

Cosmic variance of the galaxy cluster weak lensing signal

D. Gruen^{1,2*}, S. Seitz^{1,2}, M. R. Becker^{3,4}, O. Friedrich^{1,2} and A. Mana^{1,2}

¹University Observatory Munich, Scheinerstrasse 1, 81679 Munich, Germany

²Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany

³KIPAC, Physics Department, Stanford University, Stanford, CA 94305, USA

⁴KIPAC, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

ABSTRACT

Intrinsic variations of the projected density profiles of clusters of galaxies at fixed mass are a source of uncertainty for cluster weak lensing. We present a semi-analytical model to account for this effect, based on a combination of variations in halo concentration, ellipticity and orientation, and the presence of correlated haloes. We calibrate the parameters of our model at the 10 per cent level to match the empirical cosmic variance of cluster profiles at $M_{200m} \approx 10^{14} \dots 10^{15} h^{-1} M_{\odot}$, $z = 0.25 \dots 0.5$ in a cosmological simulation. We show that weak lensing measurements of clusters significantly underestimate mass uncertainties if intrinsic profile variations are ignored, and that our model can be used to provide correct mass likelihoods. Effects on the achievable accuracy of weak lensing cluster mass measurements are particularly strong for the most massive clusters and deep observations (with ≈ 20 per cent uncertainty from cosmic variance alone at $M_{200m} \approx 10^{15} h^{-1} M_{\odot}$ and $z = 0.25$), but significant also under typical ground-based conditions. We show that neglecting intrinsic profile variations leads to biases in the mass-observable relation constrained with weak lensing, both for intrinsic scatter and overall scale (the latter at the 15 per cent level). These biases are in excess of the statistical errors of upcoming surveys and can be avoided if the cosmic variance of cluster profiles is accounted for.

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

1 INTRODUCTION

The largest objects ever formed depend most sensitively on small changes in the overall properties of the Universe. It is for this reason that clusters of galaxies are a versatile probe of cosmology. The number density of clusters and its evolution with redshift is influenced by the expansion history, the density and level of inhomogeneity of matter in the Universe and the growth of structures by means of gravitation. Both the parameters of a standard Λ CDM model and deviations from primordial Gaussianity or General Relativity are therefore accessible to cluster cosmology (cf., e.g., Allen et al. 2004; Vikhlinin et al. 2009b; Rozo et al. 2010; Mantz et al. 2010; Mana et al. 2013; Benson et al. 2013; Planck Collaboration et al. 2013; Mantz et al. 2015 for individual analyses and Allen et al. 2011 for a recent review).

The building blocks of such analyses are (i) a cluster catalog with well-defined selection function based on, e.g. optical, X-ray or Sunyaev & Zel’dovich (1972, hereafter SZ) properties, (ii) a prediction, based on theory or sim-

ulations, of how the number density of clusters at given mass and redshift depends on the cosmological parameters in question (e.g. Press & Schechter 1974; Sheth & Tormen 1999; Tinker et al. 2008) and (iii) a mass-observable relation (MOR) that connects the observable from (i) to the mass-based prediction from (ii) in terms of a likelihood. The latter must not only describe the mean relation of mass and observable but, because of the observable limited nature of any practical survey, also the intrinsic scatter in the observable at fixed mass (e.g. Lima & Hu 2005).

With the advent of large cluster catalogs, our imperfect knowledge of the MOR remains the most important limiting factor of cluster cosmology. It is difficult to predict or simulate, to the level of accuracy required, all physical effects that influence the observables, particularly ones that are of baryonic nature. Thus, one needs to rely on an empirical calibration of the MOR.

The latter can in principle be done by means of self-calibration, i.e. by constraining both cosmology and MOR parameters from the cluster catalog alone (Hu 2003; Majumdar & Mohr 2004; Lima & Hu 2005). However, this approach greatly reduces the cosmological power of cluster

* E-mail: dgruen@usm.uni-muenchen.de (DG)

studies, especially when extensions to the most simple MOR and cosmological models are considered.

It is for this reason that weak lensing studies of clusters of galaxies are a particularly powerful complement to cluster cosmology. Being sensitive to all matter independent of its astrophysical state, lensing allows, in principle, an unbiased measurement of cluster mass. Several studies have used this approach to constrain MORs with lensing mass measurements of sets of individual clusters (e.g. Marrone et al. 2009, 2012; Hoekstra et al. 2012; von der Linden et al. 2012; Gruen et al. 2014; Mantz et al. 2015), with great prospects of further improving statistics with on-going and future large surveys.

However, for accurate constraints on the mean MOR and its intrinsic scatter to be derived, one needs to ensure that the mass likelihood from a lensing analysis includes all actual sources of uncertainty. For typical cluster lensing studies, the dominant uncertainty is observational, i.e. based on the limited number of intrinsically elliptical background galaxies. Since all structures between sources and observer cause a lensing signal, an additional, irreducible uncertainty results from the variance of the matter density along the line of sight (LOS; cf. e.g. Hoekstra 2001, 2003). Most cluster lensing studies include either only the first or both of these effects when fitting the signature of a mass dependent density profile to observed data.

What this neglects, either entirely or apart from a low-dimensional approximation in terms of e.g. halo concentration, is the complexity of cluster profiles. The projected density profiles of two clusters at the same virial mass can differ greatly, as could be described accurately only with detailed information about the shape and orientation of their central dark matter halo and the positions and masses of subhaloes and neighbouring structures. Due to its relatively low resolution, no weak lensing analysis is able to uncover all or even most of this information. However, the uncertainty of mass measurement will depend on the statistical properties of these variations.

The goal of this work is to provide a model for the variations in projected density profiles of cluster of galaxies at fixed mass. We construct our model using analytical templates for the expected scatter due to variations in halo concentration, ellipticity and orientation, and its substructure and cosmic neighbourhood. We re-scale these templates to match the empirical variations of cluster profiles in a cosmological simulation, where both well-defined true masses and the lensing signal without any observational noise are known. We then use this semi-analytical covariance model to test and predict the accuracy of lensing measurements of cluster mass and MORs, and to assess the effect of neglecting intrinsic variations.

The structure of this paper is as follows. In Section 2 we briefly describe the simulations used. Section 3 defines the components of our model for the mean cluster profile and, in particular, its intrinsic covariance. In Section 4 we explain how the model is fitted to the simulated data. Effects of the intrinsic covariance on weak lensing cluster mass measurements and MOR studies are shown in Section 5, before we summarize in Section 6.

Conventions

Our calculations are presented in a flat WMAP 7 year cosmology (Komatsu et al. 2011) with $(\Omega_m, \Omega_b, \sigma_8, h, n) = (0.27, 0.044, 0.79, 0.7, 0.95)$, in consistency with the simulations used. The covariances can be readily re-scaled to different sets of cosmological parameters, up to a potential cosmology dependence of our model parameters. We explicitly write appropriate factors of $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$ where applicable. We denote the radii of spheres around the cluster centre with fixed overdensity as $r_{\Delta m}$, where $\Delta = 200$ is the overdensity factor of the sphere with respect to the mean matter density ρ_m at the cluster redshift. The mass inside these spheres is labelled and defined correspondingly as $M_{\Delta m} = \Delta \times \frac{4\pi}{3} r_{\Delta m}^3 \rho_m$.

We use a linear matter power spectrum $P_{\text{lin}}(k, z) = D^2(z)P_{\text{lin}}(k, 0)$ with the transfer function model including baryonic effects of Eisenstein & Hu (1998) and normalize to our value of σ_8 . For the linear growth factor $D(z)$, normalized to $D(0) = 1$, we use the expression of Peacock & Dodds (1996). We define the corresponding linear matter two-point correlation as

$$\xi_{\text{lin}}(r, z) = \frac{1}{2\pi^2} \int dk k^2 P_{\text{lin}}(k, z) \frac{\sin(kr)}{kr}. \quad (1)$$

For the scaling of some of the cluster properties we will use the common definition of peak height

$$\nu = \delta_c / \sigma(M_{200m}, z), \quad (2)$$

where $\delta_c = 1.686$ and

$$\sigma^2(M_{200m}, z) = D^2(z) \times \frac{1}{2\pi^2} \int dk k^2 P_{\text{lin}}(k, 0) \left| w \left(k, \left(\frac{3M_{200m}}{4\pi\rho_{0,m}} \right)^{1/3} \right) \right|^2, \quad (3)$$

with the Fourier transform of the top-hat window function of radius r , $w(k, r)$.

2 SIMULATIONS

We use the simulation labeled L1000W in Tinker et al. (2008). It consists of 1024^3 dark matter particles of mass $6.98 \times 10^{10} h^{-1} M_\odot$ in a box of comoving size $1 h^{-1} \text{ Gpc}$, simulated with the parallelized Adaptive Refinement Tree algorithm (Kravtsov et al. 1997; Gottloeber & Klypin 2008) from redshift $z = 60$ to $z = 0$, at an effective spatial resolution of $30 h^{-1} \text{ kpc}$. In a snapshot at redshift $z = 0.24533$ almost 15,000 haloes at $0.95 \times 10^{14} h^{-1} M_\odot \leq M_{200m} \leq 1.5 \times 10^{15} h^{-1} M_\odot$ are identified.

The same haloes are also used in Becker & Kravtsov (2011) and we employ the lensing maps computed in that work. For the selected haloes, the mass is integrated along the LOS and the lensing signal is determined on a grid of approximately 40 comoving kpc/pixel using the Born approximation as described in Becker & Kravtsov (2011, Section 3). All matter within comoving $\pm 200 h^{-1} \text{ Mpc}$ along the LOS and, transversely, in a square with comoving side length $20 h^{-1} \text{ Mpc}$ centered on the cluster is included in the calculation. Background sources are assumed to be at a constant redshift $z_s = 1$, where $\Sigma_{\text{crit}} = 4.22 \times 10^{15} h M_\odot \text{ Mpc}^{-2}$. The density map inside a clustrocentric radius $\theta \leq 1'$ is subject

to resolution effects, which is why we discard it from our analysis.

3 MODEL DEFINITION

In the following, we introduce our model for the covariance of projected cluster profiles. Its components, namely the model for the mean profile (Section 3.1) and the contributions to the covariance (Section 3.2), are described subsequently.

Consider the convergence profile of a given cluster with mass $M_{200m} =: M$ as

$$\mathbf{K} = \kappa(M) + \mathbf{E}, \quad (4)$$

where \mathbf{K} is the observed convergence profile and $\kappa(M)$ is the mean convergence profile of clusters of fixed mass M . \mathbf{E} is the residual between the two, which arises due to observational uncertainty and variations of the noiseless observable κ profile at fixed mass. These variations could be due to uncorrelated structures along the LOS or the projected cluster density profile itself. The latter *intrinsic* variations are the focus of this work.

The profiles are vectors measured in a system of radial bins, where bin $1 \leq i \leq n$ is defined as $\theta \in [\theta_{i,\min}, \theta_{i,\max}]$ without gaps, i.e. $\theta_{i,\max} = \theta_{i+1,\min}$.

Approximating the residuals \mathbf{E} as a multivariate Gaussian with zero mean, one can write the likelihood for an observed \mathbf{K} given a mass M as

$$\begin{aligned} \mathcal{L}(\mathbf{K}|M) &= \frac{1}{\sqrt{(2\pi)^n \det(C(M))}} \\ &\times \exp\left(-\frac{1}{2}\mathbf{E}^T C^{-1}(M)\mathbf{E}\right), \end{aligned} \quad (5)$$

where we have introduced the covariance C of residuals \mathbf{E} . This $n \times n$ matrix is defined as

$$C_{ij} = \text{Cov}(E_i, E_j). \quad (6)$$

We make a semi-analytic ansatz for a parametric model of C as

$$\begin{aligned} C(M) &= C^{\text{obs}} + C^{\text{LSS}} + c^{\text{conc}}(\nu)C^{\text{conc}}(M) \\ &+ c^{\text{corr}}(\nu)C^{\text{corr}}(M) + c^{\text{ell}}(\nu)C^{\text{ell}}(M) \end{aligned} \quad (7)$$

with contributions from observational uncertainty (i.e., shape noise in shear measurement and Poisson noise in magnification studies) C^{obs} , uncorrelated large scale structure along the LOS C^{LSS} , scatter in halo concentration C^{conc} , correlated secondary haloes near the cluster C^{corr} and variations in halo ellipticity and orientation C^{ell} . The terms C^{obs} and C^{LSS} are independent of the cluster itself. In contrast to the common case where only these are considered, however, $C(M)$ inherits a mass dependence from the latter terms, which we will call *intrinsic* variations of the cluster profile. They are described in detail in Section 3.2. The empirical rescaling factors $c^*(\nu)$ of C^{corr} , C^{ell} and C^{conc} are determined by fitting the covariance model to the simulated haloes (see Section 4).

3.1 Mean profile

Several studies in the past have proposed functional forms for the three-dimensional density of dark matter haloes.

Navarro, Frenk, & White (1997, hereafter NFW) found that haloes are well-fit by the two-parametric broken power-law profile

$$\rho_{\text{NFW}}(r) = \frac{\rho_0}{(r/r_s)(1+r/r_s)^2}, \quad (8)$$

with scale density ρ_0 and a scale radius r_s that can be expressed as a fraction of r_{200m} using the concentration $c_{200m} = r_{200m}/r_s$.

Two additional effects need to be accounted for in a realistic description of halo density profiles. Firstly, the enclosed mass of an NFW profile diverges logarithmically as $r \rightarrow \infty$. The density must therefore be truncated at large radii, for which a number of approaches have been proposed (e.g. Takada & Jain 2003; Hayashi & White 2008; Baltz, Marshall, & Oguri 2009, hereafter BMO). Secondly, the halo is embedded in an overdensity of correlated matter that contributes to the overall projected profile at all, and dominantly at large, radii. We closely follow the work of Oguri & Hamana (2011), who fit a superposition of a BMO profile with a linear two-halo term to simulated haloes, i.e.

$$\Sigma(M) = \kappa(M)\Sigma_{\text{crit}} = \Sigma_{1\text{h,BMO}}(M) + \Sigma_{2\text{h,lin}}(M). \quad (9)$$

We describe the one-halo (Section 3.1.1) and two-halo term (Section 3.1.2) in the following.

3.1.1 One-halo term

BMO define a truncated version of the NFW profile as

$$\rho_{\text{BMO}}(r) = \rho_{\text{NFW}}(r) \times \left(\frac{\tau_{200m}^2}{(r/r_{200m})^2 + \tau_{200m}^2} \right)^\beta, \quad (10)$$

where we use $\beta = 2$. Here we have introduced τ_{200m} , the multiple of r_{200m} around which the halo is smoothly truncated. The BMO profile can be analytically integrated to yield the projected profile $\Sigma_{1\text{h,BMO}}$, as well as the change in mass normalisation due to missing density relative to the NFW profile (cf. BMO; Oguri & Hamana 2011 and the AMPLE code, <http://kipac.stanford.edu/collab/research/lensing/ample/ample.c>).

Besides mass, our one-halo term has two additional free parameters, the concentration c_{200m} and truncation radius τ_{200m} . We fix the former with the mass-concentration relation of Duffy et al. (2008). Oguri & Hamana (2011) note a weak dependence of $\tau_{\Delta_{\text{vir}m}}$ on mass. Our scaling relation of τ_{200m} is determined by fitting the mean κ profile of haloes in our simulation (cf. Section 2) in logarithmic bins of $\Delta \log_{10} M_{200m} = 0.05$ in the radial range of 3 to 15 arcmin using the full profile of eqn. 9, with τ as a free parameter. The uncertainties for our χ^2 minimization w.r.t. τ are derived from the ensemble variance of mean κ in the respective mass bin. We find that the relation is well described as a linear function of ν . Our fit in $\nu \in [2.2, 3.7]$ yields

$$\tau_{200m} = \begin{cases} 3.85 - 0.73\nu & \nu \leq 3.7 \\ 1.15 & \nu > 3.7 \end{cases}, \quad (11)$$

where we have assumed a constant value for τ_{200m} above the peak height range covered by our simulated haloes.

We note that the fact that more massive haloes are, on average, truncated at smaller multiples of their r_{200m} radii could be interpreted in terms of the steepening of the

density slope with increasing mass accretion rate observed by Diemer & Kravtsov (2014).

3.1.2 Two-halo term

The two-halo surface mass density due to linear evolution around a cluster of mass M at redshift z can be written as (e.g. Oguri & Hamana 2011)

$$\Sigma_{2h,lin}(\theta) = b_h(M, z) \rho_{c,0} \Omega_m W(\theta, z) D_A^{-2}(z). \quad (12)$$

Here, $b_h(M, z)$ is the linear bias of the cluster halo (which we calculate according to Tinker et al. 2010) and $D_A(z)$ is the angular diameter distance to redshift z . We have defined the projected linear excess Lagrangian depth $W(\theta, z)$ as

$$\begin{aligned} W(\theta, z) &= [D_A(z)(1+z)]^2 \\ &\times \int_{-\Delta\zeta}^{+\Delta\zeta} \xi_{lin} \left(\sqrt{[D_A(z)(1+z)\theta]^2 + \zeta^2}, z \right) d\zeta \\ &= \int \frac{ldl}{2\pi} J_0(l\theta) P_{lin} \left(\frac{l}{(1+z)D_A(z)}, z \right), \end{aligned} \quad (13)$$

where $\pm\Delta\zeta$ is the interval in comoving distance along the LOS integrated in our simulated projected density profiles. For large enough $\Delta\zeta$, almost all correlated matter is included in the integration and the latter equality holds. In eqn. 13, we have used the Bessel function of the first kind and order zero, J_0 . W has units of comoving volume per solid angle and is readily interpreted as the excess in Lagrangian volume per solid angle that has moved to the projected vicinity of a structure with unit bias due to linear evolution.

The mass measurements of our haloes from the simulations contain all particles within r_{200m} . We therefore need to correct for contributions of two-halo matter M_{2h} when defining the mass of the one-halo profile, which we do at first order as

$$M_{200m,1h} = M_{200m} - M_{2h}(r_{200m}, M_{200m}, z). \quad (14)$$

Inside a sphere of radius r around the cluster centre, the mass of two-halo matter is

$$M_{2h}(r, M, z) = b_h(M, z) \rho_{c,0} \Omega_{m,0} U(r, z). \quad (15)$$

Here we have defined the linear excess Lagrangian volume inside a sphere of radius r as

$$\begin{aligned} U(r, z) &= \int_0^{r(1+z)} dr' \xi_{lin}(r', z) 4\pi r'^2 \\ &= D^2(z) U(r(1+z), 0). \end{aligned} \quad (16)$$

3.1.3 Comparison to simulations

Figure 1 shows model and mean profiles of simulated haloes in two mass bins spanning most of the dynamic range of our halo catalog (see Section 2 for details on the simulations). The model fits the data well at projected $r > 0.3 \times r_{200m}$, but moderately overestimates projected density at smaller radii, where an un-truncated NFW profile without two-halo contribution is a better fit. Potential reasons for the discrepancy include factual deviations of the simulated dark matter haloes from the NFW profile at small radii (cf. Becker & Kravtsov 2011, their Fig. 2, for an analysis based on the same cluster sample), resolution effects at small

radii and the simplified nature of our linear superposition of the collapsed halo profile with a linear two-halo term (cf. Hayashi & White 2008 for a different approach, in which the three-dimensional density is assumed to be piecewise equal to the NFW one-halo or a linear two-halo term, only). Since we only use the model for re-normalization of the ensemble mean profile to the mass of a given cluster (cf. Section 3.3), this is not problematic for our purposes.

3.2 Components of profile covariance

We have made the ansatz for the residual between a given, noise-free cluster κ profile \mathbf{K} and the mean profile at its mass $\kappa(M)$ as a sum of multivariate Gaussian vectors corresponding to four distinct physical effects: uncorrelated large scale structure along the LOS, variations in concentration, halo asphericity and orientation, and correlated haloes in the cluster substructure and neighbourhood (cf. eqn. 7). All of these effects have the potential to bias the lensing mass measurement or, equivalently, can be interpreted as a source of intrinsic covariance between the components of $\mathbf{K} - \kappa(M)$ that causes additional intrinsic noise on a cluster-by-cluster basis. In this section, we provide analytic expressions for all four components of the intrinsic covariance.

3.2.1 Uncorrelated large-scale structure

For a given source population, random structures along the LOS introduce a covariance in convergence measured in angular bins, which can be written as (e.g. Schneider et al. 1998; Hoekstra 2003; Umetsu et al. 2011)

$$C_{ij}^{LSS} = \int \frac{ldl}{2\pi} P_\kappa(l) \hat{J}_0(l\theta_i) \hat{J}_0(l\theta_j). \quad (17)$$

Here, $\hat{J}_0(l\theta_i)$ is the area-weighted average of J_0 over annulus i . Using the identity (Abramowitz & Stegun 1965, their eqn. 9.1.30)

$$\left(\frac{1}{z} \frac{d}{dz} \right)^k [z^l J_l(z)] = z^{l-k} J_{l-k}(z). \quad (18)$$

with $k = l = 1$ and integrating yields (cf. also Umetsu et al. 2011, their eqn. 16)

$$\hat{J}_0(l\theta_i) = \frac{1}{2dl\theta_i} [(1+d)J_1(l\theta_i(1+d)) - (1-d)J_1(l\theta_i(1-d))], \quad (19)$$

where we have assumed an annulus of width $2d\theta_i$ from $(1-d)\theta_i$ to $(1+d)\theta_i$.

The convergence power spectrum P_κ in eqn. 17 is obtained from the matter power spectrum by means of the Limber (1954) approximation as

$$P_\kappa(l) = \frac{9H_0^2\Omega_m^2}{4c^2} \int_{\chi_1}^{\chi_2} d\chi \left(\frac{\chi_s - \chi}{\chi_s a(\chi)} \right)^2 P_{nl}(l/\chi, \chi), \quad (20)$$

where we have assumed a fixed comoving distance of the sources χ_s . For calculating C^{LSS} around our cluster haloes, we set the source redshift to a fixed $z_s = 1$ and limit the interval $[\chi_1, \chi_2]$ to the range of $\pm 200h^{-1}$ comoving Mpc included in the boxes around the cluster haloes (cf. Section 2). For the non-linear matter power spectrum P_{nl} we use the model of Smith et al. (2003) with the Eisenstein & Hu

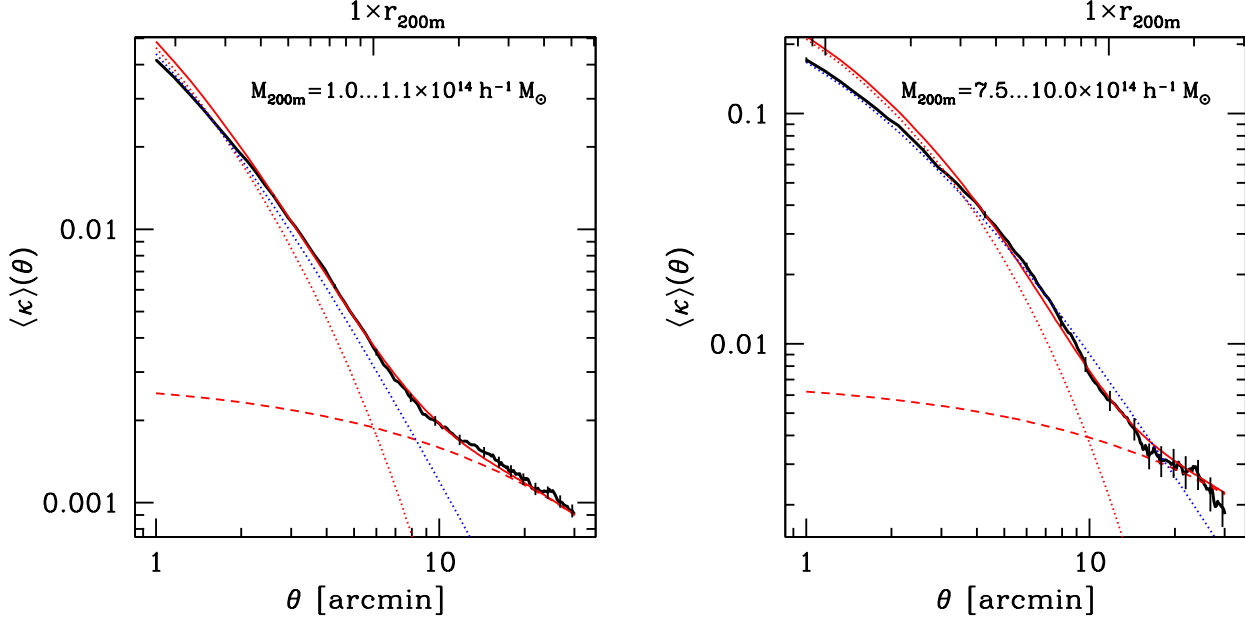


Figure 1. Mean convergence profiles of haloes at the low-mass (left) and high-mass (right) end of our sample, shown for haloes at the simulation redshift $z = 0.24533$ with corresponding angular (bottom axis) and r_{200m} (top axis) scales. Black lines show average profiles in the simulations, with error bars giving the uncertainty of the mean. Our model is shown as the red, solid line, composed of the BMO one-halo term (red, dotted) with truncation radius according to equation 11 and the two-halo term (red, dashed). The blue, dotted line shows un-truncated one-halo NFW profile for comparison.

(1998) transfer function including baryonic effects. We perform the integral using a customized version of NICAIA (Kilbinger et al. 2009). The result is shown in the lower left panel of Fig. 2.

3.2.2 Concentration

The covariance due to variations in concentration parameter at fixed mass is calculated as

$$C_{ij}^{\text{conc}} = \int dP(c) \kappa_i \kappa_j - \left[\int dP(c) \kappa_i \right] \times \left[\int dP(c) \kappa_j \right], \quad (21)$$

where κ_i is taken to be the convergence of a halo of concentration c in annulus i . For the probability density $dP(c)$, we assume a log-normal scatter of $\sigma_{\log_{10} c_{200m}} = 0.18$ (Bullock et al. 2001) about the mean concentration of Duffy et al. (2008). The result is shown in Fig. 2, lower second-to-left panel. As can be seen, it is subdominant at the radii considered here relative to other sources of intrinsic profile variation, at least for clusters of moderate mass, which illustrates the difficulty of measuring individual cluster concentrations with weak lensing.

3.2.3 Correlated large-scale structure

Clusters of galaxies are likely to form in overdense regions, where the abundance of additional (correlated) haloes is also higher than average. The mean effect of this on projected density is the two-halo term of Section 3.1.2. However,

the stochastic variation of the number of correlated haloes around a cluster also contributes to the covariance of κ .

We calculate this effect in a halo model, in analogy to Gruen et al. (2011, their Appendix A2). The idea is to split the set of all possible haloes into subsets that are alike in the sense that they cause a similar shear signal, e.g. ones that are of similar mass and projected distance from the cluster. We then apply Poissonian statistics to the number of haloes that is actually present from each subset.

Let $\mathbb{H} = \{\mathbf{h}\}$ be the set of tuples \mathbf{h} that completely characterize all possible halos, e.g. in terms of their coordinates and masses. Consider mutually exclusive and collectively exhaustive subsets $H_k \subset \mathbb{H}$ that are small enough such that $\forall H_k : \forall \mathbf{h}_a, \mathbf{h}_b \in H_k : \Sigma^i(\mathbf{h}_a) \approx \Sigma^i(\mathbf{h}_b)$ is a good approximation for the surface mass density of such a halo averaged over annulus i . Then we can write the surface mass density of correlated haloes Σ_{corr}^i in annulus i as

$$\Sigma_{\text{corr}}^i = \sum_k \lambda_k \Sigma^i(\mathbf{h}_k) \quad (22)$$

where $\lambda_k \geq 0$ is a random variable that describes the number of correlated haloes $\in H_k$ realized in a manifestation of the cluster. Defining a halo excess probability density dP_c around a cluster with properties \mathbf{h}_{cl} such that $\int_{H_k} dP_c(\mathbf{h}|\mathbf{h}_{cl}) = \langle \lambda_k \rangle$, this immediately results in the expectation value

$$\langle \Sigma_{\text{corr}}^i \rangle = \int dP_c(\mathbf{h}|\mathbf{h}_{cl}) \Sigma^i(\mathbf{h}). \quad (23)$$

For the covariance we make the assumption (discussed below) that the population random variables λ_k and λ_l of the halos in sets H_k and H_l are Poisson distributed and

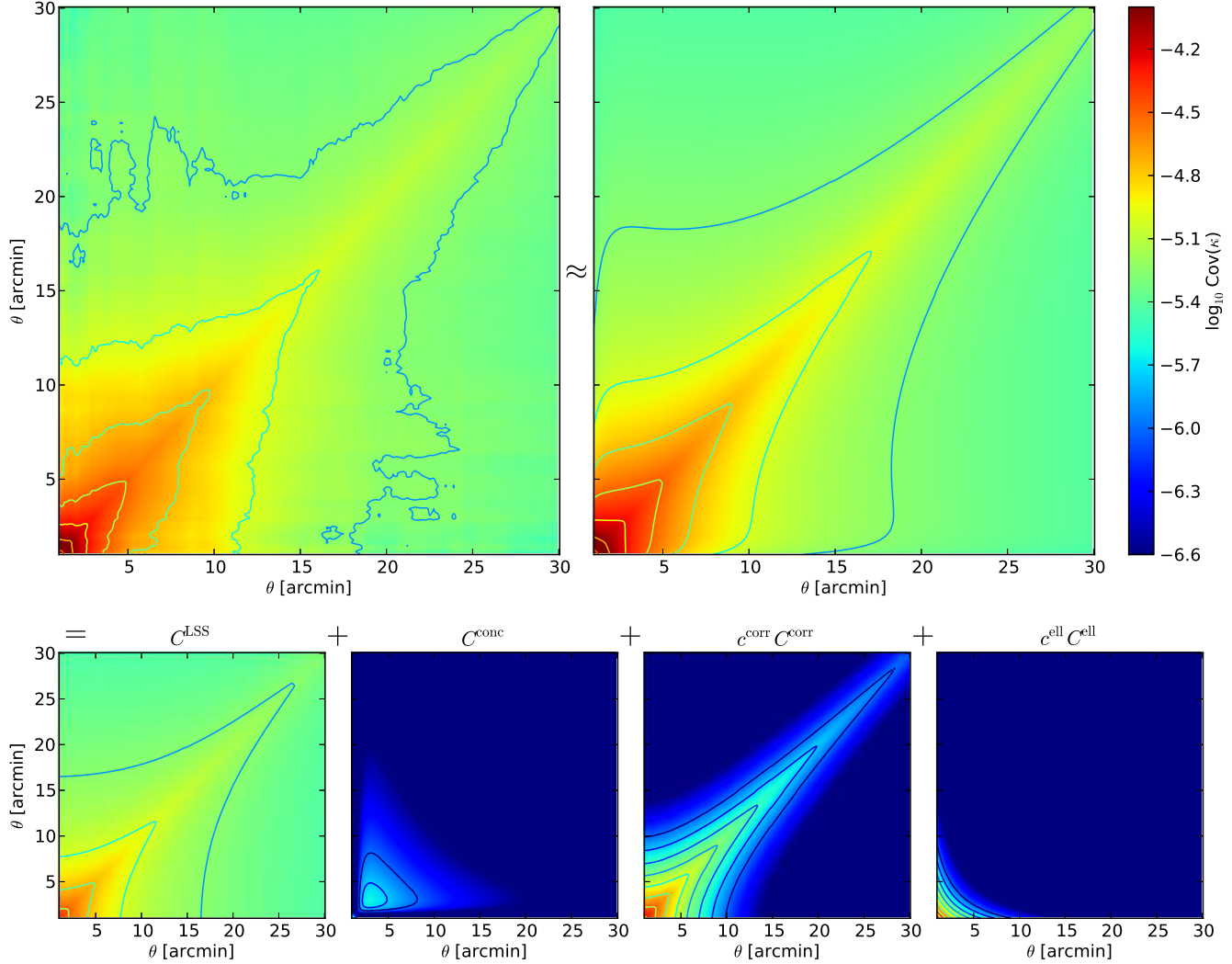


Figure 2. Illustration of our model for the profile covariance. Top left panel shows the empirical intrinsic covariance of κ profiles between 1 and 30 arcmin of simulated clusters of mass $M_{200m} = 1.8 \dots 2.2 \times 10^{14} M_{\odot}$. Our model prediction for the covariance is shown in the top right panel. It is a linear combination of the components shown in the bottom panel. Even at this relatively low mass, uncorrelated large-scale structure (left in lower panel) is an insufficient description of the observed covariance, especially in the inner regions. We model the additional covariance as a linear combination of variation of the concentration parameter, the covariance of correlated haloes and halo ellipticity and orientation (remaining plots in lower panel). All panels use the same color scale with contour lines spaced by $\Delta \log_{10} \text{Cov} \kappa = 0.25$. At the mass plotted here, $c^{\text{corr}} C^{\text{corr}}$ is the dominant component of the intrinsic (co)variance on most scales, with significant contributions from $c^{\text{ell}} C^{\text{ell}}$ in the innermost region and subdominant concentration scatter. We note, however, that the relative importance of these three components changes as a function of mass (cf. Fig. 3).

mutually independent ($\text{Cov}(\lambda_k, \lambda_l) = \delta_{kl} \lambda_k$). Then

$$C_{ij}^{\text{corr}} \times \Sigma_{\text{crit}}^2 = \int dP_c(\mathbf{h}|\mathbf{h}_{\text{cl}}) \Sigma^i(\mathbf{h}) \Sigma^j(\mathbf{h}). \quad (24)$$

If we characterize secondary haloes only by their mass and projected angular distance from the cluster centre, $\mathbf{h} = (M, \theta)$, we can write

$$dP_c(\mathbf{h}|\mathbf{h}_{\text{cl}}) = b(M_{\text{cl}})b(M) \frac{dN(M, z_{\text{cl}})}{dM dV} W(\theta, z_{\text{cl}}) 2\pi\theta d\theta dM, \quad (25)$$

Here we have introduced the halo mass function $\frac{dN(M, z_{\text{cl}})}{dM dV}$ and used the projected linear excess Lagrangian depth W from eqn. 13, which assumes that pairs of haloes of masses M_{cl} and M cluster according to the linear matter two-point

correlation and a mass dependent linear halo bias b as

$$\xi_{\text{hh}}(M_{\text{cl}}, M, r) = b(M_{\text{cl}}, z_{\text{cl}})b(M, z_{\text{cl}})\xi_{\text{lin}}(r, z_{\text{cl}}). \quad (26)$$

We perform the two-dimensional integral of equations 24 and 25, factoring out the cluster halo bias $b(M_{\text{cl}}, z_{\text{cl}})$ for later re-scaling of the covariance matrix to clusters of any mass. For the halo mass function and halo bias we use the models of Tinker et al. (2008, 2010). The haloes are modelled as BMO profiles (cf. Section 3.1.1) with concentration according to Duffy et al. (2008) and the truncation radius model of Eqn. 11. We include contributions from haloes of mass $10^8 h^{-1} M_{\odot} \leq M_{200m} \leq 10^{15.5} h^{-1} M_{\odot}$.

Finally, we note that a mutual (three-point) correlation of different correlated haloes is in fact expected, i.e. the presence of one massive, secondary halo makes the presence of

tertiary haloes more likely. This means that the approach presented above yields merely a lower limit of the true covariance, yet this can be approximately compensated by an empirical re-scaling of the covariance matrix with c^{corr} . The resulting covariance matrix is shown in the lower second-to-right panel of Fig. 2.

Non-linearly correlated subhaloes

The model for correlated large-scale structure in Section 3.2.3 assumed a linear correlation of secondary haloes at all radii. A more realistic approach might be to make this assumption only for haloes outside the virial radius and add a distinct population of subhaloes. We test this model as follows.

For the subhalo abundance inside the virial radius (Bryan & Norman 1998, their eqn. 6) we use the subhalo mass function of Jiang & van den Bosch (2014). The surface density of haloes is distributed according to an NFW profile with concentration $c_{200m} = 3.9$, adapted from the measurement of $c_{200c} = 2.6$ by Budzynski et al. (2012) for Duffy et al. (2008) halo concentration, $M_{200m} = 2 \times 10^{14} h^{-1} M_{\odot}$ and the snapshot redshift. These assumptions yield dP_c for the subhalo case.

We assume a truncation of subhaloes at their r_{200m} radius, i.e. $\tau_{200m} = c_{200m}(M_{\text{sub}})$, motivated by halo stripping. For the projected density profile of the individual subhalo, we ensure mass compensation by subtracting as much matter according to the mean subhalo density profile (eqn. 23) as is contained in the individual halo. Otherwise we apply the prescription of Section 3.2.3 to determine the subhalo shot noise covariance matrix C^{sub} . In this we are ignoring, as before for the linear correlation, both a potential dependence of the subhalo number density profile on the parent halo concentration and ellipticity (which might cause correlations between C^{sub} and the respective other components of the model) and mutual correlation of subhaloes.

The covariance for linearly correlated haloes in this model is calculated as in eqn. 24, yet using a virial sphere excised version of $W(\theta)$ to calculate $dP_c(\mathbf{h}|\mathbf{h}_{\text{cl}})$ in eqn. 25. Note that C^{corr} in this model explicitly depends on the halo mass rather than being a constant template that is re-scaled by the central halo bias.

We find that the predicted subhalo covariance is a significant contribution to the intrinsic covariance of the halo profile at the high mass end of our sample. The contribution is dominated by massive subhaloes that are sufficiently resolved by our simulations. When we fit a linear dependence of a re-scaling parameter $c^{\text{sub}} = c_0^{\text{sub}} + c_1^{\text{sub}}(\nu - \nu_0^{\text{sub}})$ together with $c_{0/1}^{\text{corr}}$ and $c_{0/1}^{\text{ell}}$ as described in Section 4 below, we find $c_0^{\text{sub}} = 0.8 \pm 0.8$ and $c_1^{\text{sub}} = -2 \pm 1.5$. Fixing $c_1^{\text{sub}} = 0$ and $c_1^{\text{corr}} = 0$ yields a best-fit $c_0^{\text{sub}} = 0$. The contribution of subhaloes is particularly degenerate with the halo asphericity covariance C^{ell} (see Section 3.2.4). We conclude that a larger sample of massive haloes would be required to determine whether the addition of C^{sub} improves the model and therefore do not include it in the following analyses.

3.2.4 Halo asphericity

Dark matter haloes are known to be triaxial in general. The axis ratios and their orientation along the LOS can change the projected κ profile significantly. It is difficult, however, to model the effect in its full generality.

We therefore make the following simplified model for the covariance due to halo asphericity. In accordance with the dominant feature of halo shapes in nature, we assume them to have a prolate shape with minor-to-major axis ratio $0 < q = b/a = c/a \leq 1$. We define a coordinate system with ellipsoidal radius r_e such that

$$\mathbf{r}_e^2 = \mathbf{x}^T \begin{pmatrix} q^{-2/3} & 0 & 0 \\ 0 & q^{-2/3} & 0 \\ 0 & 0 & q^{4/3} \end{pmatrix} \mathbf{x}, \quad (27)$$

where $\mathbf{x}^T = (x, y, z)$ is a Cartesian coordinate system centred on and aligned with z along the major axis of the halo. For the spherical coordinate system (R, θ, ϕ) with $\theta = 0$ along the major axis,

$$r_e(R, \theta)/R = \sqrt{q^{-2/3} \sin^2 \theta + q^{4/3} \cos^2 \theta}. \quad (28)$$

We find the three-dimensional density of a prolate NFW halo by evaluating equation 8 at $r = r_e$. By virtue of the unit determinant, the volume of ellipsoidal shells is $V(r_e, r_e + dr_e) = 4\pi r_e^2 dr_e$. The usual normalization of ρ_0 therefore means that the density integrated out to an ellipsoidal radius $r_e = r_{200m}$ matches M_{200m} . However, we need to re-scale ρ_0 numerically for the mass inside a sphere of radius r_{200m} to match M_{200m} (cf. Jing & Suto 2002; Oguri et al. 2003; Corless & King 2007 for definitions based on ellipsoidal overdensity and Dietrich et al. 2014 for a spherical overdensity approach to ellipsoidal haloes).

Integration of ρ along the LOS yields the surface mass density for any combination of M_{200m} , c_{200m} , q and orientation angle α , where $\alpha = 0$ puts the major axis along the LOS. We take care to use an approximation to the analytical result for the integrated density inside the innermost region, $r/r_s < 10^{-3}$, where the diverging density leads to numerical instability. Here we can approximate $\theta \approx \alpha$ for most of the matter along the LOS to find the mean surface density inside a small projected radius $\delta v = \delta r/r_s$ as (cf. Bartelmann 1996; Wright & Brainerd 2000)

$$\begin{aligned} \langle \Sigma \rangle(< \delta v) &= \rho_0 r_s \\ &\times ((1 - e^2)^{-1/3} \sin^2 \alpha + (1 - e^2)^{2/3} \cos^2 \alpha)^{-1/2} \\ &\times \frac{4}{\delta v^2} \left(\frac{2}{\sqrt{1 - \delta v^2}} \operatorname{arctanh} \left(\sqrt{\frac{1 - \delta v}{1 + \delta v}} \right) + \ln \left(\frac{\delta v}{2} \right) \right). \end{aligned} \quad (29)$$

The covariance matrix is integrated as

$$\begin{aligned} C_{ij}^{\text{ell}} &= \int dP(q, \cos \alpha) \kappa_i \kappa_j \\ &- \left[\int dP(q, \cos \alpha) \kappa_i \right] \times \left[\int dP(q, \cos \alpha) \kappa_j \right], \end{aligned} \quad (30)$$

where κ_i is the mean surface density of a halo of axis ratio q and orientation angle α in annulus i . We assume isotropic orientation, i.e. a uniform distribution of $\cos \alpha \in [0, 1]$, and a truncated Gaussian distribution of q

$$P(q) \propto \begin{cases} \mathcal{N}(\mu = 0.6, \sigma = 0.12), & 0.1 \leq q \leq 1 \\ 0, & q < 0.1 \vee q > 1 \end{cases}, \quad (31)$$

approximating the distribution of q for haloes of mass $\approx 10^{14} M_\odot$ as measured in simulations by Bett et al. (2007, cf. their Section 4.3). The resulting covariance matrix, re-scaled empirically by c^{conc} , is shown in the lower right panel of Fig. 2. We see that the influence of halo ellipticity is limited to small radii compared to correlated and uncorrelated structures along the LOS, yet quite significant in that regime.

3.3 Covariance estimation

Given a model for $\kappa(M_{200m})$ that we can subtract from the observed \mathbf{K} , any simulated cluster of mass M_{200m} yields an estimate for C as

$$\text{Cov}(E_i, E_j) = \langle E_i E_j \rangle. \quad (32)$$

While this approach heavily relies on the accuracy of κ , it is also possible to estimate the covariance matrix $C(M_{200m})$ from an ensemble of N simulated clusters of fixed mass M_{200m} without assuming a model for the mean profile. This is done by applying the ensemble covariance estimator

$$\text{Cov}(E_i, E_j) = \frac{N}{N-1} \langle \hat{E}_i \hat{E}_j \rangle, \quad (33)$$

where \hat{E} is the residual with respect to the ensemble mean profile. The variance of the estimator of eqn. 33 is larger than the one of eqn. 32 by a factor $N/(N-1)$. We decide to bin our clusters in mass in subsamples of $N = 24$, such that the loss of information due to not assuming a model for the true mean profile is negligible. However, especially at the massive end where the number of clusters in our simulations is small, this would introduce an additional variance of profiles due to the systematic change of mass within each subsample. We therefore define

$$\hat{E}_i^k = \sqrt{\frac{N}{N-1}} \left(K_i^k - N^{-1} \sum_{l=1}^N K_i^l \times \frac{\kappa_i^k}{\kappa_i^l} \right), \quad (34)$$

where K_i^j and κ_i^j are the actual and model convergence of cluster j in radial bin i . Here, we have re-scaled each of the other cluster profiles to the expected value at the mass of cluster k by means of the model and amplified the deviations from the mean by $\sqrt{N/(N-1)}$ to correct for the bias of the maximum-likelihood ensemble variance, such that $\text{Cov}(\hat{E}_i, \hat{E}_j) = \text{Cov}(E_i, E_j)$.

4 DETERMINATION OF COVARIANCE MODEL PARAMETERS

The multivariate Gaussian of equation 5 corresponds to a log-likelihood

$$\begin{aligned} -2 \ln \mathcal{L} &= \ln \det C(M) \\ &+ (\kappa(M) - \mathbf{K})^T C^{-1}(M) (\kappa(M) - \mathbf{K}) \\ &+ \text{const}. \end{aligned} \quad (35)$$

This could be interpreted as a likelihood of mass or of the parameters of both the mean profile model $\kappa(M)$ and the covariance model $C(M)$. For the latter, the first term on the right-hand side serves as a regularization that prevents run-off of the (co)variance estimate to infinity. In our case, as discussed above, we use a non-parametric profile model

to replace $\kappa(M) - \mathbf{K}$ by the $\hat{\mathbf{E}}$ of equation 34 and maximize the likelihood to constrain the parameters of the covariance model.

Equation 35 has the disadvantage that the precision matrix $C^{-1}(M)$ and $\det C(M)$ become numerically unstable due to the large conditional number and strong covariance between neighbouring radial bins. Related to this, the uncertainty of empirically estimated off-diagonal components of the covariance matrix is large (e.g. Taylor et al. 2013, their eqn. 18). We therefore decide to use equation 35, however with a diagonal model covariance matrix, i.e. one where all off-diagonal components are set to zero. We verify, using a toy model, that this yields an unbiased maximum-likelihood estimate of the covariance model parameters.

For the covariance model, we assume equation 7. The contribution of concentration becomes important only at large mass and small radii, and rather than (poorly) constraining it from the data we decide to adopt the Bullock et al. (2001) log-normal concentration scatter, i.e. we set $c^{\text{conc}} = 1$.

As a baseline for $c^{\text{corr}}(\nu)$ and $c^{\text{ell}}(\nu)$ we use the mass-independent model

$$\begin{aligned} c^{\text{corr}}(\nu) &= c^{\text{corr}} = \text{const} \\ c^{\text{ell}}(\nu) &= c^{\text{ell}} = \text{const}. \end{aligned} \quad (36)$$

As a test for mass dependence we also run our analysis with a model that allows for linear evolution in ν ,

$$\begin{aligned} c^{\text{corr}}(\nu) &= c_0^{\text{corr}} + (\nu - \nu_0^{\text{corr}}) c_1^{\text{corr}} \\ c^{\text{ell}}(\nu) &= c_0^{\text{ell}} + (\nu - \nu_0^{\text{ell}}) c_1^{\text{ell}}, \end{aligned} \quad (37)$$

where we set the pivot points $\nu_0^{\text{corr}} = 2.5$ and $\nu_0^{\text{ell}} = 2.8$ to make errors on c_0^* and c_1^* approximately uncorrelated in the bootstrap runs.

4.1 Results

The best-fitting values for the parameters of eqn. 36 with errors estimated from bootstrapping are $c^{\text{corr}} = 5.0 \pm 0.3$ and $c^{\text{ell}} = 3.7 \pm 0.3$. Allowing for a linear dependence as in eqn. 37 these values are unchanged at best fit and we get no indication (but relatively poor constraints) for a ν dependence with $c_1^{\text{corr}} = -0.04 \pm 0.88$ and $c_1^{\text{ell}} = -0.05 \pm 0.79$. For the remainder of the analysis, we therefore adopt ν independent $c^{\text{corr}} = 5.0$ and $c^{\text{ell}} = 3.7$.

Figure 2 shows the full covariance model with these parameters for $M_{200m} = 2 \times 10^{14} h^{-1} M_\odot$. Figure 3 shows model and data variance for a range of masses. The model reproduces observed intrinsic variations well over a wide range of mass. We note, however, the large uncertainty in data variance due to our limited sample size.

4.2 Redshift dependence

In order to test for a potential redshift dependence of our model parameters, we repeat the analysis with a second snapshot at $z = 0.499$, using 788 haloes at $M_{200m} \geq 2 \times 10^{14} h^{-1} M_\odot$.

The best-fitting truncation radii are consistent with eqn. 11. We fit the parameters of eqn. 36 as in Section 4 and determine errors using bootstrapping. Empirical covariances are well reproduced by a model with $c^{\text{corr}} = 4.2 \pm 1.1$

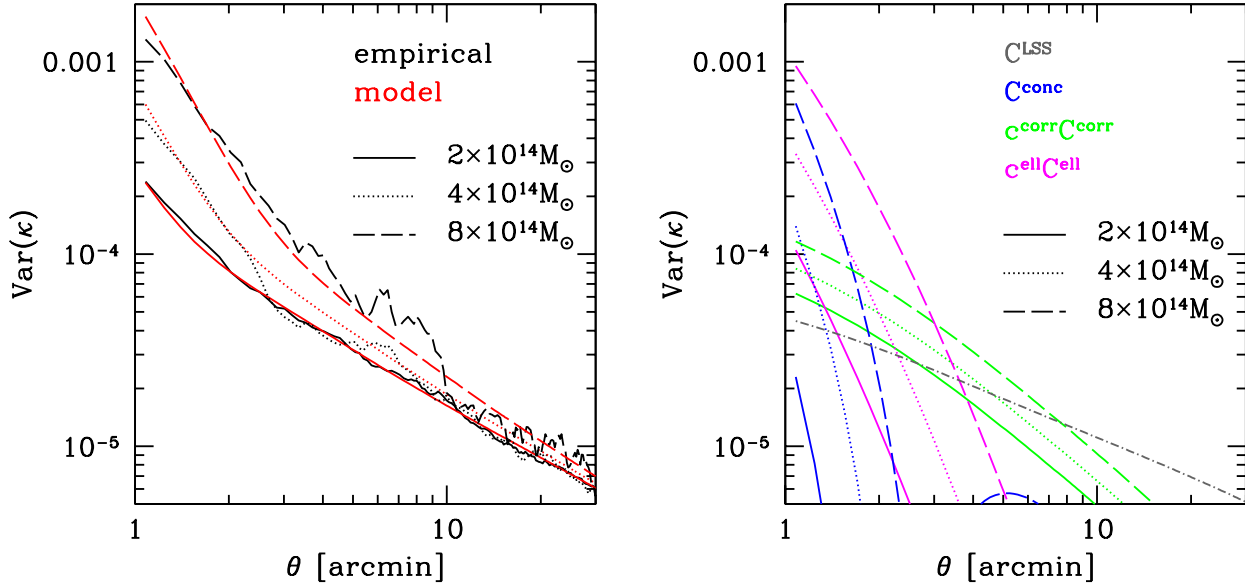


Figure 3. Empirical covariance (black, left panel) and covariance model (red, left panel) for clusters of mass $M_{200m} = 1.8 \dots 2.2 \times 10^{14} h^{-1} M_{\odot}$ (solid lines), $M_{200m} = 3.5 \dots 4.5 \times 10^{14} h^{-1} M_{\odot}$ (dotted) and $M_{200m} = 7 \dots 9 \times 10^{14} h^{-1} M_{\odot}$ (dashed lines). Right panel shows components of the model with $c^{\text{corr}} = 5.0$ and $c^{\text{ell}} = 3.7$ for all three cases (uncorrelated LSS inside the snapshot box shown as grey, dotted-dashed line).

and $c^{\text{ell}} = 3.8 \pm 0.9$. Due to the much smaller sample size and lower contrast of the haloes relative to LSS noise, the parameters are significantly less well constrained as in the $z = 0.245$ snapshot, however, consistent within the uncertainties.

We conclude that at the present level of statistical certainty the best-fit values of Section 4.1 can be used over the most relevant redshift range for cluster lensing of $z \approx 0.2 \dots 0.5$.

5 EFFECT ON WEAK LENSING CLUSTER SURVEYS

In this section, we estimate the influence of intrinsic covariance of density profiles at fixed mass

- on mass measurements of individual clusters (Section 5.1)
- and on the determination of parameters of the mass-observable relation (MOR) in a lensing follow-up of a sample of clusters (Section 5.2).

In our analyses we fix the lens redshift to $z_l = 0.24533$, the primary redshift of our simulated cluster profiles. We consider shear surveys of different depths, parametrized by the surface density \mathbf{n} of source galaxies, which we assume to lie at a fixed $z_s = 1$. The three settings chosen roughly correspond to current large ground based surveys ($\mathbf{n} = 10 \text{ arcmin}^{-2}$), the best available ground-based data ($\mathbf{n} = 50 \text{ arcmin}^{-2}$) and space-based data ($\mathbf{n} = 100 \text{ arcmin}^{-2}$).

The full covariance of equation 7 contains, apart from uncorrelated and intrinsic variations of surface mass density, the measurement uncertainty of the survey C^{obs} . For a shear survey, assume that we can measure the tangential

gravitational shear γ in each annulus i with variance

$$\sigma_{\gamma,i}^2 = \frac{\sigma_{\epsilon}^2}{\mathbf{n} A_i}, \quad (38)$$

where $\sigma_{\epsilon} \approx 0.3$ is the shape noise (including intrinsic shape dispersion and measurement noise), \mathbf{n} is the background source density and A_i the area of annulus i . The n -dimensional covariance matrix of γ (ignoring intrinsic alignment and shear systematics) is diagonal with $C_{ii}^{\text{obs},\gamma} = \sigma_{\gamma,i}^2$.

In the limit that annuli are thin, γ and κ are connected by a linear equation $\gamma = G\kappa$, where the $n \times n$ matrix G is defined such that

$$\gamma_i = \langle \kappa \rangle_{<i} - \kappa_i = \sum_{j=0}^{i-1} \kappa_j A_j / (\pi \theta_{i,\min}^2) - \kappa_i. \quad (39)$$

Our goal is to calculate the observational covariance C^{obs} of κ , which can be written as

$$C^{\text{obs}} = G^{-1} C^{\text{obs},\gamma} (G^{-1})^T. \quad (40)$$

The above derivation has omitted the technical step of breaking the mass-sheet degeneracy, which is necessary for G to be of full rank. To this end, we increase the dimensionality by 1 and let the newly introduced entry γ_{n+1} represent the convergence in the outermost radius, and the new κ_{n+1} be the mean convergence inside the innermost annulus. The measurement error of γ_{n+1} is defined to be the LSS variance of κ on the outermost angular scale (this is essentially what happens in the common assumption of $\kappa = 0$ on large scales). The mean convergence inside the innermost annulus is connected to γ by the linear equation above. We later exclude the added component from our analysis again, i.e. we consider only the sub-matrix $i, j \in [1, n]$ (similar in effect to Mandelbaum et al. 2010).

Apart from the *full model covariance* $C(M)$ of equation 7, we also define the *covariance without intrinsic variations*

$$C^{\text{point}} = C^{\text{obs}} + C^{\text{LSS}}. \quad (41)$$

For a model independent representation, we calculate the ensemble covariance of the re-scaled profiles (cf. eqn. 34) as $\hat{C}^{\text{int}}(M)$. When calculating $\hat{C}^{\text{int}}(M)$, we use 100 clusters that are nearest neighbours in a mass-ordered list of our simulated haloes to the target mass M and subtract the uncorrelated LSS covariance inside the simulation box. The *empirically estimated covariance* is

$$C^{\text{emp}}(M) = C^{\text{obs}} + C^{\text{LSS}} + \hat{C}^{\text{int}}(M). \quad (42)$$

5.1 Mass confidence intervals

We determine the effect of intrinsic variations of the density profile on the validity of confidence intervals of weak lensing mass measurements. In this procedure, we assume the empirically estimated density profile covariance from our ensemble of simulated haloes to represent the true variability. We test for the effects of either taking into account intrinsic variations with the model proposed in this work or ignoring them. We design this test as follows.

At given mass M , we define and determine the full empirically estimated covariance $C^{\text{emp}}(M)$ as described in eqn. 42. In a Monte Carlo simulation, we then generate profiles by adding a multivariate Gaussian random vector according to $C^{\text{emp}}(M)$ to the model profile $\kappa(M)$ (eqn. 9) of mass M .¹

Consider one such realization \mathbf{K} . The likelihood in eqn. 35, now interpreted as a function of mass $\mathcal{L}(M'|\mathbf{K})$, can then be run over a range of model profiles $\kappa(M')$ with different mass M' . For $C(M')$ and $C^{-1}(M')$ we use

- either the covariance without intrinsic variation C^{point} of equation 41 (colour coded red in the following figures)
- or the covariance of equation 7, including our parametric model for intrinsic variations (colour coded blue).

Note that the latter option requires to include the $\ln \det C$ term of the likelihood, since covariance is now a function of mass.

We determine 68% (and 90%) confidence intervals using mass limits where the likelihood is $2\Delta \ln \mathcal{L} = 1$ (and 2.7) worse than at best fit. We repeat the procedure for 10,000 realizations at every fixed mass and determine the empirical coverage, e.g. the fraction of cases in which the true mass lies inside these confidence intervals.

Results for three different survey depths are shown in

¹ As a test for the validity of considering only multivariate Gaussian variations over the mean profile, we perform another run where instead of a synthetic variation we use 100 nearest neighbors in mass around M , re-scale their profile to M using the model profiles of eqn. 9 and subtract their mean to generate 100 realistic random variations. We then add these (and multivariate Gaussian random vectors to simulate observational and uncorrelated LSS covariance) to the model profile at M and run the likelihood as described below. Results in this approach do not differ significantly from the synthetic multivariate Gaussian ones presented here.

Figure 4, where we now compare results of the two methods (red and blue lines) to the target levels of 68 and 90% (black lines). Increasingly with increasing mass, ignoring intrinsic variations leads to an underestimation of errors. Even for present surveys with relatively shallow depth, this manifests in $\approx 15\%$ excess outliers at the one and two sigma level for massive clusters at around $M_{200m} \approx 10^{15} h^{-1} M_{\odot}$. The effect is more severe in deeper data, where the relative importance of intrinsic over observational covariance increases.

Using our model covariance, confidence intervals are correctly estimated over a wide range of masses and even in deep data. The small number of systems above $M_{200m} > 10^{15} h^{-1} M_{\odot}$, however, limits the range of masses where we can evaluate the validity of our model. Note in this context that the scatter in the plots is due to the noise in $\hat{C}^{\text{int}}(M)$, as estimated on a limited number of clusters, rather than the number of realizations we run (the effect of the latter is indicated as the small vertical bar on the black lines).

5.1.1 Effect of intrinsic covariance on mass uncertainty

In the previous section, we found that the confidence interval based on a κ covariance matrix without contributions from intrinsic profile variation is significantly too narrow, especially for more massive clusters and larger depth of the data. Here, we perform a Fisher analysis to determine how well masses can be measured, a question particularly relevant for the design of lensing follow-up programmes.

For a mass dependent covariance matrix, the Fisher information for mass \mathcal{F} reads (e.g. Vogeley & Szalay 1996; Tegmark et al. 1997)

$$\mathcal{F}(M) = \left(\frac{d\kappa}{dM} \right)^T C^{-1} \left(\frac{d\kappa}{dM} \right) + \frac{1}{2} \text{tr} \left[\left(C^{-1} \frac{dC}{dM} \right)^2 \right], \quad (43)$$

where κ , C and their derivatives are taken to be evaluated at mass M .

We calculate the Fisher information for the model without intrinsic variations (eqn. 41) and our full model. In the first case, the second term of eqn. 43 drops because of the mass independence of the covariance. Including intrinsic variations, we also evaluate the second term, which however contributes to the Fisher information at most at the percent level. The mass uncertainty σ_M is related to the Fisher information as $\sigma_M = \mathcal{F}^{-1/2}$.

Figure 5 shows relative uncertainties in mass σ_M/M due to a shear surveys of various depth. The increase in uncertainty due to intrinsic profile variation is substantial, even for shallow data with sufficiently massive clusters, but more strongly for deeper observations. Our results are consistent with the findings of Becker & Kravtsov (2011), who used the same simulations to derive uncertainties empirically rather than with a covariance model.

5.2 Mass-observable relations

We study the effect of intrinsic variations of projected cluster density profiles on weak lensing follow-up surveys used for determining parameters of a mass-observable relation. Consider an observable Y (which could be the Compton decrement, an X-ray mass proxy or an optical richness) with a

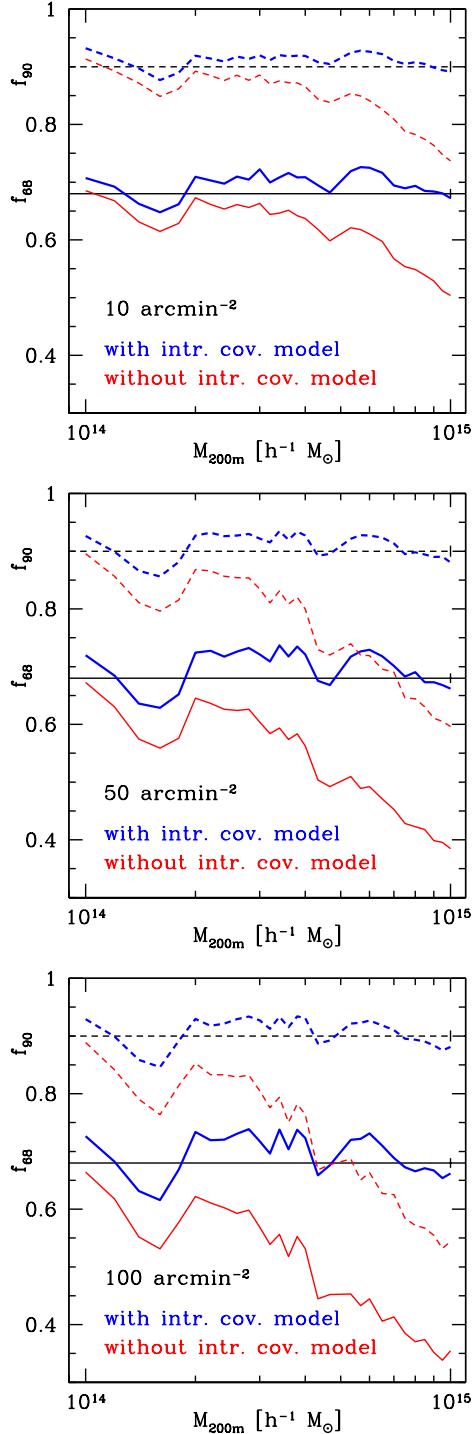


Figure 4. Fraction of weak lensing measurements of cluster mass where the true mass is inside the 68% (f_{68} , solid lines) and 90% (f_{90} , dashed lines) confidence interval (black lines at target levels for reference). In red, thin lines we show the common case where only measurement noise and uncorrelated LSS are considered for the covariance matrix (eqn. 41). Blue, thick lines show results that include our model for the intrinsic covariance (eqn. 7). Panels correspond to typical large ground based-surveys (top), best available ground based data (centre) and space-based analyses (bottom). As data quality and cluster mass increase, intrinsic variations become a significant component of the uncertainty budget of weak lensing mass measurements and should not be ignored. All plots are for lenses at $z_l = 0.24533$.

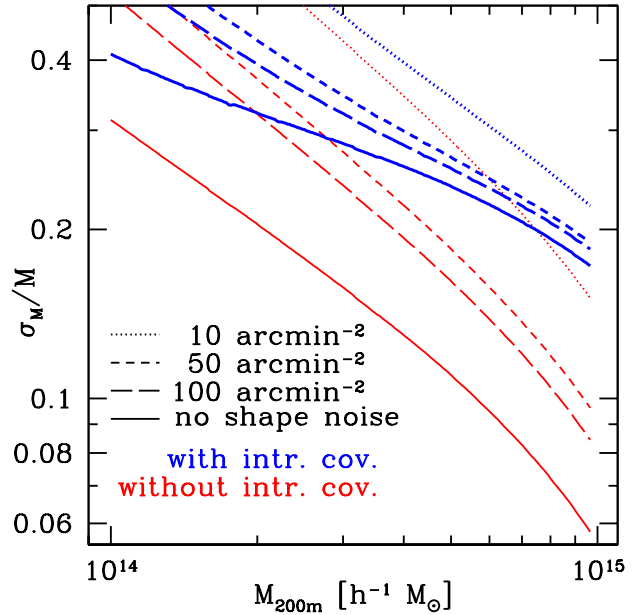


Figure 5. Relative uncertainty of weak lensing mass measurements from a Fisher analysis excluding (red) and including (blue) intrinsic covariance according to our model. Not including intrinsic covariance leads to an underestimation of uncertainty, especially at high mass and large depth. Notably, the improvement of uncertainty for massive clusters going from typical (10 arcmin^{-2} background galaxies, dotted line) to noiseless measurements (solid line) is only modest when intrinsic variations are considered. Results are for lenses at $z_l = 0.24533$.

power-law mass-observable relation that describes the fiducial value $Y_0(M)$ as a function of mass M ,

$$\ln Y_0(M)/\hat{Y} = A + B \ln M/\hat{M}. \quad (44)$$

We have introduced here pivot values \hat{Y} and \hat{M} in addition to the power-law slope B and amplitude A .

The observed value of Y for any cluster shall include a log-normal intrinsic scatter σ_{int} and, for simplicity, also log-normally distributed measurement related uncertainty σ_{obs} . The combined uncertainty $\sigma = \sqrt{\sigma_{\text{int}}^2 + \sigma_{\text{obs}}^2}$ leads to an observable

$$\ln Y(M) = \ln Y_0(M) + \mathcal{N}(\mu = 0, \sigma). \quad (45)$$

5.2.1 Likelihood

Consider a single cluster with observed convergence profile \mathbf{K} and measured observable Y . We maximize, as a function of MOR parameters, the likelihood

$$P(\mathbf{K}|Y, A, B, \sigma_{\text{int}}) = \int P(\mathbf{K}|M)P(M|Y, A, B, \sigma_{\text{int}}) dM \\ \propto \int P(\mathbf{K}|M)P(Y|M, A, B, \sigma_{\text{int}})P(M) dM. \quad (46)$$

For $P(\mathbf{K}|M)$ we insert the expression of eqn. 5 with C equal to either our full model or the covariance without intrinsic

variations. The term for the Y likelihood can be written as

$$P(Y|M, A, B, \sigma_{\text{int}}) = \frac{\exp[-(\ln Y_0(M) - \ln Y)^2 / (2\sigma^2)]}{\sqrt{2\sigma}} \times \Theta(Y - Y_{\text{lim}}) P_{\text{det}}^{-1}(M). \quad (47)$$

Here we have used the Heaviside step function Θ to impose the observable limit Y_{lim} and re-normalized by the inverse of the detection probability $P_{\text{det}}(M) = \text{erfc}[(\ln Y_{\text{lim}} - \ln Y_0(M))/(\sqrt{2}\sigma)]/2$ to compensate for the removed part of the probability distribution (cf. Vikhlinin et al. 2009a, their eqn. A10) and correct for the (Malmquist) bias due to preferential selection of objects with positive contribution from scatter.

Finally, the normalized mass prior $P(M)$ can be written for an observable limited survey as

$$P(M) \propto \frac{dn}{dM}(M) P_{\text{det}}(M) \left[\int dM' \frac{dn}{dM}(M') P_{\text{det}}(M') \right]^{-1} \quad (48)$$

to correct for (Eddington) bias due to the increase of abundance with decreasing mass. Note that $P(M)$ depends on the MOR by means of P_{det} and, by means of the halo mass function $\frac{dn}{dM}$, also on cosmology (for which we, however, assume fixed values in the simulations presented here).

5.2.2 Results

We simulate samples of clusters $1 \leq i \leq n$, where cluster i is characterized by its observable Y_i and an observed convergence profile \mathbf{K}_i .

For Y , we assume a MOR with $B = 5/3$ (the self-similar slope for Compton decrement Y_{SZ} and the X-ray equivalent Y_{X}) and choose the pivots such that measurement errors on A and B are uncorrelated and $A = 0$. We assume an intrinsic scatter $\sigma_{\text{int}} = 0.1$ and observational uncertainty of $\sigma_{\text{obs}} = 0.05$.

The convergence profile \mathbf{K} is simulated including a multivariate Gaussian deviation from eqn. 9, according to the full covariance (including uncorrelated LSS, measurement uncertainty of the survey and intrinsic variation as predicted from our model) of eqn. 7.

We take the sample to be observable limited, i.e. draw clusters with abundance proportional to the halo mass function, assign observables according to eqns. 44-45 and only accept objects where the observable exceeds the survey threshold, $Y > Y_{\text{lim}}$. The threshold Y_{lim} is chosen as $Y_{\text{lim}} = Y_0(M_{200m} = 4 \times 10^{14} h^{-1} M_{\odot})$, comparable for instance to ongoing SZ surveys. The number of clusters drawn is taken to be representative of a volume of $V = 0.15 h^{-3} \text{Gpc}^3$ (comoving, at $z = 0.24533$), which contains roughly 100 detections in our simulated survey.

Figure 6 shows the distribution of maximum likelihood estimates of the MOR parameters A , B and σ_{int} , with contours enclosing 68 and 95 per cent of the 15,000 realizations, respectively. We find that the analysis using the full covariance (blue) reproduces the input MOR well. The feature near $\sigma_{\text{int}} = 0$ is due to the hard prior $\sigma_{\text{int}} > 0$, which pushes random realization with lower empirical scatter towards this limit. The analysis using the covariance without intrinsic variations (red) significantly overestimates intrinsic scatter. As an additional bias, the overall mass scale A (and, although less strongly, the slope B) are systematically

underestimated. The move from a shallow (dotted) to deep (solid lines) surveys only moderately improves constraints and makes the bias due to using the covariance without intrinsic variations more apparent.

We note that the biases in A and B are due to a degeneracy between them and the intrinsic scatter. If the latter could be constrained externally with small uncertainty, e.g. by means of realistic simulations or by combining several observables with uncorrelated intrinsic scatter, the bias in the former would be mitigated even without a realistic model for the lensing covariance.

6 CONCLUSIONS

We have presented a model for the variation of projected density profiles of clusters of galaxies at fixed mass, constructed as a combination of the effects of variations of halo concentration, ellipticity and orientation, and correlated secondary structures.

The full covariance including our model for intrinsic variations faithfully reproduces confidence intervals in the weak lensing likelihood of cluster mass. We show that when intrinsic variations are ignored, uncertainties in lensing-derived mass are underestimated significantly (cf. Fig. 4).

Using the full covariance model we have made Fisher predictions for the accuracy of lensing measurements of cluster mass. We have shown that intrinsic variations take away some of the comparative advantage of studying a small sample of the most massive clusters with the deepest possible observations. For a massive cluster ($M_{200m} \approx 10^{15} h^{-1} M_{\odot}$ at $z = 0.25$) we find an irreducible relative uncertainty in lensing mass of ≈ 20 per-cent due to intrinsic profile variations and uncorrelated LSS along the LOS, three times higher than the uncertainty from uncorrelated LSS alone. Our results agree with the analysis of Becker & Kravtsov (2011).

With simulations of mock surveys for constraining cluster mass-observable relations with lensing data, we have shown that intrinsic variations significantly bias the derived intrinsic scatter and amplitude if they are not accounted for in the lensing mass likelihood. For a follow-up of a sample of 100 clusters selected by Y_{SZ} above the fiducial value of the observable at $M_{200m} = 4 \times 10^{14} h^{-1} M_{\odot}$ at a redshift $z = 0.25$, the bias in the MOR amplitude is ≈ 15 per cent unless tight external constraints on the intrinsic scatter are available. Accounting for the cosmic variance of cluster lensing is therefore necessary for upcoming cluster surveys that target the calibration of MORs for cluster cosmology.

ACKNOWLEDGEMENTS

We thank Gary Bernstein, Eduardo Rozo and Peter Schneider for helpful discussions and Martin Kilbinger for support with the software NICAIA used in this work.

This work was supported by SFB-Transregio 33 'The Dark Universe' by the Deutsche Forschungsgemeinschaft (DFG) and the DFG cluster of excellence 'Origin and Structure of the Universe'.

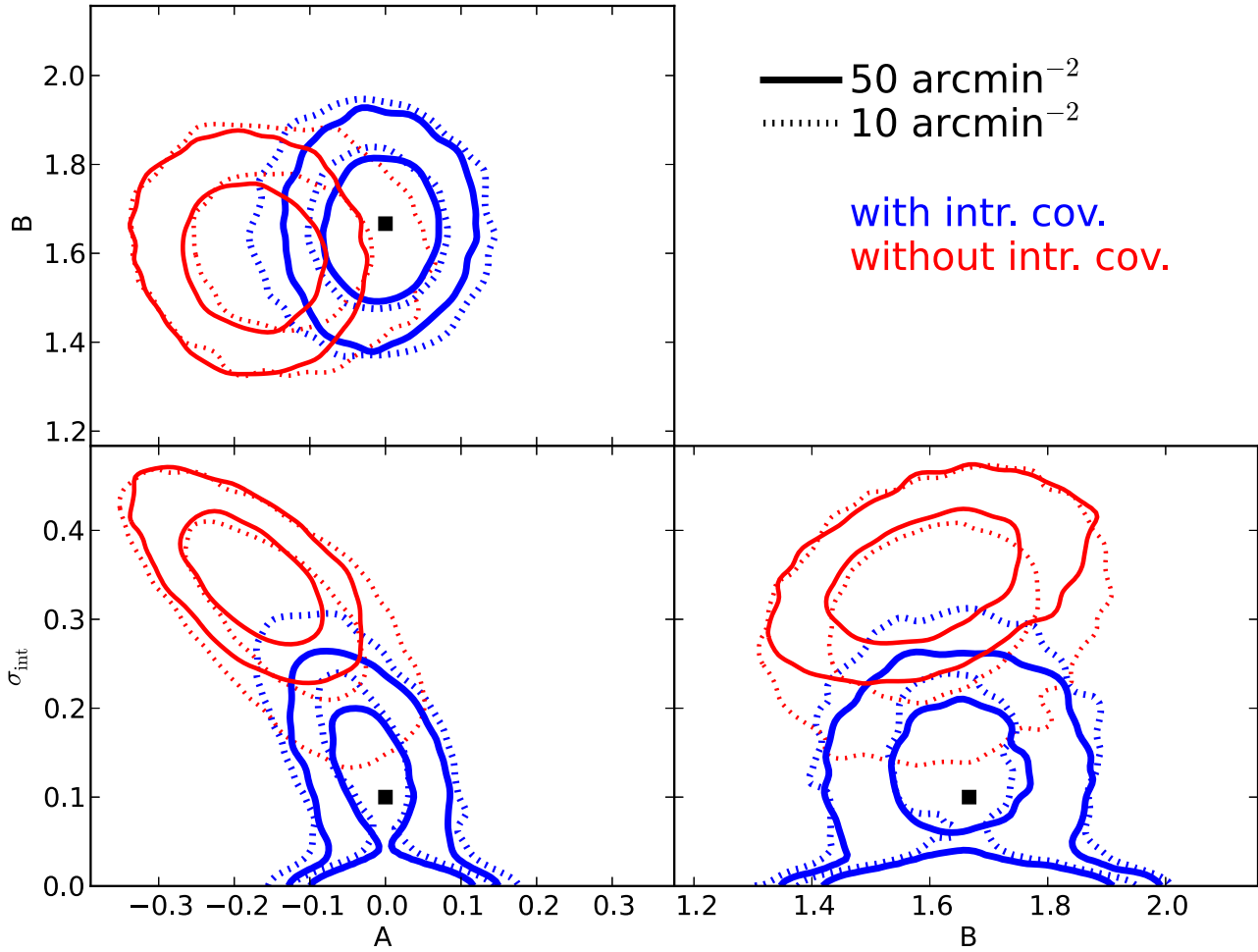


Figure 6. Distribution of maximum likelihood MOR parameters (true input indicated by black squares) estimated from a weak lensing follow-up of an observable limited sample. We show results for a shallow (dotted contours, $n = 10 \text{ arcmin}^{-2}$) and deeper (solid contours, $n = 50 \text{ arcmin}^{-2}$) lensing survey. Cluster profiles are simulated including intrinsic variation according to our model and the lensing likelihood for cluster mass is estimated including (blue, thick lines) and not including (red, thin lines) the intrinsic component of the covariance matrix in addition to shape noise and uncorrelated LSS. Inner and outer contours enclose 68% and 95% of data points, respectively. Results are for lenses at $z_l = 0.24533$.

REFERENCES

- Abramowitz M., Stegun I., 1965, Handbook of Mathematical Functions. Dover Publications
- Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
- Allen S. W., Schmidt R. W., Ebeling H., Fabian A. C., van Speybroeck L., 2004, Mon. Not. Roy. Astron. Soc., 353, 457
- Baltz E. A., Marshall P., Oguri M., 2009, JCAP, 1, 15
- Bartelmann M., 1996, A&A, 313, 697
- Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
- Benson B. A. et al., 2013, ApJ, 763, 147
- Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, Mon. Not. Roy. Astron. Soc., 376, 215
- Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
- Budzynski J. M., Kopev S. E., McCarthy I. G., McGee S. L., Belokurov V., 2012, Mon. Not. Roy. Astron. Soc., 423, 104
- Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, Mon. Not. Roy. Astron. Soc., 321, 559
- Corless V. L., King L. J., 2007, Mon. Not. Roy. Astron. Soc., 380, 149
- Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1
- Dietrich J. P. et al., 2014, Mon. Not. Roy. Astron. Soc., 443, 1713
- Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, Mon. Not. Roy. Astron. Soc., 390, L64
- Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
- Gottloeber S., Klypin A., 2008, ArXiv e-prints
- Gruen D., Bernstein G. M., Lam T. Y., Seitz S., 2011, Mon. Not. Roy. Astron. Soc., 416, 1392
- Gruen D. et al., 2014, Mon. Not. Roy. Astron. Soc., 442, 1507
- Hayashi E., White S. D. M., 2008, Mon. Not. Roy. Astron. Soc., 388, 2
- Hoekstra H., 2001, A&A, 370, 743

- Hoekstra H., 2003, *Mon. Not. Roy. Astron. Soc.*, 339, 1155
- Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, *Mon. Not. Roy. Astron. Soc.*, 427, 1298
- Hu W., 2003, *Phys. Rev. D*, 67, 081304
- Jiang F., van den Bosch F. C., 2014, *ArXiv e-prints*
- Jing Y. P., Suto Y., 2002, *ApJ*, 574, 538
- Kilbinger M. et al., 2009, *A&A*, 497, 677
- Komatsu E. et al., 2011, *ApJS*, 192, 18
- Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, *ApJS*, 111, 73
- Lima M., Hu W., 2005, *Phys. Rev. D*, 72, 043006
- Limber D. N., 1954, *ApJ*, 119, 655
- Majumdar S., Mohr J. J., 2004, *ApJ*, 613, 41
- Mana A., Giannantonio T., Weller J., Hoyle B., Hütsi G., Sartoris B., 2013, *Mon. Not. Roy. Astron. Soc.*, 434, 684
- Mandelbaum R., Seljak U., Baldauf T., Smith R. E., 2010, *Mon. Not. Roy. Astron. Soc.*, 405, 2078
- Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, *Mon. Not. Roy. Astron. Soc.*, 406, 1759
- Mantz A. B. et al., 2015, *Mon. Not. Roy. Astron. Soc.*, 446, 2205
- Marrone D. P. et al., 2012, *ApJ*, 754, 119
- Marrone D. P. et al., 2009, *ApJ*, 701, L114
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Oguri M., Hamana T., 2011, *Mon. Not. Roy. Astron. Soc.*, 414, 1851
- Oguri M., Lee J., Suto Y., 2003, *ApJ*, 599, 7
- Peacock J. A., Dodds S. J., 1996, *Mon. Not. R. Astron. Soc.*, 280, L19
- Planck Collaboration et al., 2013, *ArXiv e-prints*
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Rozo E. et al., 2010, *ApJ*, 708, 645
- Schneider P., van Waerbeke L., Jain B., Kruse G., 1998, *Mon. Not. Roy. Astron. Soc.*, 296, 873
- Sheth R. K., Tormen G., 1999, *Mon. Not. Roy. Astron. Soc.*, 308, 119
- Smith R. E. et al., 2003, *Mon. Not. Roy. Astron. Soc.*, 341, 1311
- Sunyaev R. A., Zel'dovich Y. B., 1972, *Comments on Astrophysics and Space Physics*, 4, 173
- Takada M., Jain B., 2003, *Mon. Not. Roy. Astron. Soc.*, 340, 580
- Taylor A., Joachimi B., Kitching T., 2013, *Mon. Not. Roy. Astron. Soc.*, 432, 1928
- Tegmark M., Taylor A. N., Heavens A. F., 1997, *ApJ*, 480, 22
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, *ApJ*, 688, 709
- Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, *ApJ*, 724, 878
- Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Coe D., Postman M., 2011, *ApJ*, 738, 41
- Vikhlinin A. et al., 2009a, *ApJ*, 692, 1033
- Vikhlinin A. et al., 2009b, *ApJ*, 692, 1060
- Vogeley M. S., Szalay A. S., 1996, *ApJ*, 465, 34
- von der Linden A. et al., 2012, *ArXiv e-prints*
- Wright C. O., Brainerd T. G., 2000, *ApJ*, 534, 34