(x)$ is the radial trend of a projected NFW (which we normalize at the scale radius), and $R_\lambda = (\lambda/100)^{0.2}\, h^{-1}$ Mpc. The parameters $(A_{fcl}, B_{fcl}, c_{fcl})$ are the amplitude, the richness trend and the concentration of the cluster member contamination profile, respectively. These parameters can be constrained well even in rather small samples of, e.g. SZE or X-ray selected clusters (Paulus et al., in preparation). The concentration typically takes a low value, $c \approx 2.5$, while the redshift trend is weak, $B_{fcl} \approx 0.15$, and the amplitude $A_{fcl} \approx 0.3$ represents 30 per cent contamination at the scale radius. Varying the parameters $(A_{fcl}, B_{fcl}, c_{fcl})$ then accounts for the systematic uncertainties associated with the cluster member contamination.

In our application, we produce the synthetic shear profile for the i-th cluster by diluting the measured shear profile by the cluster
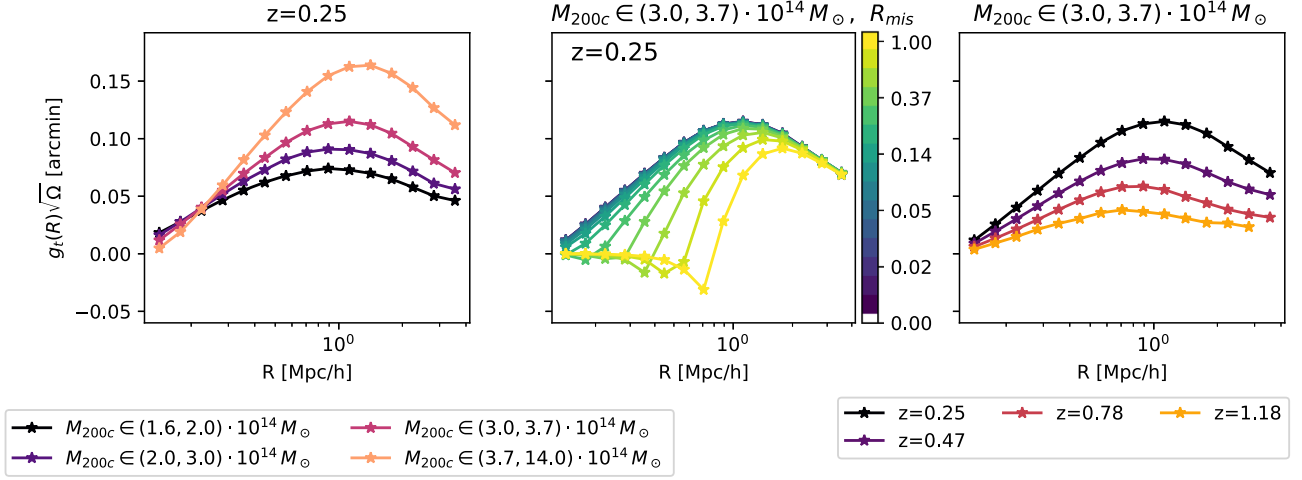
Figure 5. Shear profiles stacked in mass bins multiplied with the square root of the area in degrees, visualized for different dimensions of our shear profile library: mass bins (left-hand panel), mis-centring radius used for extraction (centre), redshift of the snapshots (right). The amplitude of the signal can be compared to the typical shape-noise of $\sim 0.3/\sqrt{n_\epsilon}$, where n_ϵ is the source density per arcmin^2. We construct this shear profile library for a given set of WL survey specifications and systematics parameters. Varying these parameters allows us to sample the space of WL systematics.

member contamination,

$$\hat{g}_t^i(R|R_{\text{mis}}) = \left(1 - f_{\text{cl}}(R|\lambda^i)\right) g_{t,\text{ meas}}^i(R|R_{\text{mis}}). \tag{14}$$

2.1.9 Library creation

Going through the steps outlined in the previous sections, we can simulate a realistic cluster catalogue with an observable richness λ and a mis-centreing distribution for each halo. Furthermore, for each halo we also compute three shear profiles – one for each cylindrical projection along each Cartesian axis – for each of the mis-centreing radii we consider. This shear library is created for a set of WL systematics and cluster parameters drawn from priors, which are presented in Section 2.3.3. Fig. 5 shows one realization of such a shear library: in the left-hand panel stacked in different mass bins, in the centre showing the impact of different mis-centreing radii, and in the right-hand panel at different redshift. We present the shear profiles weighed by the square root of the area in units of degrees. The reported magnitude of the signal can thus be directly compared to the shape noise in the tangential shear measurement, which goes as $\sim 0.3/\sqrt{n_\epsilon}$, where n_ϵ is the source galaxy density in deg^{-2}. This weighting highlights the radial range in which the number of shapes and the strength of the density contrast provide the most signal, and it cancels the redshift trend introduced by the smaller solid angle subtended by a radial bin of fixed metric size.

2.2 Shear profile model and measurement of M_{WL}

To capture the statistical and systematic uncertainties in the WL measurement, we measure the WL mass M_{WL} and quantify its relationship to the true halo mass. The WL mass results from fitting the simulated (or measured) shear profiles on to a shear profile model (Applegate et al. 2014; Schrabback et al. 2018; Dietrich et al. 2019; Stern et al. 2019). As such, the WL mass is dependent on the shear profile model. This gives us the freedom to define a model that is simple to compute at the expense of not being the most accurate on a cluster-by-cluster basis., but might not be the most accurate on a cluster by cluster basis. These inaccuracies lead to biases and scatter

between the WL and halo mass. The ultimate goal of our method is a robust determination of this bias and scatter and their respective uncertainties.

In the following we discuss the shear profile model we use to compress the complex and realistic shear profiles simulated in our shear library into WL masses. First, we extract the WL masses assuming a NFW model $\Sigma_{\text{NFW}}(R|M_{200c})$, and a constant concentration $c_{\text{NFW}} = 3.5$, a typical concentration for massive haloes (e.g. Child et al. 2018). Fixing the concentration to a constant value simplifies the WL mass extraction in the cosmological analysis at the cost of increasing the WL bias and scatter.

Secondly, we use a simplified model for the mis-centred matter surface density

$$\Sigma(R|M, R_{\text{mis}}) = \begin{cases} \Sigma_{\text{NFW}}(R_{\text{mis}}|M_{200c}) \text{ for } R < R_{\text{mis}} \\ \Sigma_{\text{NFW}}(R|M_{200c}) \text{ for } R > R_{\text{mis}}. \end{cases} \tag{15}$$

This model is expected to lead to some inaccuracy in the radial ranges $R \sim R_{\text{mis}}$, which lead to calibrateable biases between M_{WL} and the halo mass. However, our model simplifies the cosmological inference, by avoiding the explicit calculation of the azimuthal average

$$\Sigma(R|R_{\text{mis}}, M_{200c}) =$$
$$= \frac{1}{2\pi} \int_0^{2\pi} d\theta \, \Sigma\left(\sqrt{R^2 + R_{\text{mis}}^2 - 2RR_{\text{mis}}\cos\theta} \middle| M_{200c}\right). \tag{16}$$

Thirdly, instead of marginalizing over the distribution of R_{mis} we define one mis-centreing radius for the extractions

$$R_{\text{mis}}^{\text{extr},i} = R_{\lambda^i} \sqrt{\frac{\pi}{2}} \left(\bar{\rho}\bar{\sigma}_0 + (1 - \bar{\rho})\bar{\sigma}_1\right), \tag{17}$$

that is the mean mis-centreing radius for the mis-centreing distribution evaluated at the mean mis-centreing parameters $(\bar{\rho}, \bar{\sigma}_{0,1})$, evaluated for the measured richness λ^i. Other works use centred NFW profiles for the WL mass measurement, leading to less well-behaved WL–halo mass distributions (Sommer et al. 2021). Given the model for the matter surface density $\Sigma(R|R_{\text{mis}}^{\text{extr},i}, M)$, we can easily compute the density contrast $\Delta\Sigma(R|R_{\text{mis}}^{\text{extr},i}, M_{200c})$.

Fourthly, we compute the lensing efficiency $\Sigma_{\text{crit}}^{-1}$ by setting $\alpha_\beta = 0$ in equation (7), i.e. just assuming the mean photo-z bias.
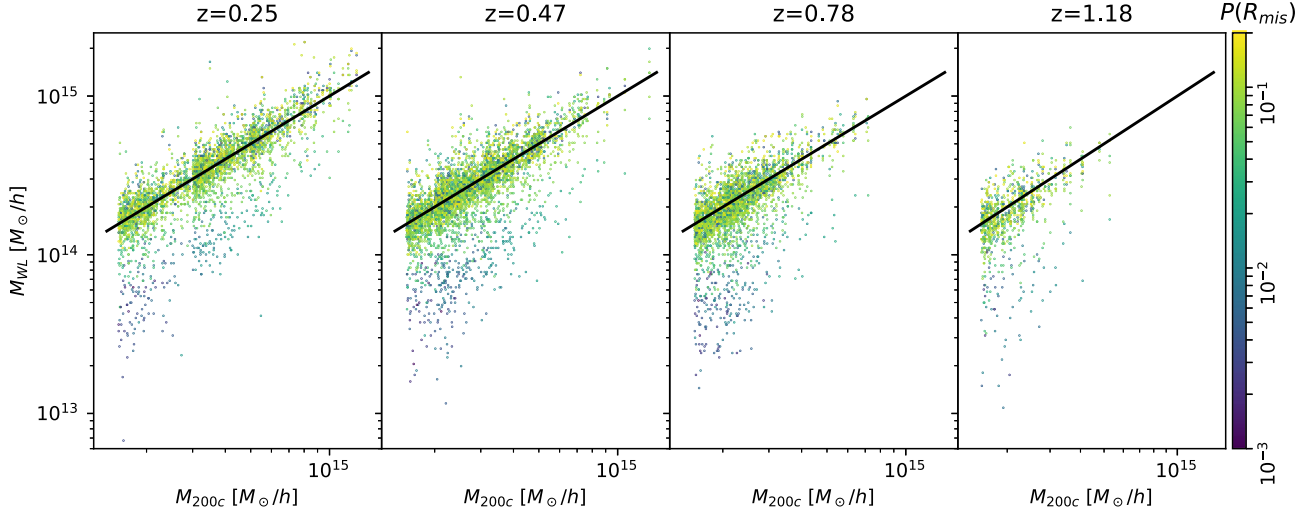
Figure 6. For a single realization of WL systematics we can extract the so called WL mass for each halo projection as a function of the mis-centreing used in the extraction by assuming a shear profile model. We show the WL mass versus halo mass in the four different simulation snapshots. Colour coded is the probability weight, derived from the probability of the mis-centreing. Strongly mis-centred simulated profiles have highly biased WL masses, but they are also very improbable. The black line shows the 1–1 relation.

Fifthly, we compute the convergence $\kappa = \Sigma_{\rm crit}^{-1} \Sigma(R | R_{\rm mis}^{{\rm extr},i}, M)$ and the shear $\gamma_{\rm t} = \Sigma_{\rm crit}^{-1} \Delta\Sigma(R | R_{\rm mis}^{{\rm extr},i}, M)$, as well as the reduced shear $g_{\rm t}(R | R_{\rm mis}^{{\rm extr},i}, M_{\rm 200c}) = \gamma_{\rm t}/(1 - \kappa)$. Our final model for the measured reduced shear then takes account of the mean cluster member contamination profile $\bar{f}_{\rm cl}(R | \lambda^i)$ and the mean multiplicative shear bias m, i.e.

$$g_{\rm t}^{\rm mdl}(R | M_{\rm 200c}, \lambda^i) = (1 + m)\big(1 - \bar{f}_{\rm cl}(R | \lambda^i)\big) g_{\rm t}(R | R_{\rm mis}^{\rm extr}, M_{\rm 200c}) \tag{18}$$

And sixthly, we measure the WL mass by minimizing the radial bin area–weighted difference between the simulated and model shear profiles. The WL mass for the i-th cluster, when taking the simulated profile mis-centred by $R_{\rm mis}$, then is

$$M_{\rm WL}^i(R_{\rm mis}) = \min_{M_{\rm 200c}} \sum_k A_k \big(g_{\rm t}^{\rm mdl}(R_k | M_{\rm 200c}, \lambda^i) - \hat{g}_{\rm t}^i(R_k | R_{\rm mis})\big)^2, \tag{19}$$

where k runs over the radial bins, and A_k is the area covered by that bin. The shape measurement variance in real measurements scales like the inverse of the bin area. Sommer et al. (2021) explicitly show how this setup produces unbiased WL mass estimates independently on the amount of shape noise. This weight guarantees that we are weighting the different scale in the same way as they will be weighted in the real data. Note also that we added the uncorrelated LSS noise to the shear profile, rather than considering it a noise source in the $M_{\rm WL}$ extraction. This implies that the uncorrelated LSS variance contributes to the WL scatter $\sigma_{\rm WL}$, as it is a statistical noise source on the shear profile. This configuration ensurers that while extracting $M_{\rm WL}$ we apply the correct relative weights to the different scales.

The choice of the innermost radial scale $R_{\rm min}$ and outermost radial scale $R_{\rm max}$ has a great impact on the accuracy of the WL mass extraction. As a baseline we choose $R_{\rm min} = 0.5 \, h^{-1}$ Mpc, and $R_{\rm max} = 3.2(1 + z_{\rm cl})^{-1} \, h^{-1}$ Mpc. We will explore the impact of varying the inner fitting radius and also argue for the redshift dependence of the outer fitting radius below (cf. Section 3.2, and Section 3.3).

This procedure provides us with 3 $M_{\rm WL}^i(R_{\rm mis})$ for each halo and for each $R_{\rm mis}$ thanks to the three projection axes. It is noteworthy
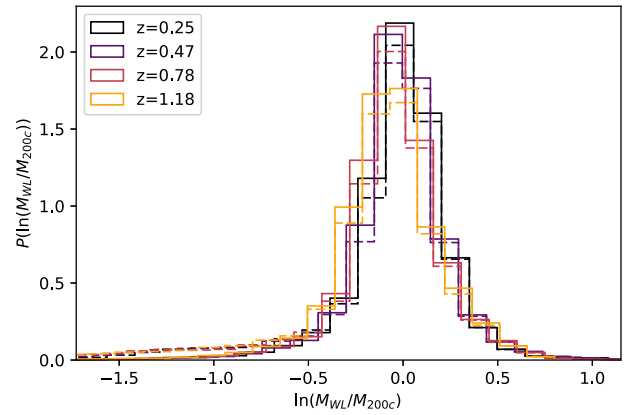
Figure 7. Distribution of WL mass $M_{\rm WL}$ around the halo mass for different redshifts (dashed line without the weighting by the mis-centreing distribution, full line with the weighting by the mis-centreing distribution) for a single realization of the shear library. Applying the weighting by the mis-centreing distribution down-weights the lowest extracted WL masses, which correspond to the largest mis-centreing radii. The resulting distribution is symmetric, although fat-tailed. We fit this distribution with a lognormal whose mean and variance can vary with halo mass and redshift, postponing the investigation of the tails to future work.

here that thanks to the mis-centreing distribution $P(R_{\rm mis} | \lambda^i)$ we also know how probable each mis-centreing radius $R_{\rm mis}$ is.

Fig. 6 shows the scatter plot between the WL mass and the halo mass in the five different snapshots we analysed. Colour coded is the associated probability $P(R_{\rm mis} | \lambda^i)$. Clearly, some WL masses underestimate the halo mass by more than an order of magnitude. These are associated to large mis-centreing radii $R_{\rm mis}$. However, such large mis-centreing radii are also highly improbable. Indeed, for the most probable mis-centreing radii, the WL mass is comparable to the halo mass, albeit with some scatter. Statistically speaking the highly biased WL masses due to strong mis-centreing are not very relevant, as can be seen in Fig. 7, which shows the probability weighted distribution of the ratio between WL mass and halo mass (solid line) in comparison to the raw distribution (dashed line). The

weighting suppresses the left tail of the distribution. The WL masses extracted for probable mis-centreing radii scatter around a ratio of ~1, indicating that our simple model for the shear profile provides an adequate fit to the data.

Following this approach, we are able to compress the complex shear profiles in our realistic shear profile library into WL masses. The model defining the WL mass does not explicitly take account of the different statistical noise sources impacting the simulated shear profile. These statistic noise sources thus manifest as scatter between the WL and the halo mass. Similarly, our model does not explicitly account for systematic uncertainties in the halo catalogue and shear profiles. Instead, fitting a relation for the WL–halo mass parameters with different realizations of WL systematics allows us to propagate their uncertainty.

2.3 Relation between WL and halo mass

2.3.1 Fitting

In the previous section, we established the WL mass such that it closely tracks the halo mass with some scatter and some (small) amount of bias. For the mean relation, we allow for a complex redshift trend $b_{WL}(z)$ and a power-law trend in mass, i.e.

$$\left\langle \ln\left(\frac{M_{WL}}{M_0}\right)\right\rangle = b(z) + b_M \ln\left(\frac{M_{200c}}{M_0}\right), \tag{20}$$

with the pivot mass $M_0 = 2 \times 10^{14}\,h^{-1}\,\mathrm{M_\odot}$. For the redshift evolution of the bias $b(z)$, we fit an independent bias parameter b_{z_i} for each simulation snapshot at $z_i \in (0.25, 0.47, 0.78, 1.18)$. For the scatter, we set up a lognormal model

$$P(\ln M_{WL}|M_{200c}, z) = \mathcal{N}\left(\ln M_{WL}|\langle \ln M_{WL}\rangle, \sigma_{WL}^2\right) \tag{21}$$

and we allow the variance to vary with mass and redshift

$$\ln \sigma_{WL}^2 = s(z) + s_M \ln\left(\frac{M_{200c}}{M_0}\right). \tag{22}$$

For the redshift evolution of the bias $s(z)$, we fit an independent bias parameter s_{z_i} for each simulation snapshot at $z_i \in (0.25, 0.47, 0.78, 1.18)$.

To fit for the parameters of the WL mass–halo mass relation we maximize the likelihood

$$\ln L = \sum_i \ln \sum_{R_{mis}} P(R_{mis}|\lambda^i) P\left(\ln M_{WL}^i(R_{mis})\big| M^i, z^i\right), \tag{23}$$

which ensures that the WL masses extracted from strongly mis-centred profiles are weighted by the mis-centreing probability. This approach implicitly marginalizes over the mis-centreing distribution, providing significant time savings in the cosmological pipeline.

We maximize the likelihood of the WL mass–halo mass relation given in equation (23), resulting in a best-fitting estimate of the parameters of the WL mass–halo mass relation (b_{z_i}, b_M, s_0, s_M, s'), and in a covariance matrix C_{sim} that describes the measurement uncertainties on these parameters that arise from the finite halo sample.

The scatter between WL mass and halo mass that we characterize here is due to several sources: variations in orientation and morphology of the cluster haloes (i.e, ellipticity, concentration, and substructure due to the diversity of dynamical states reflected in the simulations), correlated structure around the clusters, uncorrelated large-scale structure, and the variance stemming from the range of possible mis-centreing errors in choosing the cluster centre. Our simple model allows us to marginalize over all these effects,

producing posterior distributions in the WL mass to halo mass relation parameters that can then be incorporated within a WL cluster mass calibration analysis to account for systematic effects and the remaining uncertainties on those effects.

2.3.2 Accuracy of the WL mass extraction

The accuracy of the WL mass extracted using the described mapping are reflected in the statistical and systematic uncertainties at play in the approach described above. The uncertainties on the parameters (b_{z_i}, b_M, s_{z_i}, s_M) of the mapping between WL and halo mass stems in part from the finite number of haloes for which we extracted matter profiles from the simulations. We shall denote this uncertainty C_{sim} and note that these are 'statistical' in nature, in that they can be reduced by enlarging the profile library.

Another source of uncertainty comes from the systematic uncertainties in the adopted parametrizations of the WL catalogue, photometric source galaxy catalogue, and cluster properties. These include the parameters that govern the halo catalogue, i.e. the parameters of the observable–mass scaling relation (A_λ, B_λ, C_λ, $\ln \sigma_\lambda$) (cf. Section 2.1.3), and the parameters of the mis-centreing distribution (ρ, σ_0, σ_1) (cf. Section 2.1.4), and the parameters of the WL systematics, i.e. the modulation of the photometric redshift uncertainties α_β (cf. Section 2.1.5), linear and non-linear shape measurement uncertainties (α_γ, α_{NL}) (cf. Section 2.1.7), and the cluster member contamination (A_{fcl}, B_{fcl}, c_{fcl}) (cf. Section 2.1.8). By varying these parameters, we can create different realizations of shear libraries that sample the range of systematic uncertainty on the measured shear. To estimate the impact of these uncertainties, we draw from the best-fitting parameters and their uncertainties to create ~1000 different realizations of the parameters, following the prior listed below (cf. Section 2.3.3). We then use each realization to create a new shear profile library, to extract the WL masses through our fitting process and to then fit the WL to halo mass relation. This results in a estimate of WL bias and scatter for each realization of systematics. These result are then combined to build a joint posterior for the parameters of the WL to halo mass relation parameters (b_{z_i}, b_M, s_{z_i}, s_M) that reflects the uncertainties from the different systematics.

This posterior completely characterizes the impact of modelled WL systematics on the relation between WL and halo mass. The width of the posterior on the WL halo mass relation parameters can then be used for cosmological analyses. The uncertainty induced by varying the WL systematics is larger than the noise due to the limited number of haloes for which matter profiles were extracted from the simulations, $\sqrt{\mathrm{diag}(C_{sim})}$. As systematic uncertainties on the WL, photometric redshift, and halo related biases will be better characterized in the future, it will be necessary to scale up the size of the shear profile library extracted from the simulations.

2.3.3 Priors on WL systematics

Here we specify the ranges within which the WL systematics are allowed to vary. We caution the reader that the calibration of the WL mass extraction presented in the following is valid only for the specific shear profile model described in Section 2.2, and the systematics treatment described in this paper. Application to any other scenario requires a dedicated analysis, which would follow the same procedure outlined here.

For the parameters of the richness–mass relation, we follow the recent analysis by Bleem et al. (2020) $A_\lambda \sim \mathcal{N}(78.5; 8.2^2)$, $B_\lambda \sim \mathcal{N}(1.02; 0.08^2)$, $C_\lambda \sim \mathcal{N}(0.29; 0.27^2)$, $\ln \sigma_\lambda \sim \mathcal{N}(\ln 0.23; 0.16^2)$.

For the parameters of the mis-centreing distribution, we follow Saro et al. (2015) for the central values and the uncertainties for $\rho \sim \mathcal{N}(0.63; 0.06^2)$, $\sigma_0 \sim \mathcal{N}(0.07; 0.02^2)$, $\sigma_1 \sim \mathcal{N}(0.25; 0.07^2)$, but we reduce the uncertainty by half.

The parameters $\alpha_{\beta,\,\mathrm{m}}$ are introduced with the purpose of marginalizing over the multiplicative shear bias uncertainty (cf. Section 2.1.7) and the photo-z bias uncertainty (cf. Section 2.1.5). This is accomplished by sampling them $\alpha_{\beta,m} \sim \mathcal{N}(0; 1)$. For the strength of the non-linear shear response, we rely on the argument given in Sheldon & Huff (2017). For the DES Y1 data, they claim that $\alpha_{\mathrm{NL}} = 0.6$ with 40 per cent uncertainty, i.e. $\ln \alpha_{\mathrm{NL}} \sim \mathcal{N}(\ln 0.6; 0.4^2)$, would not bias the cosmic shear results.

For the cluster member contamination, we assume $\ln A_{\mathrm{fcl}} \sim \mathcal{N}(\ln 0.3; 0.1^2)$, $B_{\mathrm{fcl}} \sim \mathcal{N}(0.15; 0.1^2)$, $c_{\mathrm{fcl}} \sim \mathcal{N}(2.5; 0.1^2)$. This comes down to a ~ 30 per cent contamination at the scale radius, with a weak richness trend. The radial trend is an NFW whose concentration is comparable to the concentration of blue galaxies in clusters (see Hennig et al. 2017), as blue galaxies are plausibly expected to contribute more to the cluster member contamination due to their lower quality photo-z's.

2.3.4 Tracking the sources of uncertainty

Consider a multivariate Gaussian distribution with means 0 and a covariance $C_{i,j} = \sigma_i \sigma_j \rho_{i,j}$, with $\rho_{i,j} = 1$ if $i = j$. The variance on x_k is $\mathrm{Var}[x_k] = \sigma_k^2$, and the variance on x_k conditional on x_l is $\mathrm{Var}[x_k | x_l] = \sigma_k^2(1 - \rho_{k,l}^2)$. This can be interpreted as 'knowing x_l reduces the variance on x_k by $\rho_{k,l}^2$ per cent'. Conversely, this means that the uncertainty on x_l propagates to an increase in the variance of x_k equal to $\sigma_k^2 \rho_{k,l}^2$. Reporting the squared correlation coefficients between the WL–halo mass parameters (b_{z_i}, b_{M}, s_{z_i}, s_{M}) and the WL systematics parameters (A_λ, B_λ, C_λ, $\ln \sigma_\lambda$, ρ, σ_0, σ_1, σ_0', $\sigma_1' \alpha_\beta$, α_γ, α_{NL}, A_{fcl}, B_{fcl}, c_{fcl}) thus provides a means of quantifying the impact of each systematics parameter on each of the WL to halo mass parameters.

We estimate the squared correlation coefficient between the parameters x and y from the sample of $\{x_i, y_i\}$ as

$$\rho_{x,y}^2 = \frac{N^{-1} \sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\mathrm{Var}[x]\,\mathrm{Var}[y]}}, \qquad (24)$$

with the variance on x estimated as

$$\mathrm{Var}[x] = N^{-1} \sum_i (x_i - \bar{x})^2. \qquad (25)$$

Here N is the number of samples used, and $\bar{x}$, and $\bar{y}$ are the means of x, and y, respectively.

2.3.5 Skewness and kurtosis

To test whether the assumed lognormal model in equation (21) is adequate, we compute the skewness and kurtosis of the residuals between the extracted WL mass $\ln M_{\mathrm{WL}}^i(R_{\mathrm{mis}})$ and the predicted WL mass $\langle \ln M_{\mathrm{WL}} \rangle(M_{200c}^i, z^i)$. For simplicity we define the residual $\Delta \ln M_{\mathrm{WL}}^i(R_{\mathrm{mis}}) = \ln M_{\mathrm{WL}}^i(R_{\mathrm{mis}}) - \langle \ln M_{\mathrm{WL}} \rangle(M_{200c}^i, z^i)$. We also need to consider that each data point has a weight $w(R_{\mathrm{mis}}, i) = P(R_{\mathrm{mis}} | \lambda^i)$.

The skewness of a weighted sample is given by

$$\gamma_{\mathrm{skw}} = \frac{1}{W} \sum_{R_{\mathrm{mis}},i} w(R_{\mathrm{mis}}, i)\big(\Delta \ln M_{\mathrm{WL}}^i(R_{\mathrm{mis}})\big)^3, \qquad (26)$$

while the kurtosis is given by

$$\gamma_{\mathrm{krt}} = \frac{1}{W} \sum_{R_{\mathrm{mis}},i} w(R_{\mathrm{mis}}, i)\big(\Delta \ln M_{\mathrm{WL}}^i(R_{\mathrm{mis}})\big)^4, \qquad (27)$$

with $W = \sum_{R_{\mathrm{mis}},i} w(R_{\mathrm{mis}}, i)$.

We estimate the statistical error on the skewness and kurtosis via bootstrapping. We also study the impact of the uncertainty on the WL systematics on skewness and kurtosis, by tracking them for each realization of shear library we study. We then consider deviations of the skewness and kurtosis from the Gaussian case ($\gamma_{\mathrm{skn}} = 0$ and $\gamma_{\mathrm{krt}} = 3$) as indicators if our assumed lognormal scatter model is adequate.

3 RESULTS AND DISCUSSION

In the following we will present the results from the calibration of the WL bias and scatter of our generic WL survey (Section 3.1). We then discuss the choice of the inner and outer fitting radius (Section 3.2 and Section 3.3, respectively), and the impact of further systematics that are harder to model to date (Section 3.4). Finally, we put our framework for dealing with systematics related to hydrodynamical effects in context with other works (Section 3.5).

3.1 Calibration for generic catalogue and WL survey

We analyse the generic survey described above with three different settings for the minimal fitting radius: $R_{\mathrm{min}} = 0.2, 0.5, 0.8\,h^{-1}\,\mathrm{Mpc}$. This results in the three posterior distributions whose 1-d and 2-d marginals are shown in Fig. 8. The 1-d median parameter values and their difference with the 16th and 84th percentile values are reported in Table 1 and shown in Fig. 9, left-hand panel. Our WL mass is biased low by 0–13 per cent, depending mainly on redshift. We find a mass trend close to linear.

The exact numerical values of the biases are subject to the details of the systematics implemented in the creation of the generic survey (cf. Section 2.1) and the specific model choices used to measure the WL mass (cf. Section 2.2). In particular, the biases could be reduced by adjusting our shear model to better represent the simulated shear profiles, but as long as the bias remains relatively small it is the uncertainty on this mean bias rather than the bias itself that serves as a limitation on the ensuing WL mass calibration.

The intrinsic scatter about the WL to halo mass relation is between 0.2 and 0.3 in the natural logarithm of the WL mass. It depends mainly on the inner fitting radius; for smaller inner fitting radii the intrinsic scatter is larger. The intrinsic scatter shows no clear trend with mass, but the uncertainties are large. A shear model that minimizes this variance would be crucial to any WL mass calibration, because a smaller variance effectively increases the mass information coming from each WL mass measurement.

For the smallest fitting radius considered ($R_{\mathrm{min}} = 0.2\,h^{-1}\,\mathrm{Mpc}$) the distribution of WL masses given halo mass is negatively skewed, while it shows no skewness for $R_{\mathrm{min}} = 0.5\,h^{-1}\,\mathrm{Mpc}$ and a positive skewness for $R_{\mathrm{min}} = 0.8\,h^{-1}\,\mathrm{Mpc}$. As shown in Appendix A, the WL mass in absence of mis-centreing is also positively skewed, which matches the large minimal fitting radius case (in which the impact of mis-centreing is minimal). As can be seen in Fig. 6, strong mis-centreing introduces WL masses that scatter low very strongly. When moving to smaller fitting radii, the impact of mis-centreing increases. This manifests itself in the negative skewness for the $R_{\mathrm{min}} = 0.2\,h^{-1}\,\mathrm{Mpc}$ case. For the current setting, the two effects seem to cancel in the $R_{\mathrm{min}} = 0.5\,h^{-1}\,\mathrm{Mpc}$ case, leading to
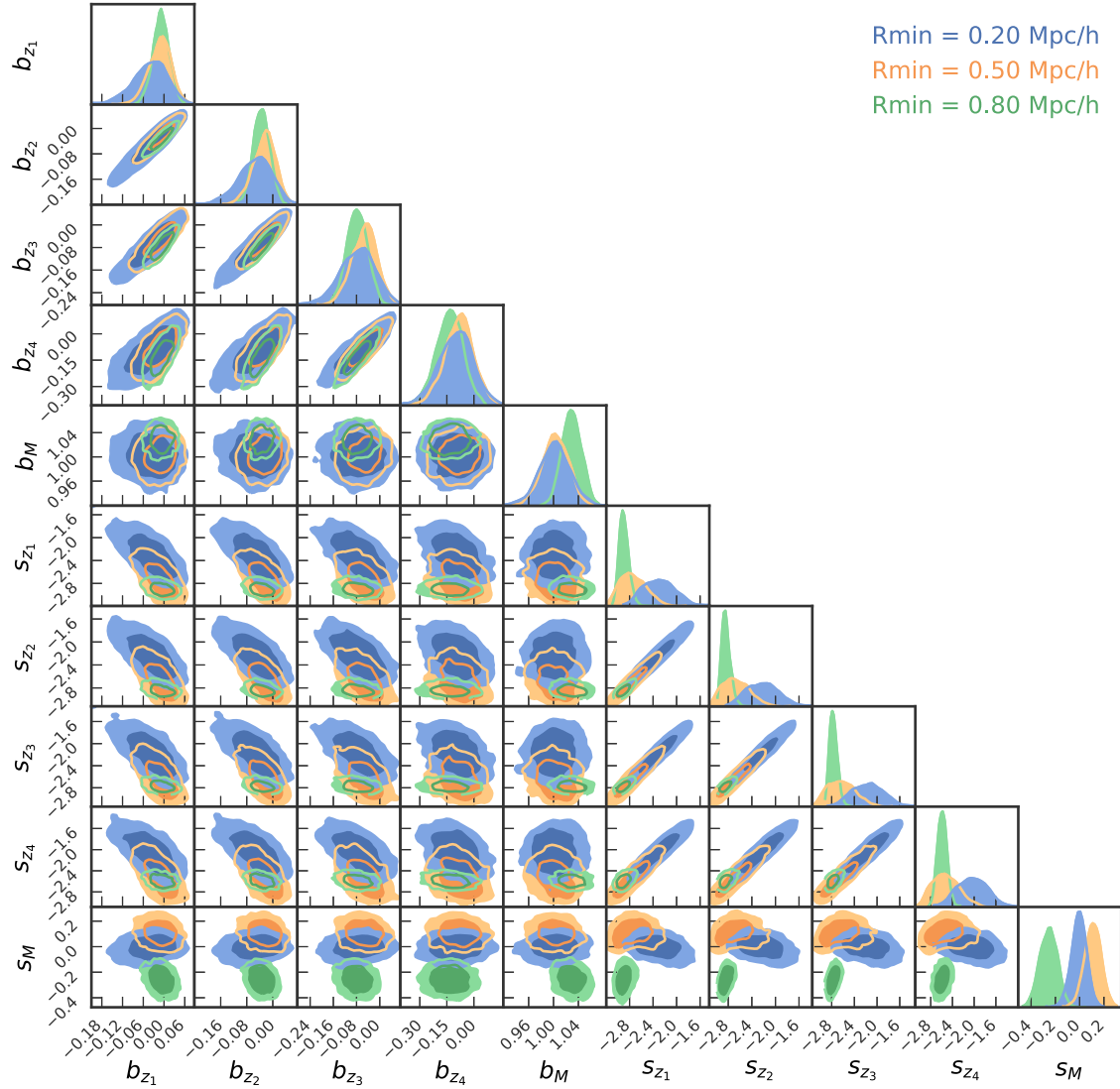
Figure 8. Posterior on the WL to halo mass relation parameters (equations 20–22) extracted from the analysis of ∼1000 different realizations of shear libraries with different WL and cluster catalogue systematics. b_{z_i} are the biases between the WL and the halo mass for each simulation snapshot at $z_i \in$ (0.25, 0.47, 0.78, 1.18), b_M is the mass trend of these biases, while s_{z_i} are the natural logarithms of the variance of the WL mass at given halo mass, and s_M is its mass trend. The posterior displays strong degeneracies among different parameters. It encapsulates the systematic uncertainty on the mapping between WL and halo mass, and can be used as a prior for cosmological analyses. We present the posterior for different choices of radial fitting range. Generally, using smaller scales for the fit increases the uncertainty on the WL to halo mass scaling parameters because central cluster regions are more affected by, e.g. mis-centreing and cluster contamination.

no detectable skewness, but this might not be the case for different mis-centreing distributions.

Note also that the kurtosis is larger than 3 (the Gaussian case) for all settings studied here. That means that distribution of WL masses given halo mass is fat-tailed. The impact of these tails will be investigated in future work.

Perhaps the most important result is the uncertainty on the different parameters of the WL to halo mass relation. They reflect uncertainties on the WL systematics parameters, and these ultimately limit the accuracy and precision with which a WL mass calibration can be done. We find that for a inner fitting radius of $R_{\min} = 0.2$ (0.5, 0.8) h^{-1} Mpc the uncertainty on the bias parameters increases with redshift from 4.8 per cent (3.1 per cent, 2.3 per cent) at z = 0.25–9.4 per cent (7.8 per cent, 7.4 per cent) at z = 1.18. The same trend can also be seen for the mass trend of the bias and the intrinsic

variances. This quantitatively confirms an important aspect of the treatment of systematics in cluster WL studies: the closer one fits to the cluster centre, the larger the uncertainties on the WL to halo mass relation parameters and therefore the larger the systematics budget in a WL mass calibration analysis.
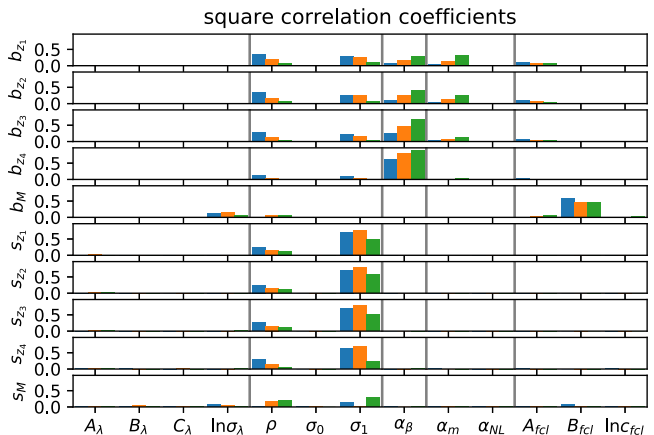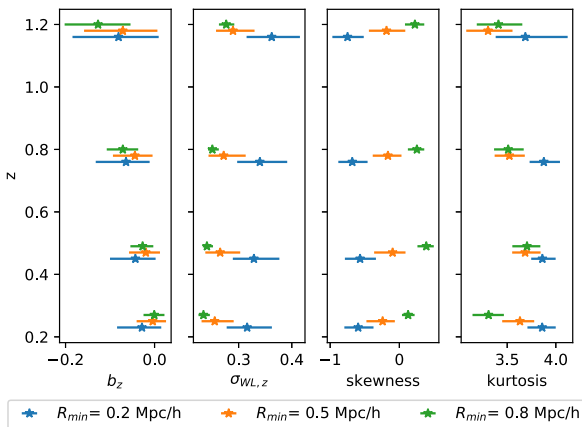
Finally, we explore the correlation between the parameters of the WL to halo mass relation and the WL systematics parameters. This allows us to gauge which systematic effects are most important in determining the mapping from WL to halo mass (cf. Section 2.3.4).

The squared correlation coefficients are shown in Fig. 9 (right), where the WL to halo mass parameters are on the vertical axis and the WL systematics parameters are organized by function into five groups on the horizontal axis. From left to right these are the parameters from the λ–mass relation, mis-centreing, photometric redshifts, shear bias, and finally cluster member contamination.

Table 1. Medians with difference to the 16th and 84th percentile of the parameters of the WL to halo mass relation obtained from several realizations of generic surveys assuming different inner fitting radii R_{min}.

R_{min}	$0.2\ h^{-1}$ Mpc	$0.5\ h^{-1}$ Mpc	$0.8\ h^{-1}$ Mpc
b_{z_1}	$-0.037^{+0.044}_{-0.053}$	$-0.010^{+0.029}_{-0.033}$	$-0.008^{+0.024}_{-0.023}$
b_{z_2}	$-0.052^{+0.048}_{-0.054}$	$-0.024^{+0.032}_{-0.037}$	$-0.035^{+0.025}_{-0.025}$
b_{z_3}	$-0.074^{+0.056}_{-0.062}$	$-0.051^{+0.040}_{-0.047}$	$-0.079^{+0.036}_{-0.034}$
b_{z_4}	$-0.096^{+0.087}_{-0.100}$	$-0.079^{+0.074}_{-0.081}$	$-0.132^{+0.070}_{-0.077}$
b_M	$1.002^{+0.024}_{-0.024}$	$1.003^{+0.021}_{-0.021}$	$1.029^{+0.016}_{-0.016}$
s_{z_1}	$-2.298^{+0.279}_{-0.277}$	$-2.758^{+0.248}_{-0.206}$	$-2.901^{+0.084}_{-0.069}$
s_{z_2}	$-2.224^{+0.277}_{-0.269}$	$-2.682^{+0.240}_{-0.222}$	$-2.847^{+0.083}_{-0.059}$
s_{z_3}	$-2.156^{+0.272}_{-0.291}$	$-2.637^{+0.253}_{-0.241}$	$-2.776^{+0.080}_{-0.061}$
s_{z_4}	$-2.038^{+0.284}_{-0.282}$	$-2.531^{+0.258}_{-0.247}$	$-2.588^{+0.084}_{-0.087}$
s_M	$-0.005^{+0.057}_{-0.061}$	$0.110^{+0.065}_{-0.068}$	$-0.253^{+0.069}_{-0.076}$
$\gamma_{skw,\,1}$	$-0.578^{+0.217}_{-0.229}$	$-0.215^{+0.166}_{-0.231}$	$0.088^{+0.108}_{-0.129}$
$\gamma_{skw,\,2}$	$-0.582^{+0.239}_{-0.229}$	$-0.069^{+0.205}_{-0.252}$	$0.402^{+0.104}_{-0.121}$
$\gamma_{skw,\,3}$	$-0.661^{+0.236}_{-0.215}$	$-0.109^{+0.232}_{-0.231}$	$0.303^{+0.091}_{-0.123}$
$\gamma_{skw,\,4}$	$-0.783^{+0.230}_{-0.213}$	$-0.156^{+0.236}_{-0.247}$	$0.230^{+0.097}_{-0.107}$
$\gamma_{krt,\,1}$	$3.888^{+0.203}_{-0.159}$	$3.616^{+0.332}_{-0.225}$	$3.426^{+0.220}_{-0.190}$
$\gamma_{krt,\,2}$	$3.918^{+0.149}_{-0.119}$	$3.729^{+0.169}_{-0.148}$	$3.849^{+0.168}_{-0.167}$
$\gamma_{krt,\,3}$	$3.771^{+0.216}_{-0.152}$	$3.489^{+0.148}_{-0.150}$	$3.609^{+0.157}_{-0.139}$
$\gamma_{krt,\,4}$	$3.750^{+0.370}_{-0.269}$	$3.280^{+0.223}_{-0.207}$	$3.405^{+0.179}_{-0.166}$

By inspecting the first four rows, we can see that at larger redshifts the biases b_z are impacted by the photometric redshift uncertainty (α_β column) the most. We can also see that the amplitude of the cluster member contamination (A_{fcl} column), the multiplicative shear bias (α_m column), and the weight and amplitude of the larger mis-centring distribution component (ρ and σ_1 columns) correlate with the biases (and thus contribute to their systematic uncertainty). We can also see that the contribution from mis-centreing and cluster member contamination diminishes for larger inner fitting radii, while the impact of multiplicative shear bias increases.

Noticeably, we find no correlation with the parameters governing the richness–mass relation. Remember that we introduced these parameters because some parts of our shear model depend on the measured richness. The accuracy of our WL mass calibration is thus unaffected by changes in the parameters of the richness mass relation. It is also worth pointing out that the uncertainty on the non-linear shear response, albeit very large, plays no role in the overall systematics budget. Other cluster WL systematics have larger impacts. This is reassuring, as the uncertainty on the non-linear shear response was chosen in accordance with the limits required by cosmic shear experiments.
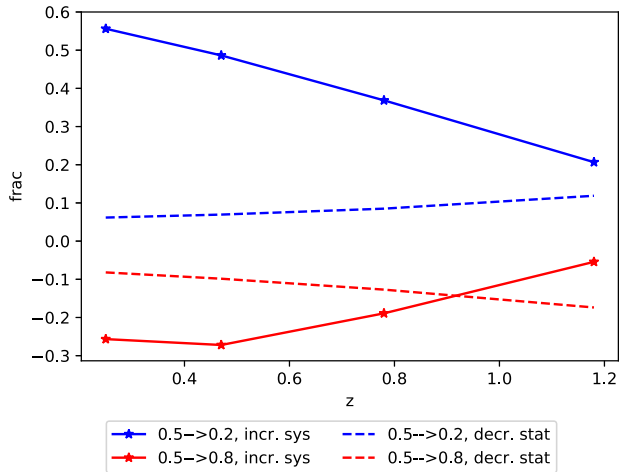
3.2 Inner fitting radius

We presented above the impact of the inner fitting radius choice on the calibration of the WL to halo mass relation. We found that using a larger inner fitting radius reduces the systematic uncertainty on the WL to halo mass bias parameters. In the context of a cosmological study, this means that in an experiment similar to the generic Stage III case adopted here, the systematic floor on the WL mass calibration is reduced by avoiding the inner core region of the cluster.

Using a larger fitting radius will, however, result in the exclusion of some shear data, reducing the precision of the measurement. This effect scales more quickly than the mere change in number of background galaxies, because the shear signal is stronger in the inner regions. Fig. 5 shows the shear signal-to-noise (shear signal scaled by square root of area on the sky) profiles for clusters at a range of redshifts and masses. In the range between 0.2-0.8 h^{-1} Mpc the signal-to-noise is quite constant with radius (which means that the shear signal falls off approximately linearly with radius). This suggests that the loss of signal from excluding the central core is not as severe as one might first think.

In Fig. 10 we contrast the increase in systematic uncertainty (ratio between the uncertainties of the biases) with the decrease of the statistical signal-to-noise (ratio of the total signal-to-noises in the respective fitting ranges). We can see that the change in signal-to-noise is 10–20 per cent when going from 0.5 to 0.8 h^{-1} Mpc or from 0.5 to 0.2 h^{-1} Mpc. The change is slightly larger at higher redshifts

Figure 9. Left-hand panel: WL bias, WL intrinsic scatter, skewness, and kurtosis (0 and 3, respectively, for a perfectly lognormal model) of the WL mass–halo mass distribution. Right-hand panel: Squared correlation coefficients between the WL–halo mass parameters and the WL and cluster catalogue systematics parameters (i.e. the paramaters of the observable mass scaling relation (A_λ, B_λ, C_λ, $\ln\sigma$), of the mis-centreing distribution (ρ, σ_0, σ_1), the modulation of the photometric redshift uncertainties α_β, linear and non-linear shape measurement uncertainties (α_γ, α_{NL}), and the cluster member contamination (A_{fcl}, B_{fcl}, c_{fcl}). These correlations can be interpreted as the impact of a given systematic on WL–halo mass parameters. The colours stand for different choices of fitting range. Clearly, the mis-centreing and cluster member contamination contribute more to the systematics budget if smaller radial scales are considered. Furthermore, the photo-z bias has as a stronger impact at high redshift, where it is less well constrained.

Figure 10. Fractional increase in the systematic uncertainty on the bias parameters (solid line) when changing the inner fitting radius, compared to fractional change in statistical power (dashed line). Changing the inner fitting radius from $R_{min} = 0.5\,h^{-1}$ Mpc to $R_{min} = 0.2\,h^{-1}$ Mpc results in a up to 50 per cent increase in the systematics floor but only a ~10 per cent gain in statistical precision. The optimal choice depends on the relative scale of the statistical and systematic errors, which in turn depends on the source galaxy and cluster number and redshift distributions.

because our outer fitting radius gets smaller with larger redshift (see discussion in next section). Note that this estimate is independent of the number of clusters and the background source density (which would be the same no matter the choice of inner fitting radius).

This modest change in statistical constraining power needs to be contrasted with the larger change in systematic uncertainty. Going further in ($R_{min} = 0.2\,h^{-1}$ Mpc instead of $0.5\,h^{-1}$ Mpc) increases the systematic uncertainty by 20–50 per cent. Staying further out ($R_{min} = 0.8\,h^{-1}$ Mpc instead of $0.5\,h^{-1}$ Mpc) reduces the systematic uncertainty by 10–30 per cent. Determining which fitting range is optimal depends also on the relative scale of the systematic and statistical noise, which for a given set of WL systematics inputs depends on the statistical constraining power of the WL calibration. This follows from the number of source galaxies and clusters and their respective redshift distributions. However, these estimates show that choosing a smaller fitting radius provides a benefit in constraining power that is fractionally smaller than its fractional cost in systematic accuracy. On the other hand, choosing a larger inner fitting radius roughly offsets gains and losses in the fractional precision and accuracy.

3.3 Outer fitting radius

The choice of the outer fitting radius is motivated by the attempt to avoid the so-called 'two-halo' term. While the halo matter profile is the dominant source of shear on small scales, at larger scales the shear signal around clusters is dominated by surrounding haloes. Indeed, haloes are biased tracers of the matter distribution, and – to first approximation – the two-halo term can be computed from the matter power spectrum and the halo bias. Because the halo bias is mass dependent, a bias measurement provides mass information. But the mass dependence of the bias is weak (e.g. Mo & White 1996).

Moreover, while these larger scales hold much potential for precise measurements of the shear signal, the modelling of this shear signal is more complicated (no longer an NFW profile) and cosmologically dependent. Furthermore, the neighbouring halo is on average a low

mass halo, and these are more strongly impacted by hydrodynamical effects. The mass distribution in the regime outside massive haloes, but on much smaller scales than accurately predicted by linear theory, is an important source of systematic uncertainty for a variety of cluster cosmological experiments.

To study the transition radius at which the cluster shear profile is no longer dominated by the central halo matter profile, we inspect the logarithmic slope of the density contrast $\Delta\Sigma$ stacked in mass bins for the different simulation snapshots in Fig. 11. We also plot a black vertical line at $R_{max}(z) = 3.2/(1+z)\,h^{-1}$ Mpc. For all considered redshifts and masses, this line falls approximately at the minimum of the logarithmic slope. We identify this transition as the transition from the 'one-halo'-regime (where the central halo matter profile dominates the shear signal), to the 'two-halo' term (which leads to a flattening of the shear profile).

The redshift evolution of the transition radius is somewhat surprising. A distance that scales with redshift like $(1 + z)^{-1}$ corresponds to a fixed physical distance between two points that does not scale with the universal expansion. This scaling might, however, well be coincidental. Halos are gravitationally bound objects, detached from the background expansion. The radius enclosing a given overdensity of matter grows with redshift as the density is diluted. Simultaneously, the amplitude of fluctuations in the Universe grows, increasing the two-halo term. These two effects counteract each other. Further study of this interplay is surely warranted, but for our current purpose we can establish that limiting our fitting range to $R < 3.2/(1+z)\,h^{-1}$ Mpc excludes radial ranges not dominated by the central halo matter density profile. This avoids the complex cosmological dependencies and the modelling uncertainties of the two-halo regime.

3.4 Additional systematic uncertainties

While the numbers above provide the systematics budget for the most prominent systematic effects in a quantitative manner, some known systematics are left out because they are harder to quantify. These include the systematic uncertainties in the modelling of hydrodynamical effects in simulations and the impact of the – possible – correlation between the mis-centreing and the dynamical state of the cluster. Therefore, we discuss our estimates for these systematic effects in the Appendices A and B and provide a summary of our these estimates in Tables 2 and 3.

At low redshift, the systematic uncertainty from the treatment of hydrodynamical effects contributes to the error budget of the bias and its mass trend at a level that is comparable to other systematics. At high redshift, it is significantly smaller than the impact of photometric redshift biases (for the generic Stage III survey adopted here). Overall, we estimate that hydrodynamical effects contribute to a ~ 2 per cent systematic uncertainty floor on the WL mass bias. The uncertainty on the variance – and its mass trend – are clearly dominated by the modelling of hydrodynamical effects, independently on the choice of the inner fitting radius. This results in a systematic floor for the intrinsic scatter of WL mass estimates of about ~ 10 per cent. These estimates are quite crude as they rely on the comparison between two hydrodynamical simulations (Magneticum and Illustris TNG). Further exploration of this uncertainty is clearly needed, especially in light of the goal of achieving 1 per cent WL mass accuracy in Stage IV surveys. These efforts would likely focus on (1) enlarging the number of hydrodynamical simulations to better probe the space of possible hydrodynamical effects, and (2) aiming to develop quantitative tools
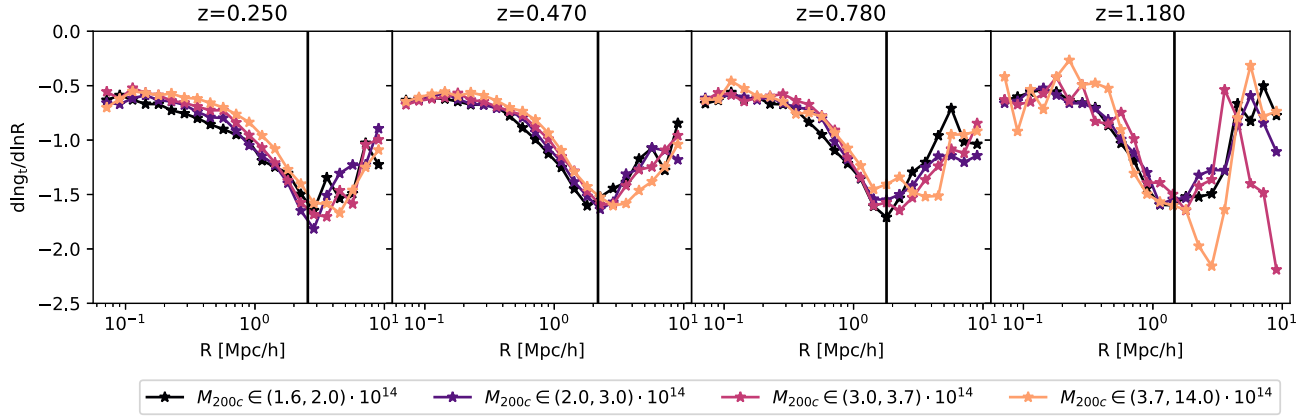
Figure 11. Logarithmic slope of the density contrast profiles stacked in mass bins for different redshifts. The vertical line indicates the upper limit of our fitting range. It clearly limits the mass fit to the range dominated by the one-halo term and avoids fitting the two-halo term, which can be recognized by the increase of the logarithmic slope, i.e. a flattening of the profile. In order to avoid the two-halo regime, with its complex dependence of cosmology and its yet unknown sensitivity to hydrodynamical modelling effect, we propose to measure the halo mass within $3.2 (1 + z)^{-1} h^{-1}$ Mpc.

Table 2. Estimates of the systematic uncertainty resulting from hydrodynamical modelling for different inner fitting radii $R_{\min}$. These values are extracted from the comparison of a small number of independent simulations and should therefore be considered order of magnitude estimates. For the bias Δb and the variance Δs, we use the maximum deviation found over different redshifts. The impact of hydrodynamical modelling uncertainties on the parameters of the WL to halo mass mapping cannot be ignored.

$R_{\min}$	$0.2\,h^{-1}$ Mpc	$0.5\,h^{-1}$ Mpc	$0.8\,h^{-1}$ Mpc
Δb	0.06	0.02	0.03
Δb_{M}	0.016	0.018	0.017
Δs	0.15	0.25	0.2
Δs_{M}	0.26	0.59	0.60

Table 3. Estimates of the systematic uncertainty resulting from the correlation between the cluster dynamical state and the strength of the mis-centreing for different inner fitting radii $R_{\min}$. These values are extracted from the comparison of two limiting cases of very strong and very weak correlations, and should therefore be considered upper limits. For the bias Δb and the variance Δs, we use the maximum deviation found over different redshifts. Using larger inner fitting radii ensures that these uncertainties are smaller than other known uncertainties.

$R_{\min}$	$0.2\,h^{-1}$ Mpc	$0.5\,h^{-1}$ Mpc	$0.8\,h^{-1}$ Mpc
Δb	0.02	0.008	0.002
Δb_{M}	0.04	0.02	0.01
Δs	0.08	0.06	0.04
Δs_{M}	0.25	0.24	0.15

to assess the agreement between simulation predictions and observed data.

In contrast to the uncertainties from hydrodynamical modelling, the uncertainties stemming from the unknown degree of correlation between mis-centreing and cluster dynamical state scales strongly with the inner fitting radius. This is to be expected, because mis-centreing impacts the inner cluster regions much more. For the generic Stage III survey adopted here, these uncertainties can be reduced to a level that is smaller than the impact of other systematics by choosing a large inner fitting radius.

Among the systematic effects we did not discuss in this work, we find the impact of intra-cluster light on the photometric redshift

estimation. Gruen et al. (2019) show that this effect is not relevant for current wide photometric surveys. We also did not consider the impact of blending on the shear measurement accuracy in crowded regions (Hernández-Martín et al. 2020; Sheldon et al. 2020). Both of these effects can be safely expected to be stronger in the cluster centre, providing additional reasons to select a larger inner fitting radius.

3.5 Comparison of modeling approaches

The systematic biases and associated uncertainties sourced by the impact of hydrodynamical effects on different quantities predicted by cosmological simulations have received increased attention over the last years. Studies have been carried out on the impact of hydrodynamical effects on the matter power spectrum (e.g. White 2004; Jing et al. 2006; Guillet, Teyssier & Colombi 2010; Fedeli et al. 2014; Mead et al. 2015; Schneider & Teyssier 2015; Springel et al. 2018; Aricò et al. 2021), the halo mass function (e.g. Cui et al. 2012; Martizzi et al. 2014; Velliscig et al. 2014; Bocquet et al. 2016; Beltz-Mohrmann & Berlind 2021; Castro et al. 2021), and the halo lensing signal (e.g. Lee et al. 2018; Debackere et al. 2021). The latter two – halo mass function and halo WL signal – are of greatest interest to WL-calibrated cluster number counts experiments, especially when the WL analysis is restricted to the inner central halo-dominated region (cf. Section 3.3).

Concerning the impact of hydrodynamical effects on the halo mass function, both Castro et al. (2021) and Beltz-Mohrmann & Berlind (2021) constrain the relation between the halo masses in paired gravity-only and hydrodynamical simulations run with the same initial conditions. Simply adopting this approach for WL calibrated cluster cosmology would require one to calibrate the mapping between the halo mass in hydrodynamical simulation and the matter density profiles that are themselves affected by the same hydrodynamical effects. This likely leads to correlations between the corrections applied to the halo mass function and the WL to halo mass mapping. This problem can be circumvented, and the overall set-up of the analysis simplified, if – as we propose – the mapping between the gravity-only halo mass and the hydrodynamic-impacted WL shear signal is calibrated. While facilitating the cosmological inference from cluster number counts, this approach has the disadvantage of reducing the astrophysical information contained in inferred scaling

relations, as they now relate to the abstract gravity-only mass, and not to the physically motivated halo mass, impacted by hydrodynamical effects.

Several recent works have suggested to absorb the impact of hydrodynamical effects on the WL shear signal by fitting for the concentration alongside the WL mass (Lee et al. 2018; Debackere et al. 2021). While this approach indeed produces less biased WL masses, it opens new challenges to the WL calibrated number counts experiments. First, it requires one to marginalize over the concentration in the WL calibration likelihood. In our tests this significantly slows down the likelihood computation, making the analysis of O(100) clusters already computationally challenging. Furthermore, the signal-to-noise ratio on the concentration measurement for individual clusters is low, and the concentration is degenerate with the mass. In addition, we find that using flat priors on the concentration is not an uninformative choice, and leads to mass biases by artificially preferring a particular range of concentrations. On the other side, imposing informative priors requires accurate mass concentration relations. The accuracy of the mass–concentration simulations is likely going to be limited by the accuracy of hydrodynamical simulations, thus likely delivering no increase in the overall accuracy compared to the calibration of the WL to halo mass relation using a fixed concentration, as done here.

Considering this together with the increased computational cost of allowing concentration to float, it is clear that there are benefits of extracting WL masses with a model that has a fixed concentration or a fixed concentration–mass relation (see, e.g. Mantz et al. 2014; Schrabback et al. 2018; Dietrich et al. 2019; Stern et al. 2019). Notably, Debackere et al. (2021) find that the bias between the WL mass and the gravity-only mass does not change significantly when fixing the concentration to the correct concentration–mass relation (Fig. 6, central panel), further justifying our approach.

Generally, introducing additional observables (like fitting for the concentration, including the brightest central galaxy (BCG) orientation to estimate the cluster orientation (Herbonnet et al. 2019), or using BCG–X-ray offsets to estimate the amount of mis-centreing) with the aim of increasing the precision (i.e. reducing the WL scatter σ_{WL}) has the risk of resulting in a loss of accuracy. The link between the additional observable and the underlying profile property would generally introduce additional systematic uncertainties. Hydrodynamical modelling uncertainties only slightly perturb the 2D projected matter density profile for massive haloes, because the matter distribution is dominated by dark matter and it is projected along the line of sight. This reduces its sensitivity to baryon redistribution effects. Other cluster observables, like the BCG position, the X-ray centre, or cluster concentration, depend more strongly on the hydrodynamical modeling. Thus, they are inherently more challenging to calibrate with simulations. In the limit of an analysis that is systematics-dominated (large cluster samples, dense shear galaxy samples) it is crucial only to control the systematic uncertainties. On the other hand, in the limit of a small sample or a single cluster analysis or a large cluster sample with excellent control of systematic uncertainties, a higher fidelity model profile that employs additional information to reduce the scatter in the WL to halo mass relation could be advantageous.

4 CONCLUSIONS

In this work we describe a method and provide a proof-of-concept application to quantitatively determine the bias and accuracy of a

WL mass calibration of a cluster sample. The concept of the WL was previously used by Becker & Kravtsov (2011) to quantify the cluster-to-cluster variation w.r.t. to an NFW profile (see also Oguri & Hamana 2011; Bahé et al. 2012; Lee et al. 2018). Then, others followed (Applegate et al. 2014; Schrabback et al. 2018; Dietrich et al. 2019; Stern et al. 2019) who sought to also account for the other systematic uncertainties in the WL measurement and folded these into the mapping between the WL and halo mass. In this work, we systematize this concept, by explicitly marginalizing over known WL systematics and using hydrodynamical simulations to simulate realistic shear profiles. By careful modelling choices, we ensure that the computational cost of the cosmological analysis using our framework is sufficiently small to allow us to analyse large cluster samples O(10 000) with individual cluster WL measurements.

Besides providing a clear procedure for deriving the uncertainties on the WL to halo mass, we make a conceptual improvement in the accounting of hydrodynamical effects in cluster number counts cosmology. Prior to this work, the impact of hydrodynamical effects on the abundance of haloes and the impact of hydrodynamical effects of the matter profiles have been discussed separately. Here we show how using suites of simulations run with and without hydrodynamical effects from identical initial conditions allows us to relate the gravity-only halo mass to the WL mass extracted in the presence of hydrodynamical effects. Using the calibration of this mapping, the cosmological dependence of the halo mass function extracted from large suites of gravity-only simulations can be exploited in a self-consistent manner that does not ignore hydrodynamical effects.

Our method starts from a library of 2D projected matter density profiles extracted from hydrodynamical simulations and the halo masses from the matched gravity-only runs. The projected matter density profiles are extracted for different mis-centred positions. We then specify a set of cluster properties and WL survey properties (such as an observable–mass relation, a multiplicative shear bias, an uncertainty on the lensing efficiency due to photometric redshift uncertainties, etc.) and transform the surface matter density profiles into shear profiles. A one-parameter model of the shear profile, inspired by a mis-centred NFW-profile, is then fit to the simulated shear profiles. The resulting best-fitting mass is called the WL mass. We fit the relation between WL and halo mass accounting for the mis-centreing by appropriately weighing each mis-centred profile by the mis-centreing distribution. This results in a set of WL to halo mass parameters.

We then repeat the extraction of the WL to halo mass parameters ∼1000 times with varying realizations of cluster and WL survey systematics. This results in ∼1000 realizations of shear profile libraries sampling the entire space of WL systematics, and ∼1000 realizations of WL to halo mass parameters. The variance in this sample of WL to halo mass parameters encapsulates the systematic uncertainty on the WL mass measurement and can be used as a prior in cosmological analyses. Exploring the correlations between the WL to halo mass parameters and the cluster and WL systematics allows us to assess which systematic effects have the largest impact on the accuracy of the WL mass measurement. For this exercise we adopt a Stage III-like generic cluster survey modeled on the combination of SPT and DES.

We perform the calibration of the WL to halo mass relation for different inner fitting radii. We generally find that smaller inner fitting radii, including scales closer to the cluster centre, lead to larger systematic uncertainties in the WL mass measurement. This loss in accuracy is fractionally larger than the gain in precision due to the inclusion of more stronger sheared background galaxies. Selection of an optimal inner radius depends on the details of a cluster WL

experiment, but in the limit of analyses of large cluster samples that are systematics-dominated, there are clear advantages in avoiding the cluster core region.

Inspecting the logarithmic slope of the shear profile, we determine the outer fitting radius $R_{max} = 3.2\,(1 + z)^{-1}\,h^{-1}$ Mpc which excludes the signal from correlated neighbouring structures. That signal – the so-called two-halo term – has complex dependencies on cosmological parameters and hydrodynamical effects, and for that reason we have excluded it from our analysis.

Finally, we also provide estimates for the uncertainty stemming from the hydrodynamical modelling and from the possible correlations between mis-centreing and cluster dynamical state. The former leads to a ~ 2 per cent systematics floor, comparable to other systematics for Stage III surveys, but clearly above the 1 per cent goal for Stage IV surveys. Further reducing this uncertainty would require that a wide range of hydrodynamical models fully sampling the plausible range of sub-grid physics be carried out and then compared to observational data to assess the likelihood of each hydrodynamic model (e.g. Chisari et al. 2019; Aricò et al. 2021). That ensemble of simulations and the associated likelihoods could then be used to reduce the hydrodynamics related systematics floor.

In summary, this work discusses a new framework for deriving the systematic uncertainty on the WL mass measurements for cluster surveys. The proposed method – shown in the context of a generic Stage III cluster and WL survey – accounts for the impact of hydrodynamical effects, different data systematics, and marginalizes over mis-centreing effects, while ensuring an efficient mass extraction that is well suited for large cluster surveys. Our method is being applied to the WL mass calibration of *South Pole Telescope* selected clusters with Dark Energy Survey WL data (Bocquet et al., in preparation), and the WL mass calibration of clusters selected in the eROSITA full survey depth equatorial field with Hyper-Supreme Cam WL data (Chiu et al. 2021). Our framework can naturally be extended to incorporate the properties of upcoming Stage-IV lensing surveys such as those by Vera C. Rubin Observatory (Ivezic et al. 2008) and Euclid (Laureijs et al. 2011).

DATA AVAILABILITY

The data underlying this article will be shared upon request to the corresponding author.

REFERENCES

Abbott T. M. C. et al., 2020, Phys. Rev. D, 102, 023509
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Applegate D. E. et al., 2014, MNRAS, 439, 48
Aricò G., Angulo R. E., Hernández-Monteagudo C., Contreras S., Zennaro M., 2021, MNRAS, 503, 3596
Bahé Y. M., McCarthy I. G., King L. J., 2012, MNRAS, 421, 1073
Bartelmann M., 1996, A&A, 313, 697
Beck A. M. et al., 2016, MNRAS, 455, 2110
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Beltz-Mohrmann G. D., Berlind A. A., 2021, preprint (arXiv:2103.05076)
Benson B. A. et al., 2013, ApJ, 763, 147
Bleem L. E. et al., 2015, ApJS, 216, 27
Bleem L. E. et al., 2020, ApJS, 247, 25
Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361
Bocquet S. et al., 2019, ApJ, 878, 55
Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5
Castro T., Borgani S., Dolag K., Marra V., Quartin M., Saro A., Sefusatti E., 2021, MNRAS, 500, 2316
Child H. L., Habib S., Heitmann K., Frontiere N., Finkel H., Pope A., Morozov V., 2018, ApJ, 859, 55
Chisari N. E. et al., 2019, Open J. Astrophys., 2, 4
Chiu I.-N. et al., 2021, preprint (arXiv:2107.05652)
Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, MNRAS, 423, 2279
Debackere S. N. B., Schaye J., Hoekstra H., 2021, MNRAS, 505, 593
Dietrich J. P. et al., 2019, MNRAS, 483, 2871
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Dolag K., Mevius E., Remus R.-S., 2017, Galaxies, 5, 35
Drlica-Wagner A. et al., 2018, ApJS, 235, 33
Fedeli C., Semboloni E., Velliscig M., Van Daalen M., Schaye J., Hoekstra H., 2014, J. Cosmol. Astropart. Phys., 2014, 028
Grandis S. et al., 2020, MNRAS, 498, 771
Gruen D. et al., 2019, MNRAS, 488, 4389
Guillet T., Teyssier R., Colombi S., 2010, MNRAS, 405, 525
Haiman Z., Mohr J. J., Holder G. P., 2001, ApJ, 553, 545
Hennig C. et al., 2017, MNRAS, 467, 4015
Herbonnet R., von der Linden A., Allen S. W., Mantz A. B., Modumudi P., Morris R. G., Kelly P. L., 2019, MNRAS, 490, 4889
Hernández-Martín B. et al., 2020, A&A, 640, A117
Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304
Hoekstra H., 2003, MNRAS, 339, 1155
Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, MNRAS, 427, 1298
Hoyle B. et al., 2018, MNRAS, 478, 592
Ivezic Z. et al., 2008, ApJ, 873, 44
Jing Y. P., Zhang P., Lin W. P., Gao L., Springel V., 2006, ApJ, 640, L119
Kaiser N., Squires G., 1993, ApJ, 404, 441
Komatsu E. et al., 2011, ApJS, 192, 18
Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lee B. E., Le Brun A. M. C., Haq M. E., Deering N. J., King L. J., Applegate D., McCarthy I. G., 2018, MNRAS, 479, 890
Lin Y., Mohr J. J., Stanford S. A., 2004, ApJ, 610, 745
McClintock T. et al., 2019a, MNRAS, 482, 1352
McClintock T. et al., 2019b, ApJ, 872, 53
Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, MNRAS, 440, 2077
Mantz A. B. et al., 2015, MNRAS, 446, 2205
Marinacci F. et al., 2018, MNRAS, 480, 5113
Martizzi D., Mohammed I., Teyssier R., Moore B., 2014, MNRAS, 440, 2290
Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Murata R. et al., 2019, PASJ, 71, 107
Naiman J. P. et al., 2018, MNRAS, 477, 1206

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Nelson D. et al., 2018, MNRAS, 475, 624
Nelson D. et al., 2019, Comput. Astrophys. Cosmol., 6, 2
Nishimichi T. et al., 2019, ApJ, 884, 29
Oguri M., Hamana T., 2011, MNRAS, 414, 1851
Pillepich A. et al., 2018, MNRAS, 475, 648
Planck Collaboration, 2016, A&A, 594, A13
Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe
    N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25
Saro A. et al., 2015, MNRAS, 454, 2305
Schneider A., Teyssier R., 2015, J. Cosmology Astropart. Phys., 2015, 049
Schrabback T. et al., 2018, MNRAS, 474, 2635
Schrabback T. et al., 2021, MNRAS, 505, 3923
Sheldon E. S., Huff E. M., 2017, ApJ, 841, 24
Sheldon E. S., Becker M. R., MacCrann N., Jarvis M., 2020, ApJ, 902, 138
Smail I., Ellis R. S., Fitchett M. J., 1994, MNRAS, 270, 245
Sommer M. W., Schrabback T., Applegate D. E., Hilbert S., Ansarinejad B.,
    Floyd B., Grandis S., 2021, preprint (arXiv:2105.08027)
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328,
    726
Springel V. et al., 2018, MNRAS, 475, 676
Stern C. et al., 2019, MNRAS, 485, 69
Teklu A. F., Remus R.-S., Dolag K., Beck A. M., Burkert A., Schmidt A. S.,
   Schulze F., Steinborn L. K., 2015, ApJ, 812, 29
Umetsu K., 2020, A&AR, 28, 7

Varga T. N. et al., 2019, MNRAS, 489, 2511
Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le
    Brun A. M. C., Dalla Vecchia C., 2014, MNRAS, 442, 2641
von der Linden A. et al., 2014, MNRAS, 443, 1973
Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G.,
    Rozo E., 2013, Phys. Rep., 530, 87
White M., 2004, Astropart. Phys., 22, 211
Zhang Y.-Y. et al., 2017, A&A, 599, A138

APPENDIX A: SYSTEMATIC BUDGET FOR HYDRODYNAMICAL MODELLING

Each simulation we use has a specific set of assumptions for the hydrodynamical modelling. As such, each simulation is a point estimate of the hydrodynamical modelling uncertainties. Despite ongoing efforts, it is to date very costly to sample the entire range of possible hydrodynamical modelling systematics (several thousand simulations would have to be run). In this work, we bracket the remaining uncertainty by determining the difference in WL to halo mass scaling relation parameters between three different simulations: Magneticum with hydrodynamical effects, Magneticum gravity-only, and Illustris TNG with hydrodynamical effects.
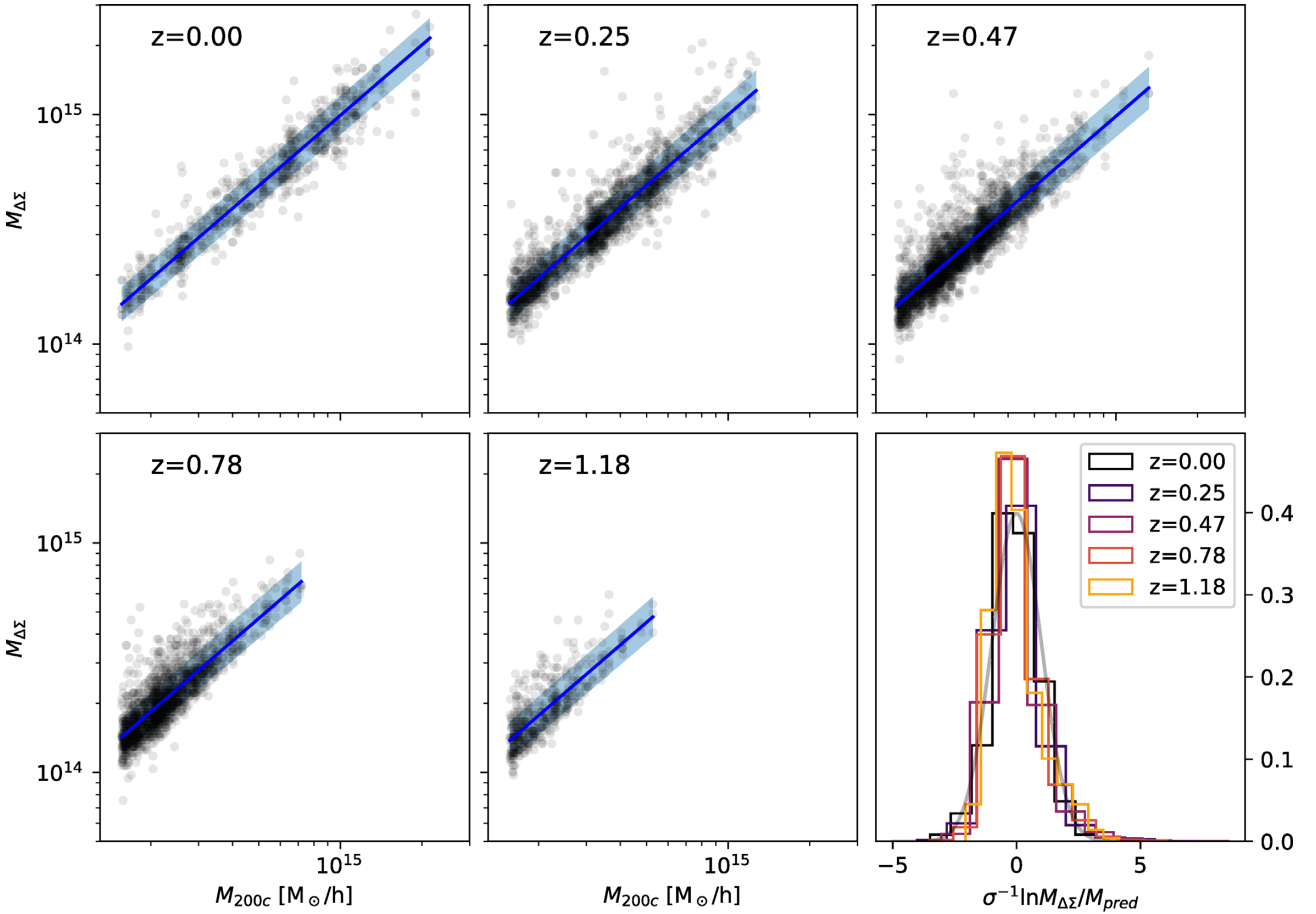
Figure A1. First five panels: Mass $M_{\Delta\Sigma}$ (extracted from fitting an NFW profile to the perfectly centred density contrast profiles) plotted against the halo mass M_{200c} for haloes extracted from different snapshots (different panel) in the Magneticum hydrodynamical simulation, together with a lognormal fit to the distribution as a blue line (intrinsic scatter in shaded blue). Lower left-hand panel: Residuals of the fits in the different snapshots, normalized by the intrinsic scatter. In grey a standard normal distribution. Some positive skewness in the mass distribution is noticeable. It might be related to line-of-sight project effects. Comparing the $M_{\Delta\Sigma}$– halo mass relation from different simulation provides an estimate on the residual uncertainty stemming from the treatment of hydrodynamical effects.
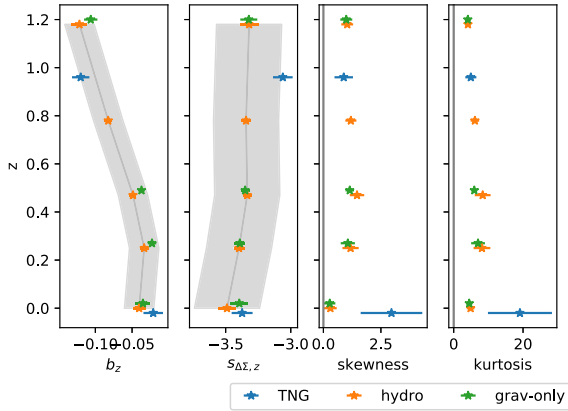
Figure A2. Bias between the density contrast mass and the halo mass, natural logarithm of the variance of the the density contrast mass, $s_{\Delta\Sigma} = \ln \mathrm{Var}[\ln M_{\Delta\Sigma}|M_{200c}]$, and skewness and kurtosis of the distribution of density contrast masses around the halo masses for different simulations at different redshift. The grey band indicate the maximal difference between any simulation and Magneticum hydro. We use this quantity to estimate the systematic error introduces by the modelling of the hydrodynamical effects (grey band). In the two rightmost panels, the grey lines indicate the reference value for Gaussian distributions (skewness = 0, kurtosis = 3). We find a positive skewness and large kurtosis in all simulations.

For simplicity, we specifically look at centred density contrast profiles $\Delta\Sigma(R)$ and extract the density contrast mass $M_{\Delta\Sigma}$ by fitting an NFW profile with constant concentration ($c = 3.5$) to the density contrast profile of each halo. We use a minimal fitting radius of $0.5\ h^{-1}$ Mpc, and an outer fitting radius following the prescription discussed above. The scatter plot between the halo mass and the density contrast mass in the Magneticum hydrodynamical run is shown in Fig. A1. We then fit a lognormal distribution with mass trend in the mean and the variance (as specified in equations 20, 21, and 22). This results in bias and intrinsic scatter parameters shown in the two leftmost panels of Fig. A2. On the Magneticum hydro simulations we find a mass trend of the bias of $b_M = 1.019 \pm 0.006$, and a mass trend of the variance $s_M = 0.11 \pm 0.04$.

The same analysis is repeated also for the Magneticum gravity-only simulations, and for the Illustris-TNG simulations. The resulting biases and intrinsic scatters are shown in the two leftmost panels of Fig. A2. For Magneticum gravity-only (Illustris-TNG) simulation we find a mass trend of the bias of $b_M = 1.001 \pm 0.006$ (1.037 ± 0.012), and a mass trend of the variance $s_M = 0.28 \pm 0.04$ (-0.48 ± 0.08). We then estimate the systematic uncertainty induced by different – or no – hydrodynamical treatment by taking the difference between the parameters from Illustris-TNG and the parameters from Magneticum hydro. Note that this difference is always larger than the difference between Magneticum hydro and Magneticum gravity-only. The resulting estimates are shown in Table 2 for different inner fitting radii (the same analysis as above is repeated also for $R_{\min} = 0.2,\ 0.8\ h^{-1}$ Mpc). Note that these numbers are estimated from a very limited number of simulations. As such, they are to be considered order of magnitude estimates. The three simulations we compared also have quite different universal baryon fractions: 16.8 per cent for Magneticum–hydro, 15.7 per cent for Illustris-TNG–hydro. This impacts the strength of the hydrodynamical feedback. Our estimated uncertainties include this effect, making them conservative estimates. Proper marginalization over a sufficiently wide range of hydrodynamical modelling and cosmological parameters currently eludes our possibilities, due to the large computational cost.

At low redshift, the hydrodynamical modelling introduces as error on the biases that is comparable in magnitude to the uncertainty derived from the better quantified systematics (compare Tables 1 and 2). At higher redshift, the large uncertainties introduced by the photometric redshift uncertainty surpass the error resulting from the hydrodynamical modelling. In the case of the intrinsic variance, the error budget from hydrodynamical modelling is rather constant with inner fitting radius, while the effect from the other systematics is reduced strongly when using a larger inner fitting radius. At $R_{\min} = 0.8\ h^{-1}$ Mpc, the effects of hydrodynamical modelling are the dominant source of systematic uncertainty. In the case of the mass trend of the bias, the uncertainty from hydrodynamical modelling is of the same order of magnitude as the effect of other systematics. In contrast, the uncertainty on the mass trend of the variance is dominated by the impact of hydrodynamical modelling. In summary, hydrodynamical modelling contributes strongly to the uncertainty on the variance and its mass trend, while it plays an important role in the systematics budget of the bias and its mass trend. We add these contributions in quadrature to the errors resulting from the marginalization over the other systematics.

APPENDIX B: EXTRA SYSTEMATICS RELATED TO MIS-CENTREING

Our treatment of mis-centreing rests on some rather strong assumptions. We make assumptions on the direction of the mis-centreing and on the correlation between its strength and the cluster dynamical state. We are going to estimate the impact of the latter on the uncertainty of the WL bias and scatter.

B1 Direction of the mis-centreing

We assigned the offset isotropically in the plane of the sky. Depending on the sort of mis-centreing one plans to emulate, this assumption might be wrong. If the mis-centreing is due to a physical reason (physical offset between BCG and halo centre, or physical offset between peak in the X-ray surface brightness and halo centre, etc.) it stands to reason that such an offset would correlate with the triaxial structure of the halo: it would be more pronounced along the major axis, and less pronounced along the minor axis. How strong this effect actually is, is yet to be determined empirically – and in our opinion very hard to measure.

Similarly, the mis-centreing can be due to mis-association or other forms of ambiguity. The automated codes could have selected a galaxy far from the cluster centre as the BCG, or the system could be undergoing a merger, offsetting its gas component from the stellar and dark matter components. In such cases, the matter profile around the observed centre might be very different from that of a relaxed cluster, artificially mis-centred by an equal amount. Qualitatively, such systems only make a few per cent of typical cluster samples.

Finally, a 2D isotropic mis-centreing is actually a correct description of the mis-centreing due to the measurement uncertainty induced by survey instruments with a comparatively bad angular resolution (typically X-rays or SZ surveys). In this case, noise fluctuation scatter the measured centre around the true centre in an isotropic manner on the plane of the sky.

In summary, the correct directional assignment of the mis-centreing is highly dependent on which centres one uses for the WL analysis. There remain also many uncertainties and unknowns concerning the quantitative description of the direction of the mis-centreing. In our opinion, this provides an additional reason to choose large inner fitting radii.

B2 Correlation between strength of mis-centreing and cluster dynamical state

The second assumption we made, is that the strength of the mis-centreing (as parameterized by equation 6) does not correlate with any cluster property. A quantitative description of how much cluster centres correlate with cluster properties is, however, both highly dependent on the type of centres one chooses, and also very uncertain, as – to the authors' knowledge – no empirical study in the context of large surveys has found a conclusive measurement of such correlation.

In order to bracket the possible impact of such a correlation, we construct a case in which the strength of mis-centreing and the cluster dynamical state are highly correlated. For each of the three projections of a halo, we measure the concentration by fitting a two parameter NFW model to the centred density contrast. We then rank all halo-projections based of their 2D concentration. The ρ steepest halo-projections then get assigned the narrower component of the mis-centreing distribution, while the $1 - \rho$ flattest halo-projections get assigned the larger component of the mis-centreing distribution. This way, the question whether the halo is strongly or weakly mis-centred along that projection is made to correlate perfectly with concentration of the density contrast profile in that projection.

Operationally, this requires us to alter the mis-centreing distribution used in the fitting of the WL–halo mass parameters (equation 23). In this case, it depends on the 2D concentration. The extraction of the WL mass is kept agnostic to the cluster 2D concentration, as in most large survey applications the signal-to-noise of cluster-by-cluster WL measurements is too low to measure the concentration at all (considering that in many cases, even the mass measurement is only about 1–2σ with fixed concentration).

This analysis then results in a new set of bias and variance parameters for the WL–halo mass relation. In Table 3 we report the absolute difference between the baseline results and the results where the strength of the mis-centreing is perfectly correlated with the 2D concentration of the halo-projection. We perform the analysis for different inner fitting radii. We find that using large inner fitting radii reduces the difference to a level clearly smaller than the systematic uncertainty from other sources.

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