Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **405,** 2078–2102 (2010) doi:10.1111/j.1365-2966.2010.16619.x

# Precision cluster mass determination from weak lensing

Rachel Mandelbaum,[1⋆] Uroš Seljak,[2,3] Tobias Baldauf[2] and Robert E. Smith[2]

[1]*Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*
[2]*Institute for Theoretical Physics, University of Zurich, Zurich, Switzerland*
[3]*Department of Physics, University of California, Berkeley, CA 94720, USA*

Accepted 2010 February 27. Received 2010 February 19; in original form 2009 November 25

**ABSTRACT**

Weak gravitational lensing has been used extensively in the past decade to constrain the masses of galaxy clusters, and is the most promising observational technique for providing the mass calibration necessary for precision cosmology with clusters. There are several challenges in estimating cluster masses, particularly (a) the sensitivity to astrophysical effects and observational systematics that modify the signal relative to the theoretical expectations, and (b) biases that can arise due to assumptions in the mass estimation method, such as the assumed radial profile of the cluster. All of these challenges are more problematic in the inner regions of the cluster, suggesting that their influence would ideally be suppressed for the purpose of mass estimation. However, at any given radius the differential surface density measured by lensing is sensitive to all mass within that radius, and the corrupted signal from the inner parts is spread out to all scales. We develop a new statistic $\Upsilon(R; R_0)$ that is ideal for estimation of cluster masses because it completely eliminates mass contributions below a chosen scale (which we suggest should be about 20 per cent of the virial radius), and thus reduces sensitivity to systematic and astrophysical effects. We use simulated and analytical profiles including shape noise to quantify systematic biases on the estimated masses for several standard methods of mass estimation, finding that these can lead to significant mass biases that range from 10 to over 50 per cent. The mass uncertainties when using the new statistic $\Upsilon(R; R_0)$ are reduced by up to a factor of 10 relative to the standard methods, while only moderately increasing the statistical errors. This new method of mass estimation will enable a higher level of precision in future science work with weak lensing mass estimates for galaxy clusters.

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

## 1 INTRODUCTION

Many scientific applications require robust measurements of the mass in galaxy clusters. One such application is the use of the dark matter halo mass function to constrain cosmological model parameters, including the amplitude of matter density perturbations, the average matter density and even the equation of state of dark energy (e.g., most recently, Rines, Diaferio & Natarajan 2007; Mantz et al. 2008; Vikhlinin et al. 2009; Rozo et al. 2010). Another example is validation and refinement of models of cluster formation and evolution, which predict relations between the more easily measured optical and X-ray emission, and the underlying dark matter halo (Kravtsov, Vikhlinin & Nagai 2006; Nagai, Kravtsov & Vikhlinin 2007; Zhang et al. 2008; Borgani & Kravtsov 2009). Currently, there are thousands of known clusters selected in various ways that can be used for these applications. Future surveys such as the dark energy survey (DES),[1] Pan-STARRS[2] and the Large Synoptic Survey Telescope (LSST)[3] will provide even larger and deeper samples that can be used for this purpose, requiring greater systematic robustness in the mass measures to complement the smaller statistical errors.

Many different methods have been used to measure the halo profile of clusters and thereby estimate their masses. Kinematic tracers such as satellite galaxies, in combination with a Jeans analysis or caustics analysis, can give information over a wide range of physical scales and halo masses. While the issues of relaxation, velocity bias, anisotropy of the orbits and interlopers need to be carefully addressed, recent results suggest a good agreement with theoretical

⋆E-mail: rmandelb@astro.princeton.edu

[1] http://www.darkenergysurvey.org/
[2] http://pan-starrs.ifa.hawaii.edu/public/
[3] http://www.lsst.org/lsst

predictions for the form of the density profile (Biviano & Girardi 2003; Katgert, Biviano & Mazure 2004; Rines et al. 2003; Diaferio, Geller & Rines 2005; Rines & Diaferio 2006; Salucci et al. 2007). Hydrostatic analyses of X-ray intensity profiles of clusters use X-ray intensity and temperature as a function of radius to reconstruct the density profile and estimate a halo mass. The advantage of thermal gas pressure being isotropic is partially lost due to the possible presence of other sources of pressure support, such as turbulence, cosmic rays or magnetic fields. These extra sources of pressure support cannot be strongly constrained for typical clusters with present X-ray data (Schuecker et al. 2004), but could modify the hydrostatic equilibrium and affect the conclusions of such analyses. Recent results are encouraging and are in a broad agreement with predictions, although most require concentrations that are higher than those predicted by a concordance cosmology (Vikhlinin et al. 2006; Buote et al. 2007; Schmidt & Allen 2007). While the above-mentioned systematic biases cannot be excluded, the small discrepancy could also be due to baryonic effects in the central regions, due to selection of relaxed clusters that may be more concentrated than average (Vikhlinin et al. 2006), or due to the fact that at a given X-ray flux limit, the more concentrated clusters near the limiting mass are more likely to be included in the sample (Fedeli et al. 2007).

Gravitational lensing is by definition sensitive to the total mass, and is therefore one of the most promising methods to measure the mass profile independent of the dynamical state of the clusters. Many previous weak lensing analyses have focused on individual clusters (e.g. Hoekstra 2007; Pedersen & Dahle 2007; Abate et al. 2009; Okabe et al. 2009). Measuring the matter distribution of individual clusters allows a comparison with the combined baryonic (light and gas) distribution on an individual basis, and so can constrain models that relate the two, such as Modified Newtonian Dynamics (MOND; Milgrom 1983) versus cold dark matter (CDM; Clowe et al. 2006). However, these measurements can be quite noisy for individual clusters. Stacking the signal from many clusters can ameliorate this problem, since shape noise and the signal due to correlated structures will be averaged out. Such a statistical approach is thus advantageous if one is to compare the observations to theoretical predictions, which also average over a large number of haloes in simulations. A final advantage of stacking is that it allows for the lensing measurement of lower-mass haloes, where individual detection is impossible due to their lower shears relative to more massive clusters. Individual high signal-to-noise ratio (S/N) cluster observations and those based on stacked analysis of many clusters are thus complementary to each other at the high-mass end, with the stacked analysis drastically increasing the available baseline in mass.

Extraction of cluster dark matter halo masses from the weak lensing signal is subject to a number of uncertainties, which we discuss in this paper in detail, including the ways in which the uncertainties differ for individual versus stacked cluster lensing analyses. In brief, these uncertainties are: (i) biased calibration of the lensing signal; (ii) modification of the lensing profile in the inner cluster regions due to accidental inclusion of cluster member galaxies in the source sample, intrinsic alignments of those galaxies, non-weak shear, magnification, baryonic effects that modify the initial cluster dark matter halo density profile and cluster centroiding errors; (iii) contributions to the lensing signal from non-virialized local structures and large-scale structure (LSS). Furthermore, parametric modelling of the mass requires the assumption of a form for the dark matter halo profile, which may differ from the intrinsic profile and/or have poorly constrained parameters. Non-parametric modelling, while not subject to this weakness, results in projected masses that must be converted to three-dimensional (3D) enclosed masses to be compared against the theory predictions, all of which are currently phrased in terms of 3D masses. We quantify the degree to which this conversion depends on assumptions about the density profile. Generally, we show the effects of many of these uncertainties on the estimated masses from cluster weak lensing analyses, both in the stacked and in individual cases, using parametric and non-parametric mass modelling.

Effects that modify the cluster density profile in the inner regions ($\lesssim 0.5\,h^{-1}$ Mpc) are particularly problematic given that the weak lensing signal $\Delta\Sigma(R)$ is sensitive to the density profile not just at a projected separation $R$, but also at all smaller separations. We propose a modified statistic, denoted by $\Upsilon(R; R_0)$, that removes the dependence on the projected density between $R = 0$ and $R = R_0$, with $R_0$ chosen to avoid scales with systematic uncertainties. The decrease in systematic errors that results from removing scales below $R_0$ comes at the expense of somewhat increased statistical errors. We explore the optimal choice of $R_0$, and quantify the degree to which our use of this new statistic to estimate cluster masses lessens systematic biases and increases statistical errors. Our tools for this investigation include simple, idealized cluster density profiles; more complex and realistic density profiles from $N$-body simulations and, finally, real cluster lensing data from the Sloan Digital Sky Survey (SDSS; York et al. 2000) that was previously analysed by Mandelbaum, Seljak & Hirata (2008a).

We begin in Section 2 with a discussion of the theoretical aspects of cluster-galaxy weak lensing, including a detailed discussion of the challenges of mass determination, and a summary of typical approaches to parametric and non-parametric mass estimation, with the introduction of a new statistic from which to derive parametric mass estimates. In Section 3, we describe the $N$-body simulations that we use to provide sample cluster density profiles. Section 4 has a description of the SDSS cluster lensing data we use to test for some of the effects that we find using the simulations. Results for both the theoretical profiles and the real data are presented in Section 5. We conclude with a discussion of our findings and their implications in Section 6.

## 2 THEORY

This section includes theoretical background related to cluster-galaxy weak lensing, modelling of cluster masses using lensing, and the new statistic that we propose is optimal for cluster mass estimation.

### 2.1 Standard lensing formalism

Cluster-galaxy weak lensing provides a simple way to probe the connection between clusters and matter via their cross-correlation function $\xi_{\rm cl,m}(\boldsymbol{r})$, defined as

$$\xi_{\rm cl,m}(\boldsymbol{r}) = \langle \delta_{\rm cl}(\boldsymbol{x})\delta^*_{\rm m}(\boldsymbol{x}+\boldsymbol{r})\rangle_{\boldsymbol{x}}, \quad (1)$$

where $\delta_{\rm cl}$ and $\delta_{\rm m}$ are overdensities of clusters and matter, respectively ($\delta_{\rm m} = \rho_{\rm m}/\overline{\rho}_{\rm m} - 1$). This cross-correlation can be related to the projected surface density

$$\Sigma(R) = \overline{\rho}\int [1 + \xi_{\rm cl,m}(\sqrt{R^2+\chi^2})]{\rm d}\chi, \quad (2)$$

where $\overline{\rho}$ is the mean matter density, $R$ is the transverse separation and $\chi$ is the line-of-sight direction over which we are projecting. Here, we ignore the line-of-sight window function, which is hundreds of mega-parsec broad and not relevant at cluster scales. For this paper,

we are primarily interested in the contribution to the cluster-matter cross-correlation from the cluster halo density profile $\rho_{\rm cl}$ itself, rather than from other structures, and hence

$$\Sigma(R) \approx \int_{-\infty}^{\infty} \rho_{\rm cl}(r = \sqrt{\chi^2 + R^2})\, {\rm d}\chi. \quad (3)$$

The surface density is then related to the observable quantity for lensing, called the differential surface density,

$$\Delta\Sigma(R) = \gamma_{\rm t}(R)\Sigma_c = \overline{\Sigma}(< R) - \Sigma(R), \quad (4)$$

where $\gamma_{\rm t}$ is the tangential shear (a weak but coherent distortion in the shapes of background galaxies) and $\Sigma_c$ is a geometric factor,

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{D_{\rm S}}{D_{\rm L} D_{\rm LS} (1 + z_{\rm L})^2}. \quad (5)$$

Here $D_{\rm L}$, $D_{\rm S}$ and $D_{\rm LS}$ are (physical) angular diameter distances to the lens, to the source and between the lens and source, respectively. In the second relation in equation (4), $\overline{\Sigma}(< R)$ is the average value of the surface density within some radius $R$,

$$\overline{\Sigma}(< R) = \frac{2}{R^2} \int_0^R R'\, \Sigma(R')\, {\rm d}R'. \quad (6)$$

The second equality of equation (4) is true in the weak lensing limit, for a matter distribution that is axisymmetric along the line of sight (which is naturally achieved by the procedure of stacking many clusters to determine their average lensing signal), or in the non-axisymmetric case, provided that $\Sigma$ is averaged azimuthally. For individual cluster analyses, profiles can be fit either using average shears in annuli or with full, 2D shear maps.

Unless otherwise noted, all computations assume a flat ΛCDM universe with matter density relative to the critical density $\Omega_{\rm m} = 0.25$ and $\Omega_\Lambda = 0.75$. Distances quoted for transverse lens-source separation are comoving (rather than physical) $h^{-1}$ Mpc, where the Hubble constant $H_0 = 100\,h\,{\rm km\ s^{-1}\,Mpc^{-1}}$. Likewise, the differential surface density $\Delta\Sigma$ is computed in comoving coordinates, equation (5), and the factor of $(1 + z_{\rm L})^{-2}$ arises due to our use of comoving coordinates.

### 2.2 Theoretical challenges in cluster mass modelling

In this section, we discuss theoretical challenges in cluster mass modelling. By 'theoretical' challenges, we refer to issues that cause the underlying cluster density profile (surface density $\Sigma$) to be unknown. This uncertainty in $\Sigma$ at a given scale $R$ is propagated to larger scales in $\Delta\Sigma(R)$ because of its dependence on $\overline{\Sigma}(< R)$ (equations 4 and 6).

#### *2.2.1 Unknown density profile*

When attempting to extract 3D enclosed masses from the projected lensing data, the unknown density profile may lead to a biased mass estimate. For example, even for the latest generation of simulations, the concentration parameter (defined more precisely below) of clusters remains somewhat uncertain, with differences at the level of 20 per cent at the high-mass end (Dolag et al. 2004; Neto et al. 2007; Zhao et al. 2009). The concentration parameter at a given mass is also affected by the assumed cosmological model, especially the amplitude of perturbations. For a given halo mass, the differences between the profiles increase towards the inner parts of the cluster, and if only those scales are used in parametric fits for mass estimation, this can result in a significant error on the halo mass. In this paper, we investigate bias due to unknown cluster concentration extensively, including the use of parametric mass estimators with assumptions about the form of the profile, and the use of non-parametric projected mass estimates that require an assumption about the profile to get a 3D enclosed mass.

#### *2.2.2 Baryonic effects*

The effect of baryons on the cluster mass distribution is unclear, but may be significant in the inner cluster regions (Blumenthal et al. 1986; Gnedin et al. 2004; Naab et al. 2007; Rudd, Zentner & Kravtsov 2008; Zentner, Rudd & Hu 2008; Barkana & Loeb 2010). Baryon cooling not only brings significant mass into the inner regions of the cluster, but may also redistribute the dark matter out to much larger scales than the scale of baryon cooling. These works suggest that the effect of baryons is to change the cluster matter profile in the inner regions in a way that roughly mimics a change in the halo concentration; however, the extent of this effect in reality, and the affected scales, is unknown.

#### *2.2.3 Offsets from minimum of cluster potential*

The cluster centre about which the lensing signal should be computed can be determined using a variety of methods. The most reliable approach is to use the peak in X-ray or Sunyaev–Zeldovich flux. For optically identified clusters, the usual method is to find the brightest cluster galaxy (BCG). The offsets from the true cluster centre arise due to two effects: (1) BCGs may be slightly perturbed from the minimum of the cluster potential well by some real physical effect, such as an infalling satellite, and (2) photometric redshift errors and/or limitations in the cluster detection technique (when detecting clusters using imaging data) may lead to the wrong galaxy being chosen as the BCG. This latter effect might occur, for example, with red-sequence cluster finding algorithms, in cases of BCGs with bluer colours (estimated to be ∼25 per cent of the BCG population in reality; Bildfell et al. 2008). As was discussed quantitatively in Johnston et al. (2007), the effect of BCG offsets on stacked cluster lensing data is to convolve the surface density $\Sigma(R)$ with some BCG offset distribution, which tends to suppress the lensing signal in the inner regions (similar, qualitatively, to the effect of the previous two systematic issues we have discussed). Consequently, fitted cluster masses and concentrations will be reduced due to these centroiding errors (Guzik & Seljak 2002; Yang et al. 2006). Note that while cluster centroiding errors arise due to observational limitations, we classify them as a theoretical issue because of their impact on $\Sigma(R)$ which leaks to larger scales in $\Delta\Sigma(R)$.

Studies comparing the BCG position to the cluster centre defined by either the X-ray intensity or the average satellite velocity have found that the typical displacement is about 2–3 per cent of the virial radius when the BCG is properly identified (van den Bosch et al. 2005; Koester et al. 2007a,b; Bildfell et al. 2008). The last of these studies finds that for about 10 per cent of BCGs, the displacement is >10 per cent of the virial radius. Another study that includes red galaxy photometric errors (i.e. both causes of offsets rather than just the first) finds that the median displacement is 10 per cent of the virial radius (Ho et al. 2009).

Because the real data we use as a test case uses the maxBCG lens sample, we focus in more detail on the issue of BCG offsets for that cluster catalogue. The maxBCG group uses mock catalogues to estimate the distribution of BCG offsets resulting from the use of their algorithm (Johnston et al. 2007). The accuracy of the distribution they find is quite sensitive to the details of how the simulations are

populated with galaxies. In brief, their result includes a richness-dependent fraction of misidentified BCGs (from 30 per cent at low richness to 20 per cent at high richness), and those that are misidentified have a Gaussian distribution of projected separation from the true cluster centre, with a scale radius of $0.42\,h^{-1}$ Mpc.

A full discussion of how these results from mocks compare with observations can be found in Mandelbaum et al. (2008a). To summarize, at high masses (more than a few times $10^{14}\,h^{-1}\,\mathrm{M}_{\odot}$), a comparison with X-rays (Koester et al. 2007b; Ho et al. 2009) suggests that the mocks may overestimate the fraction of offsets greater than 250 $h^{-1}$ kpc. However, the true level of offsets for the majority of the cluster catalogue is poorly constrained from the real data.

### 2.3 Observational challenges in cluster mass modelling

In this section, we discuss observational challenges in cluster mass modelling. We define 'observational' challenges as those that result in difficulty in properly measuring $\Delta\Sigma(R)$ for a given density profile $\Sigma(R)$.

#### *2.3.1 Lensing signal calibration*

The cluster-galaxy lensing signal overall calibration is an important issue for cluster mass estimates. The signal may be miscalibrated due to shape measurement systematics (e.g. Heymans et al. 2006; Massey et al. 2007; Bridle et al. 2009), unknown lens and/or source redshift distributions (e.g. Kleinheinrich et al. 2005; Mandelbaum et al. 2008b) and contamination of the 'source' sample by stars. The effect of miscalibration typically is to multiply the signal on all scales by a single multiplicative factor. We will investigate the effect of changes in lensing signal calibration on the estimated masses when fitting both parametrically and non-parametrically.

#### *2.3.2 Signal dilution due to cluster member galaxies*

In principle, in the absence of intrinsic alignments, contamination of the source sample by cluster member galaxies will dilute the cluster lensing signal, since the cluster member galaxies are not lensed. Thus, they suppress the cluster lensing signal, with the strongest effect towards the cluster centre where the member galaxies are most numerous. For stacked cluster lensing data, this effect may be effectively removed by cross-correlating random points with the source catalogue, and boosting the signal by the scale-dependent ratio of the weighted number of sources around the real clusters to that around the random points (Hirata et al. 2004; Sheldon et al. 2004; Mandelbaum et al. 2005a).

For measurements of individual cluster lenses, the best way around this problem is to use some colour-based criterion that removes the cluster member galaxies. Without multicolour imaging, contamination of the lensing signal can be several tens of per cent on a few hundred $h^{-1}$ kpc scales (Broadhurst et al. 2005; Limousin et al. 2007), and even with it, there may be residual dilution of the signal of approximately 10 per cent on those scales (Hoekstra 2007; Okabe et al. 2009). This scale-dependent suppression of the signal results in underestimation of the cluster mass and concentration.

#### *2.3.3 Intrinsic alignments*

Intrinsic alignments of galaxy shapes with the local tidal field can affect cluster lensing measurements when cluster member galaxies that are treated as sources actually have some mean alignment of their shapes radially towards the cluster centre. This effect, which leads to a suppression of the lensing signal that is worse at smaller transverse separations, has been detected observationally in several contexts (Agustsson & Brainerd 2006; Mandelbaum et al. 2006a; Faltenbacher et al. 2007; Hirata et al. 2007; Siverd, Ryden & Gaudi 2009). Its amplitude varies with cluster mass, member galaxy type and separation from the cluster centre.

The best way to avoid this effect is to remove cluster member galaxies from the source catalogue, but a perfect removal is often not possible, as described in Section 2.3.2 and references therein. When using a very large stacked sample, the amplitude of the effect may be roughly estimated using the estimated shear from the sample of galaxies that were chosen based on the colour–redshift relation to be cluster member galaxies. This test, however, is only possible with good colour information for the source galaxies. We defer a detailed discussion of the effects of intrinsic alignments on weak lensing cluster mass estimates to future work, but the sign is always to lower the signal (and therefore mass) in a way that is worse at smaller cluster-centric radius.

#### *2.3.4 Non-weak shear and magnification effects*

The measured weak lensing signal is not precisely the tangential shear $\gamma_{\rm t}$, but rather the reduced shear $g = \gamma_{\rm t}/(1-\kappa)$, where $\kappa = \Sigma/\Sigma_c$ is the convergence. For a typical cluster density profile, the difference between $g$ and $\gamma_{\rm t}$ is of the order of unity at the critical radius where $\kappa = 1$ (that depends on the redshift, but can be as large as $100\,h^{-1}$ kpc) reducing to a few per cent out to transverse separations of $\sim 500\,h^{-1}$ kpc, beyond which the assumption that $g \approx \gamma_{\rm t}$ is quite accurate. This distinction may be explicitly taken into account using parametric mass models (Mandelbaum et al. 2006b), but is typically ignored in non-parametric mass estimation (though since those estimates usually do not rely on the shear on small scales, this neglect is not necessarily a problem).

A related effect is magnification, which alters the source galaxy population by changing the measured fluxes and sizes[4] (Mandelbaum et al. 2005a; Schmidt et al. 2009). As a result, their redshift distribution may change, and the number density of sources near lens galaxies typically differs from that in the field. Furthermore, if a correction is made to the observed weak lensing signal for stacked clusters to account for the dilution due to cluster member galaxies included in the source sample, as suggested in Section 2.3.2, then this correction must be carried out by using the observed source number densities relative to that around random points. The boost factor is supposed to only correct for changes in source number density due to clustering (which introduces unlensed galaxies into the source sample). Since the number density of lensed galaxies may legitimately be altered by magnification, magnification can lead to incorrect boost factors. This effect may be accounted for using parametric mass modelling, provided that the properties of the source sample at the flux and apparent size limits is reasonably well understood (Mandelbaum et al. 2006b).

### 2.4 Summary of the challenges and how we model them

The challenges discussed in the previous two sections result in three types of changes in the lensing signal. One type of change is

[4] The change in apparent size may not be important for typical photometric data, but weak lensing measurements require imposition of an apparent size cut on the galaxies to ensure that they are well-resolved relative to the point spread function (PSF).

an elevation (suppression) of the lensing signal on small scales that changes sign at some value of transverse separation to become a suppression (elevation). For example, this change may result from an unknown dark matter concentration and baryonic effects. The second type of change is a uniform suppression or elevation of the lensing signal in the inner cluster regions, such that the lensing signal gradually reaches the expected value at and above some value of transverse separation. This change may result from cluster centroiding errors, dilution of the lensing signal due to cluster member galaxies and/or intrinsic alignments, non-weak shear, and magnification-induced errors in the source redshift distribution and number density. The exact functional forms for and magnitudes of these changes, and their characteristic scale radii, vary depending on the situation. However, we will use two models, one for each type of change. The final type of change in the lensing signal that we consider is a uniform calibration offset.

The profiles we use for our test cases include pure NFW profiles, and the cluster lensing signal observed in $N$-body simulations. We modify the concentrations of these test profiles, apply a model for the effects of cluster centroiding errors based on mock catalogues (Johnston et al. 2007) and rescale them all to mimic calibration offsets. However, we do not want to rely too much on our modelling of these effects (as concentrations and a specific centroid error model) being correct in detail. Thus, if cluster mass determination is to be robust, we need estimators that are as insensitive to these types of changes in cluster profile as possible. Note that a key feature of all three types of changes in profile is that they affect the inner cluster regions. This fact leads to the requirement that the small-scale information is suppressed, which will motivate a new statistic introduced in this paper.

In all cases, we use spherically symmetric profiles, as is appropriate for stacked cluster lensing analyses. The observed lensing profile is roughly equivalent to the spherical average of the underlying triaxial density profiles of the dark matter haloes, so that the cluster masses can be recovered to few per cent accuracy with mass estimation assuming spherical profiles (Mandelbaum et al. 2005b; Corless & King 2009). For individual cluster lensing estimates, however, there is an additional level of complication due to the assumption of a spherical profile: individual deviations in the form of the profile from the assumed form due to mergers, substructure (King, Schneider & Springel 2001), and deviations from a spherical shape (Clowe, De Lucia & King 2004; Corless & King 2007) can cause tens of per cent uncertainties in cluster mass model parameters. We do not attempt to estimate the uncertainties for individual cluster lensing analyses due to these effects, relying instead on previous work.

### 2.5 Signal due to other mass

The measured lensing signal is caused by the projected mass distribution around the cluster, and consequently it includes some contributions that are not part of the cluster halo, which will affect the mass estimates. In the case of stacked cluster lensing analyses, the average over these contributions from all clusters in the stack results in the so-called halo–halo term, which can be modelled simply using the cluster-matter cross-power spectrum as in, e.g., Seljak (2000) and Mandelbaum et al. (2005b). This term becomes dominant on several $h^{-1}$ Mpc scales. While here we use scales where this term is sub-dominant, we will consider the question of how the estimated masses may be biased if this term is not explicitly modelled but is instead neglected. This failure to model the halo–halo term should tend to pull the mass estimates upwards, since mass that is not part of the cluster mass distribution will be attributed to the cluster. Our approach is to simply use the cluster lensing signal from simulations without explicitly decomposing it into one- and halo–halo terms; thus, mass that is not part of the cluster mass distribution itself is implicitly included in our numerical predictions of the cluster lensing signal.

For individual cluster lensing analyses, the effect of matter that is not part of the cluster on the lensing signal is more complex, because unlike for stacked analyses, no averaging process occurs over the structures around many clusters. As a result, local non-virialized structure (Metzler et al. 1999; Metzler, White & Loken 2001) and LSS (Hoekstra 2001, 2003; Dodelson 2004) can appear in the cluster lensing signal on all scales, not just large scales, causing both an average bias and significant scatter in the mass estimates. A recent numerical study of LSS projection effects on weak lensing cluster counts (Marian, Smith & Bernstein 2009) has shown that, whilst there is scatter and bias in the $M_{\rm 2D}$–$M_{\rm 3D}$ relation, the utility for such data to constrain cosmological parameters through the mass function is not impaired. Moreover, if one uses carefully constructed aperture mass shear filters, then the bias arising from 'correlated' LSS can be reduced to the per cent level (Marian, Smith & Bernstein 2010). However, the impact of 'chance' projections along the line-of-sight on the mass estimates is still relatively poorly quantified. While we use simulations to assess the effect of the halo–halo term on stacked cluster analyses that neglect it, a detailed treatment of this issue for individual cluster lensing analyses is beyond the scope of this paper.

### 2.6 Parametric modelling of cluster masses

In principle, we can model the cluster-galaxy weak lensing signal as a sum of two terms, the first due to the BCG stellar component, only important on scales below $\sim 100\,h^{-1}$ kpc, and the second due to the dark matter halo. Typically, the halo is modelled using the broken power-law NFW density profile (Navarro, Frenk & White 1996):

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

where the scale radius $r_{\rm s}$ is the scale at which the logarithmic slope, $\mathrm{d}\ln\rho/\mathrm{d}\ln r$, is equal to $-2$. While this approach to cluster mass estimation is fairly standard, recent work (Merritt et al. 2006; Gao et al. 2008) suggests that the Einasto profile (Einasto 1965),

$$\rho(r) = \rho_{\rm s} e^{(-2/\alpha)[(r/r_{\rm s})^{\alpha} - 1]}, \tag{8}$$

(where $\alpha$ has a weak mass dependence with a value around 0.15) may better describe the dark matter halo profiles. We note here that on the scales we use for modelling in this work, the two profiles agree to within a few per cent. Thus, the NFW profile is sufficient for our purposes.

It is convenient to parametrize the NFW profile by two parameters, the concentration $c_{200b} = r_{200b}/r_{\rm s}$ and the virial mass $M_{200b}$. The virial radius $r_{200b}$ and $\rho_{\rm s}$ can be related to $M_{200b}$ via consistency relations. The first is that the virial radius is defined such that the average density within it is $200\overline{\rho}$:

$$M_{200b} = \frac{4\pi}{3} r_{200b}^3\,(200\overline{\rho})\,. \tag{9}$$

The second relation, used to determine $\rho_{\rm s}$ from $M_{200b}$ and $c_{200b}$, is simply that the volume integral of the density profile out to the virial radius must equal the virial mass (though when computing the lensing signal, we do not truncate the profiles beyond $r_{200b}$). The

NFW concentration is a weakly decreasing function of halo mass, with a typical dependence as

$$c_{200b} = \frac{c_0}{1+z}\left(\frac{M}{M_0}\right)^{-\beta}, \quad (10)$$

with $\beta \sim 0.1$ (Bullock et al. 2001; Eke, Navarro & Steinmetz 2001; Neto et al. 2007), making this profile a one-parameter family of profiles. The normalization of equation (10) depends on the non-linear mass (and hence cosmology), but for the typical range of models, one expects $c_{200b} = 5$–$8$ at $M_0 = 10^{14}\,h^{-1}\,\mathrm{M}_\odot$. Some work (Zhao et al. 2003, 2009; Neto et al. 2007) suggests that this mass dependence levels off to a constant concentration above some high value of mass. The precise value remains somewhat controversial, with $c_{200b} \sim 5$–$6$ in Neto et al. (2007) and Zhao et al. (2009), but some other analyses suggest a significantly higher value around $c_{200b} \sim 7$–$8$ at $z = 0$ (J. Tinker, private communication). In addition, if one applies the typical concentration–mass relation assumed in, e.g., Hoekstra (2007) to very high mass clusters, one finds very small concentration values, e.g., $c_{200b} \sim 4$ at $M \sim 10^{15}\,\mathrm{M}_\odot$. In this paper, we will assess the effect of assuming the wrong concentration value in parametric mass estimates, taking $c_{200b} = 4$–$7$ as the plausible range given the current level of uncertainties.

While we demonstrate our cluster mass estimation procedure using stacked lensing data for which a spherical model is appropriate, one can easily apply the same techniques using lensing data for individual clusters. In that case, parametric profile fitting may use a circular average of the shear profile, or a full shear map with the inclusion of a projected ellipticity and position angle among the fit parameters. Here, for simplicity, we assume the former.

There are two significant practical differences between stacked survey data versus data for individual clusters: first, survey data are typically available to large transverse separations, whereas data for individual clusters are limited by the field of view (FOV) of the telescope used for the observations. For typical cluster redshifts in cluster lensing analyses, $2\,h^{-1}$ Mpc is a typical maximum radius to which the lensing signal can be measured. Secondly, stacked lensing data typically yield a concentration that is around the mean concentration of the sample used for the stacking (Mandelbaum et al. 2005b). As a result, the main uncertainty in what concentration to assume for parametric mass estimation comes from differences between the published concentration–mass relations from $N$-body simulations, the uncertainty in cosmological parameters and the uncertainty about how baryonic cooling may have changed the halo concentration. In contrast, for individual cluster data the concentration is likely to vary significantly from cluster to cluster due to the intrinsic lognormal concentration distribution at fixed mass; this variation of $\sim$0.15 dex (Bullock et al. 2001) is non-negligible compared to the sources of systematic uncertainty about halo concentration.

In this paper, when studying the effects of parametric models on fits for the mass, we choose to fix the halo concentration as in some individual analyses, such as Hoekstra (2007), and some stacked analyses, such as Reyes et al. (2008). Other works have fit simultaneously for a concentration and a mass (e.g. Mandelbaum et al. 2008a; Okabe et al. 2009). In the latter case, there is no concern about biases in the mass due to assumption of the wrong concentration, but small biases may remain due to deviations of the profile from NFW, and there is a loss of statistical power so that the mass estimates become noisier. Furthermore, if there are systematic errors in the data (such as centroiding errors or intrinsic alignments) that do not perfectly mimic a change in concentration, those analyses may still find a biased result for the mass. For the most part, we wish to characterize systematic biases that can occur when the concentration is fixed, but we will mention the effects of allowing it to vary.

Finally, we note that parametric mass estimation lends itself easily to corrections for effects such as non-weak shear and magnification bias (Mandelbaum et al. 2006b). These effects can simply be incorporated into the model before comparing with the data.

### 2.7 Non-parametric modelling of cluster masses

Another common approach to cluster mass estimation is the non-parametric aperture mass statistic. In this work, we present tests of the $\zeta_c$ statistic (Clowe et al. 1998), which is related to the $\zeta$ statistic of Fahlman et al. (1994). $\zeta_c$ has been used in several recent cluster modelling papers, including Hoekstra (2007) and Okabe et al. (2009). This statistic is defined using three radii: the first, $R_1$, is the transverse separation within which we wish to estimate the projected mass and the second and third, $R_{o1}$ and $R_{o2}$, define an outer annulus. $\zeta_c$ is equal to the mean surface density within $R_1$ relative to that in the outer annulus:

$$\zeta_c(R_1) = \overline{\kappa}(R < R_1) - \overline{\kappa}(R_{o1} < R < R_{o2}), \quad (11)$$

where $\kappa$ is the scaled surface density or convergence, $\kappa = \Sigma/\Sigma_c$. The aperture mass statistic can be measured using the observed shear $\gamma_{\rm t}(R)$ using

$$\zeta_c(R_1) = 2\int_{R_1}^{R_{o1}} \mathrm{d}\ln R\,\gamma_{\rm t}(R) + \frac{2}{1-(R_{o1}/R_{o2})^2}\int_{R_{o1}}^{R_{o2}} \mathrm{d}\ln R\,\gamma_{\rm t}(R). \quad (12)$$

The 2D (projected or cylindrical) mass $M_{\rm 2D}(R_1)$ within $R_1$ can be estimated from $\zeta_c$ by

$$M_{\rm 2D}(R_1) = \pi R_1^2\,\Sigma_c\,\zeta_c(R_1). \quad (13)$$

Typically, $R_1$ is chosen to be either a fixed physical scale or a spherical overdensity radius (determined either using a parametric model to estimate the appropriate radius or iteratively using the aperture mass estimate from the data). Various approaches are taken to the second term in equation (12), which should ideally be subdominant to the first, given the scaling of shear with radius. For example, Hoekstra (2007) use the parametric fits to an NFW profile with fixed concentration parameter to estimate the amplitude of the second term. In contrast, Okabe et al. (2009) neglect it, after choosing $R_{o1}$ to be 10–15 arcmin, depending on where in the cluster field there appeared to be significant structures that they wished to avoid.[5] For their typical cluster redshifts, this choice corresponds to roughly 2–2.5 comoving $h^{-1}$ Mpc in transverse separation. We will consider the effect of both approaches in our tests below.

The aperture mass statistic is often used because of its insensitivity to the details of the cluster mass profile. Furthermore, because it estimates the mass within $R_1$ using the shear on scales larger than $R_1$, it is not very sensitive to systematics that affect the signal in the inner parts, such as contamination by cluster member galaxies, intrinsic alignments and centroiding errors. This decreased sensitivity to systematics comes at a price, however: as shown in equation (12), the determination of $\zeta_c$ requires integration over the measured shear profile in logarithmic annular bins, which can often be quite

[5] M. Takada, private communication.

noisy. Our tests will help quantify the extent to which this noisiness increases the statistical error on the mass estimates relative to parametric modelling.

An additional disadvantage to the use of the aperture mass statistic and the derived $M_{2D}$ is that cosmological analyses using the mass function, and any comparison against X-ray-derived masses, requires the use of a 3D (enclosed) mass, $M_{3D}$. The conversion from $M_{2D}$ to $M_{3D}$ requires the assumption of a profile, such as NFW (for which a concentration parameter must either be assumed or be derived from parametric fits). This conversion factor may be derived analytically from expressions for the enclosed $M_{2D}$ and $M_{3D}$ as in Wright & Brainerd (2000). Okabe et al. (2009) show that the conversion factor only weakly depends on the concentration, but for analyses that seek to determine the mass to 10 per cent, this dependence on concentration is still important. A way of avoiding this necessity would be to determine the mass function in terms of projected masses in the simulations, rather than the typical practise of using $M_{3D}$ within some spherical overdensity; however, given that this has not yet been done, we also test the effect of this $M_{2D}$ to $M_{3D}$ conversion.

### 2.8 New statistic for mass estimation

As noted previously, one complication in parametric modelling of the lensing signal $\Delta\Sigma(R)$ is the sensitivity to the mass profile on small scales, which is particularly prone to theoretical and observational uncertainty. We wish to avoid sensitivity to small scales, which comes from the first term on the right-hand side of equation (4), via $\overline{\Sigma}(< R)$ (defined in equation 6).

Thus, we must turn the lower limit of integration in equation (6) from $R = 0$ to some larger scale that is not strongly affected by small-scale systematics such as intrinsic alignments and centroiding errors. We refer to this new minimum scale as $R_0$, and achieve our goal by defining the annular differential surface density (ADSD)

$$\Upsilon(R;R_0) = \Delta\Sigma(R) - \Delta\Sigma(R_0)\left(\frac{R_0}{R}\right)^2 = \frac{2}{R^2}\int_{R_0}^{R}\Sigma(R')R'\,\mathrm{d}R' - \Sigma(R) + \Sigma(R_0)\left(\frac{R_0}{R}\right)^2. \quad (14)$$

As shown in equation (14), by subtracting off $\Delta\Sigma(R_0)(R_0/R)^2$ from the observed lensing signal, we achieve our goal of removing the sensitivity to scales below $R_0$. The resulting robustness of the analysis to systematic errors comes at the expense of introducing slight ($\sim$10 per cent level) anti-correlations between the signal around $R_0$ and the signal at larger scales, plus increased statistical errors.

For some of this paper, we model theoretical and observational uncertainties in $\Delta\Sigma$ as changes in the NFW concentration parameter. However, as already discussed, some systematics are manifested in different ways (e.g. centroiding errors) that must be modelled rather differently. If one truly believes that unknown concentration is the dominant systematic uncertainty, then the simplest solution would be to fit $\Delta\Sigma$ to an NFW profile and then marginalize over the concentration. Since we do not believe that this procedure is adequate for all theoretical and observational systematics, parametric modelling of $\Upsilon(R;R_0)$ to remove all small-scale information is a better solution that will give more accurate mass estimates. For some systematics, we will see in Section 5 that we do not, in general, have to select $R_0$ to be completely above the affected scales, because the errors in $\Upsilon(R;R_0)$ change sign and thus nearly cancel out of the mass estimation, despite contributing to $\Delta\Sigma(R)$ with the same sign at all scales.

In practice, to use the ADSD $\Upsilon(R;R_0)$ we must estimate $\Delta\Sigma(R_0)$ from the data themselves. In this work, we have tried two methods of doing so based on fits to the following functional form for $\Delta\Sigma$ in the neighbourhood of $R_0$:

$$\Delta\Sigma(R) = \Delta\Sigma(R_0)\left(\frac{R}{R_0}\right)^{p+q(R/R_0)}. \quad (15)$$

In the simpler method, $q = 0$, whereas in the more complex method it is a free parameter in the fit (which generally allows for a better fit to broken power-law profiles such as NFW, but also increases the statistical errors on the mass). We primarily present results of the latter procedure, but discuss the trade-offs between the two in Section 5.

Finally, we note that the ADSD $\Upsilon(R;R_0)$ is well-suited not only to estimating cluster masses, but also to cosmological studies, where the choice of $R_0$ to be outside of the host dark matter halo virial radius allows contributions to the lensing signal from small-scale information to be suppressed (Baldauf et al. 2010).

## 3 SIMULATIONS

To obtain realistic cluster lensing profiles for our tests of mass inference methods, we use the 'zHORIZON' simulations, a suite of 30 pure dissipationless dark matter simulations of the $\Lambda$CDM cosmology (Smith 2009). Each simulation models the dark matter density field in a box of length $L = 1500\,h^{-1}$ Mpc, using $N_p = 750^3$ dark matter particles with a mass of $M_{dm} = 5.55 \times 10^{11}\,h^{-1}\,M_\odot$. The cosmological parameters for the simulations in Table 1 are inspired by the results of the *WMAP* cosmic microwave background experiment (Spergel et al. 2003, 2007). For this work, we use eight of the 30 simulations, and probe a volume of $27\,h^{-3}$ Gpc$^3$ at redshift $z = 0.23$. The initial conditions were set up at redshift $z = 50$ using the 2LPT code (Scoccimarro 1998). The evolution of the $N_p$ equal-mass particles under gravity was then followed using the publicly available *N*-body code GADGET-II (Springel 2005). Finally, gravitationally bound structures were identified in each simulation snapshot using a Friends-of-Friends (FoF; Davis et al. 1985) algorithm with linking length of 0.2 times the mean inter-particle spacing. We rejected haloes containing fewer than 20 particles, and identified the potential minimum of the particle distribution associated with the halo as the halo centre. We note that using the FoF halo finder might cause some problems with the halo profile, since FoF tends to link together nearby haloes. In total, we identify haloes in the mass range $1.1 \times 10^{13}\,h^{-1}\,M_\odot \leq M_{200b} \leq 4 \times 10^{15}\,h^{-1}\,M_\odot$.

**Table 1.** Cosmological parameters adopted for the simulations: matter density relative to the critical density, dark energy density parameter, dimensionless Hubble parameter, matter power spectrum normalization, primordial power spectrum slope and dark energy equation of state $p = w\rho$.

| $\Omega_m$ | $\Omega_\Lambda$ | $h$ | $\sigma_8$ | $n_s$ | $w$ |
|---|---|---|---|---|---|
| 0.25 | 0.75 | 0.7 | 0.8 | 1.0 | $-1$ |

### 3.1 Calculation of the signal

We calculate the spherically averaged correlation function in the simulations using direct counts of mass particles in spherical shells about the halo centres of the cluster stack, $N_{\rm cl,m}(r_i)$. Our estimator for the correlation function is

$$\xi_{\rm cl,m}(r_i) = \frac{N_{\rm cl,m}(r_i)}{N^{\rm (rand)}_{\rm cl,m}(r_i)} - 1, \tag{16}$$

where $N^{\rm (rand)}_{\rm cl,m}(r_i) = N_{\rm cl} N_{\rm m} V_{\rm shell}/V_{\rm box}$ is the expected number of pairs for a purely random sample (for $N_{\rm cl}$ and $N_{\rm m}$ defined as the total number of clusters and matter particles in the box, respectively), and $V_{\rm shell} = 4\pi(r^3_{i+1} - r^3_i)/3$ is the volume of the spherical shell at $r_i$. To reduce the computational cost of this calculation, we dilute the dark matter density field by a factor of 24, using only $20 \times 10^6$ dark matter particles. We have confirmed the convergence of this procedure.

### 3.2 Correction for resolution effects

Despite the large dynamical range of our simulations, our resolution is still limited on small scales. The force softening length was set to $70\,h^{-1}$ kpc, so our results may not be reliable for $r \lesssim 200\,h^{-1}$ kpc. This resolution problem limits our ability to predict the excess surface mass density $\Delta\Sigma(R)$ on small scales, since this quantity is affected by the average over the correlation function on even smaller scales. Therefore, to correct for this problem, we continue the profile towards small scales using the NFW profile as follows:

$$1 + \xi^{\rm (stitch)}_{\rm cl,m}(r) = \begin{cases} \rho^{\rm (NFW)}_{\rm cl,m}(r)/\bar{\rho}, & \text{for } r < r_{\rm stitch} \\ \rho^{\rm (sim)}_{\rm cl,m}(r)/\bar{\rho}, & \text{for } r \geq r_{\rm stitch}. \end{cases} \tag{17}$$

We used the combinations ($r_{\rm stitch} = 0.2\,h^{-1}$ Mpc, $c_{200b} = 5$) and ($r_{\rm stitch} = 1.0\,h^{-1}$ Mpc, $c_{200b} = 7$).

Virial radii and masses are calculated by imposing the constraint

$$\frac{3}{r^3_{200b}\delta} \int_0^{r_{\rm vir}} (r')^2\,{\rm d}r' [1 + \xi_{\rm cl,m}(r')] = \frac{3M_{200b}}{4\pi r^3_{200b}\bar{\rho}\delta} = 1. \tag{18}$$

The overdensity of haloes is assumed to be $\delta = 200$ times the background density. The profile is then spline fitted and integrated along the line of sight over separations $-50 \leq \chi \leq 50\,h^{-1}$ Mpc from the cluster.

## 4 DATA

The SDSS (York et al. 2000) imaged roughly $\pi$ steradians of the sky, and followed up approximately one million of the detected objects spectroscopically (Eisenstein et al. 2001; Richards et al. 2002; Strauss et al. 2002). The imaging was carried out by drift-scanning the sky in photometric conditions (Hogg et al. 2001; Ivezić et al. 2004), in five bands (*ugriz*) (Fukugita et al. 1996; Smith et al. 2002) using a specially designed wide-field camera (Gunn et al. 1998). These imaging data were used to create the cluster and source catalogues that we use in this paper. All of the data were processed by completely automated pipelines that detect and measure photometric properties of objects, and astrometrically calibrate the data (Lupton et al. 2001; Pier et al. 2003; Tucker et al. 2006). The SDSS was completed with its seventh data release (Stoughton et al. 2002; Abazajian et al. 2003, 2004, 2005; Finkbeiner et al. 2004; Adelman-McCarthy et al. 2006, 2007, 2008; Abazajian et al. 2009).

In this paper, the only data that we use are the maxBCG cluster lensing data previously analysed in Mandelbaum et al. (2008a). Because the data were described there in detail, here we simply give a brief summary.

The parent sample from which our lens samples were derived consists of 13 823 MaxBCG clusters (Koester et al. 2007a,b), identified by the concentration of galaxies in colour-position space using the well-known red galaxy colour–redshift relation (Gladders & Yee 2000). The sample is based on 7500 deg$^2$ of imaging data in SDSS. There is a tight mass–richness relation that has been established using dynamical information (Becker et al. 2007) and weak lensing (Johnston et al. 2007; Mandelbaum et al. 2008a; Reyes et al. 2008) across a broad range of halo mass. The redshift range of the maxBCG sample is $0.1 < z < 0.3$; within these redshift limits, the sample is approximately volume-limited with a number density of $3 \times 10^{-5}\ (h/\,{\rm Mpc})^3$, except for a tendency towards higher number density at the lower end of this redshift range (Reyes et al. 2008). In this paper, we use scaled richness in red galaxies above $0.4L_*$ within $R_{200}$, known as $N_{200}$, as a primary tracer of halo mass. For the data in Mandelbaum et al. (2008a) that we use here, the richness range is $12 \leq N_{200} \leq 79$ divided into six bins ($12 \leq N_{200} \leq 13$, $14 \leq N_{200} \leq 19$, $20 \leq N_{200} \leq 28$, $29 \leq N_{200} \leq 39$, $40 \leq N_{200} \leq 54$, and $55 \leq N_{200} \leq 79$).

The source sample with estimates of galaxy shapes is the same as that originally described in Mandelbaum et al. (2005a). This source sample has over 30 million galaxies from the SDSS imaging data with *r*-band model magnitude brighter than 21.8, with shape measurements obtained using the REGLENS pipeline, including PSF correction done via re-Gaussianization (Hirata & Seljak 2003) and with cuts designed to avoid various shear calibration biases. The overall calibration uncertainty due to all systematics was originally estimated to be eight per cent (Mandelbaum et al. 2005a), though the redshift calibration component of this systematic error budget has recently been decreased due to the availability of more spectroscopic data (Mandelbaum et al. 2008b). The absolute mass calibration is not a critical issue for this paper, in which we study the changes in estimated mass for a given observed signal when using different estimation procedures.

## 5 RESULTS

### 5.1 Purely analytical profiles

In this section, we add realistic levels of noise to pure NFW profiles to create simplified mock cluster density profiles. The profiles that we use have $\log_{10}[hM_{200b}/{\rm M}_\odot] = 14.0$ and 14.8, with $c_{200b} = 4$ and $c_{200b} = 7$ (see properties of these profiles listed in Table 2). Using these profiles, we can test the dependence of parametric and non-parametric modelling on assumptions about the NFW concentration parameter. We caution that these profiles cannot be used to test for the effects of deviations from an NFW profile on the extracted masses when fitting assuming NFW profiles, or for the effects of LSS contributions to the lensing signal. These are discussed in the next section.

These values of concentration were selected as the extremes of the variation allowed with cosmology, and with the various determinations of the concentration–mass relation in the literature, including results suggesting that the concentration stops decreasing with mass at the high-mass end (Zhao et al. 2003, 2009; Neto et al. 2007). In addition, we consider that baryonic effects may increase the concentration of the dark matter profile (for an extreme example, see Rudd et al. 2008). Furthermore, for individual cluster lensing analyses, we must consider the fact that dark matter haloes exhibit a large scatter in concentration (0.15 dex; Bullock et al. 2001), so

**Table 2.** Properties of cluster lensing profiles, both analytical (pure NFW) and those from *N*-body simulations. We show the mean number density of the sample for the mass-selected samples from *N*-body simulations; the virial mass and radius $M_{200b}$ and $r_{200b}$ (exact value for the pure NFW profiles, and the ensemble mean for the samples from *N*-body simulations); the analytical profiles used for resolution corrections of the *N*-body simulations and the best-fitting NFW profiles when fitting the simulation lensing signals $\Delta\Sigma(R)$ for scales $0.2 \leq R \leq 2\,h^{-1}$ Mpc.

| $\overline{n}$ $[10^{-6}\ (h/\mathrm{Mpc})^3]$ | $M_{200b}$ $[10^{14}\,h^{-1}\,\mathrm{M}_\odot]$ | $r_{200b}$ $[\,h^{-1}\,\mathrm{Mpc}]$ | Stitching | Best-fitting NFW parameters $M_{200b}$ $[10^{14}\,h^{-1}\,\mathrm{M}_\odot]$ | $c_{200b}$ |
|---|---|---|---|---|---|
| | | | Pure NFW profiles, $c_{200b} = 4$ and 7 | | |
| - | 1.0 | 1.2 | - | - | - |
| - | 6.3 | 2.2 | - | - | - |
| | | | *N*-body simulation profiles | | |
| 0.25 | 7.9 | 2.4 | $c_{200b} = 7$ at $1\,h^{-1}$ Mpc | 7.9 | 6.6 |
| 0.25 | 7.5 | 2.3 | $c_{200b} = 5$ at $0.2\,h^{-1}$ Mpc | 7.8 | 4.6 |
| 2 | 4.2 | 1.9 | $c_{200b} = 7$ at $1\,h^{-1}$ Mpc | 4.2 | 6.6 |
| 2 | 4.0 | 1.9 | $c_{200b} = 5$ at $0.2\,h^{-1}$ Mpc | 4.3 | 4.6 |
| 16 | 1.6 | 1.4 | $c_{200b} = 7$ at $1\,h^{-1}$ Mpc | 1.7 | 6.5 |
| 16 | 1.6 | 1.4 | $c_{200b} = 5$ at $0.2\,h^{-1}$ Mpc | 1.7 | 4.5 |

the variation we have used is not as extreme in this case as it may be for a stacked cluster analysis. The change in concentration from 4 to 7 is less than $2\sigma$ of this intrinsic scatter.

To generate the profiles, we begin with the cluster halo density profile $\rho_{\rm cl}(r)$, which is defined in very narrow logarithmic (3D) radial bins. We then numerically integrate this profile along the line-of-sight, for comoving line-of-sight separations $|\chi| \leq r_{200b}$, to define $\Sigma(R)$ in very narrow logarithmic bins in transverse separation $R$. We calculate $\overline{\Sigma}(< R)$ by converting the integral in equation (6) to a summation. $\Delta\Sigma(R)$ can then be computed directly from $\overline{\Sigma}(<R) - \Sigma$.

To make this theoretical signal, defined in very narrow bins without any noise, look like an observed signal, we then do the following. First, we use a spline to determine the values of $\Delta\Sigma$ at the centre of the bins in $R$ used to calculate the real signal for maxBCG clusters in Mandelbaum et al. (2008a). Secondly, we choose a cluster richness subsample from that paper with roughly comparable mass to the theoretical signal we are using. We estimate a power-law function for the (bootstrap-determined) errors as a function of radius from our selected cluster subsample to avoid the influence of any noise in the determination of the covariances. We use this power law to assign a variance to the theoretical signal as a function of transverse separation. Finally, since the signal in the different radial bins was found to be nearly uncorrelated for all scales used in that paper, we add noise to our theoretical signals using a Gaussian distribution with a diagonal covariance matrix. This procedure was performed 1000 times to generate 1000 realizations of the lensing data. For context, the input level of noise is typically sufficient to achieve $\sim$20 per cent statistical uncertainty on the best-fitting masses at the $1\sigma$ level, when using $\Delta\Sigma$ with $R < 4\,h^{-1}$ Mpc to fit for the mass.

The input lensing signals $\Delta\Sigma(R)$ and $\Upsilon(R;R_0)$ (before the addition of noise) with several $R_0$ values are shown in Fig. 1 for the higher mass value, $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.8$. Since we will also test the effect of centroiding errors, which were discussed in detail in Section 2.2, we apply the offset model from Johnston et al. (2007). For offset fractions, we have chosen 20 per cent for this mass scale; for $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.0$, we will use 35 per cent (roughly in accordance with the trends with richness in that paper).

As expected, $\Delta\Sigma(R)$ for $c_{200b} = 7$ is higher than that for $c_{200b} = 4$ on small scales; the radius at which they cross over is relatively large because $\Delta\Sigma(R)$ includes information from $\Sigma(R)$ for small $R$. For the 3D $\rho(r)$, the cross-over radius is within the virial radius by necessity, since the masses are the same. As we increase $R_0$ in $\Upsilon(R;R_0)$, the trend going from $c_{200b} = 4$ to $c_{200b} = 7$ gets less pronounced, because even though $\Delta\Sigma(R)$ is larger on small scales for $c_{200b} = 7$, that also means the value that is subtracted off to obtain $\Upsilon(R;R_0)$ is larger. Thus, by the time we reach $R_0 = 1\,h^{-1}$ Mpc, $\Upsilon(R;R_0)$ is actually higher for $c_{200b} = 4$ than for $c_{200b} = 7$ for all $R > R_0$.

As shown in the bottom panel, the effect of centroiding errors is quite pronounced on $\Delta\Sigma(R)$. The characteristic scale of the offsets is $0.42\,h^{-1}$ Mpc, and the signal is noticeably suppressed out to three times this scale. The use of $\Upsilon(R;R_0)$ ameliorates this effect, and it even gets reversed for larger $R_0$, similar to what happens with the different concentration values. While for $\Delta\Sigma(R)$, the offsets cause suppression of the signal for all affected scales, for $\Upsilon(R;R_0)$, the signal is suppressed on smaller scales and elevated on larger scales, which suggests that biases in parametric mass modelling due to these offsets may be smaller because the small- and large-scale changes in sign may cancel out.

### 5.1.1 *Parametric modelling*

In this section, we begin by fitting the pure NFW lensing signals for $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.8$ to pure NFW profiles. This procedure allows us to assess the systematic uncertainty due to the assumption of a fixed concentration when using various parametric fit procedures. For each noise realization, we attempted to determine a mass using several fitting procedures:

(i) Assuming an NFW profile with $c_{200b} = 4$ and $c_{200b} = 7$.

(ii) Using $\Delta\Sigma(R)$ with minimum fit radii ($R_{\rm min}$) values ranging from 0.1 to $2\,h^{-1}$ Mpc, maximum fit radii of $R_{\rm max} = 1$, 2 and $4\,h^{-1}$ Mpc.

(iii) Using $\Upsilon(R;R_0)$ with $R_0 = 0.25$, 0.5 and $1\,h^{-1}$ Mpc, again with a variety of $R_{\rm min}$ values (always with $R_{\rm min} > R_0$). The value of $\Delta\Sigma(R_0)$ was determined on each noisy realization rather than from the well-determined mean over those scenarios, consistent with a real measurement for which we only have one observation of the lensing signal for a given sample. The estimation was done by fitting the data to the three-parameter functional form in equation

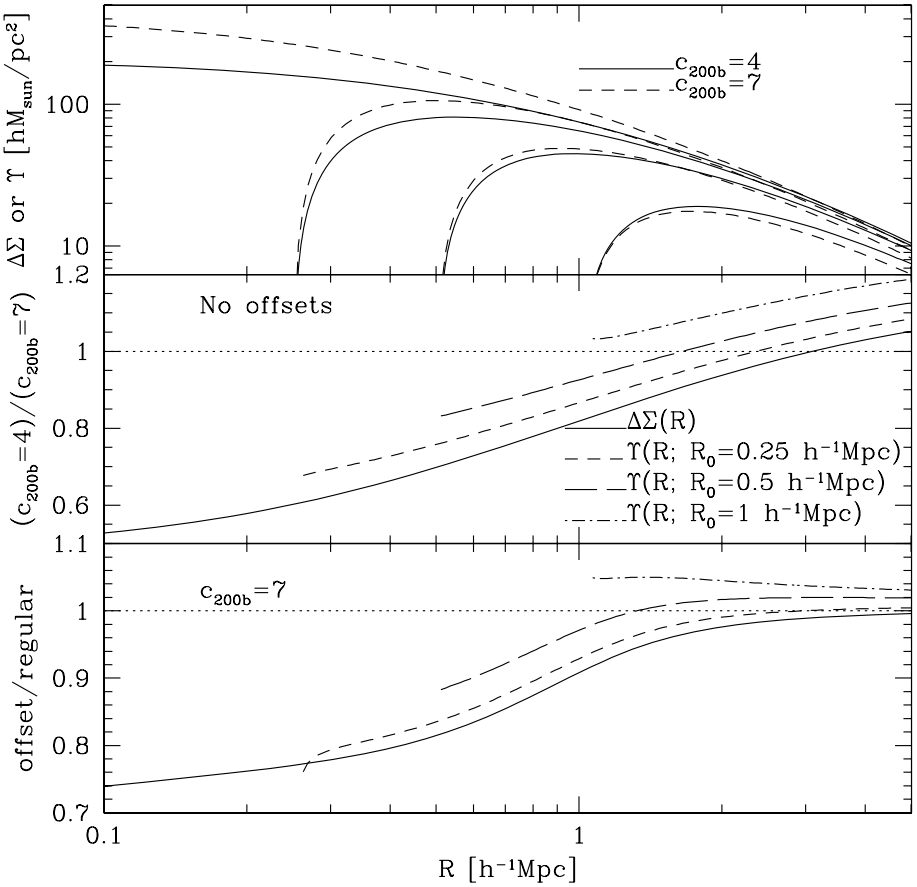


**Figure 1.** Top panel: from top to bottom, we show $\Delta\Sigma(R)$ and $\Upsilon(R;R_0)$ with $R_0 = 0.25$, 0.5 and $1\,h^{-1}$ Mpc. The solid lines are for $c_{200b} = 4$ and the dashed lines are for $c_{200b} = 7$; in both cases, $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.8$. Middle panel: without inclusion of centroid offsets, we show the ratio of these four quantities for $c_{200b} = 4$ versus $c_{200b} = 7$, where the line types indicate which quantity is used to construct the ratio, and the horizontal dotted line indicates a ratio of 1. Bottom panel: assuming $c_{200b} = 7$, we show the ratio of these four quantities when including centroiding offsets versus not, with the same line styles as in the middle panel.

(15) from $0.1 < R < 0.5$, $0.3 < R < 1$, and $0.7 < R < 1.3\,h^{-1}$ Mpc for $R_0 = 0.25$, 0.5 and $1\,h^{-1}$ Mpc, respectively.

In detail, the fits to $\Delta\Sigma(R)$ are performed via $\chi^2$ minimization in comparison with theoretical signals that were generated via the procedure described at the start of Section 5.1. Thus, for each of the lensing signal realizations $j$, denoted by $\Delta\Sigma_j^{(\mathrm{data})}(R_i)$ (for bins in transverse separation with index $i$ such that $R_{\mathrm{min}} \le R_i \le R_{\mathrm{max}}$) with noise variance $\sigma^2(\Delta\Sigma_j(R_i))$, we use the Levenberg–Marquardt algorithm (Levenberg 1944; Marquardt 1963; Press et al. 1992) to find the NFW profile mass that minimizes

$$\chi^2_j = \sum_i \frac{\left[\Delta\Sigma_j^{(\mathrm{data})}(R_i) - \Delta\Sigma^{(\mathrm{model})}(R_i|M_{200b}, c_{200b})\right]^2}{\sigma^2(\Delta\Sigma_j(R_i))} \qquad (19)$$

at fixed $c_{200b}$.

The fits to $\Upsilon(R;R_0)$ require an additional step: the conversion of both the theoretical signals [$\Delta\Sigma^{(\mathrm{model})}$, defined without noise in very narrow bins in $R$] and the mock data [$\Delta\Sigma^{(\mathrm{data})}$, defined in realistically broad bins with added noise] from $\Delta\Sigma(R)$ to $\Upsilon(R;R_0)$. In practice, the theoretical signal is defined such that we can very accurately interpolate to determine the value of $\Delta\Sigma(R_0)$, which is then used to construct $\Upsilon(R;R_0)$ directly using equation (14). For the noisy mock data, we must use a different procedure. We fit to $\Delta\Sigma(R)$ to estimate $\Delta\Sigma(R_0)$ using equation (15), so that $\Upsilon(R;R_0)$ can be constructed. We will shortly discuss more details of this procedure, because we find that the exact way of getting $\Delta\Sigma(R_0)$ is important: some methods introduce a bias on the mass, others add extra noise, neither of which is desirable. Once $\Upsilon(R;R_0)$ is determined for the mock signals, we then determine its covariance matrix using the distribution of values for all data sets. Finally, we minimize the $\chi^2$ function for each mock realization using equation (19) with $\Upsilon(R;R_0)$ in place of $\Delta\Sigma(R)$.

We then examined the distribution of best-fitting masses for the 1000 noise realizations to find the mass at the 16th, 50th (median) and 84th percentile. We define the spread in the masses, $\sigma_{\mathrm{M}}$, as being half the difference between the 84th and 16th percentile (which would be the standard deviation for a Gaussian distribution). The mass distributions are sufficiently close to Gaussian that using the mean rather than the median, and using the standard deviation directly, would not change the plots substantially. The median best-fitting mass $M_{200b,\mathrm{est}}$ relative to the input mass $M_{200b,\mathrm{true}}$ and the spread in the best-fitting masses are shown for both input profiles and each fit method as a function of $R_{\mathrm{min}}$ in Fig. 2. The criterion that we apply when selecting a robust mass estimator is that the ratio $M_{200b,\mathrm{est}}/M_{200b,\mathrm{true}}$ should not depend strongly on the input or output $c_{200b}$(though a systematic offset independent of input and output $c_{200b}$ is acceptable, since simulations can be used to correct for it).

We begin by considering the trends in the ratio $M_{200b,\mathrm{est}}/M_{200b,\mathrm{true}}$ with fitting method. When assuming $c_{200b} = 4$ while fitting to the

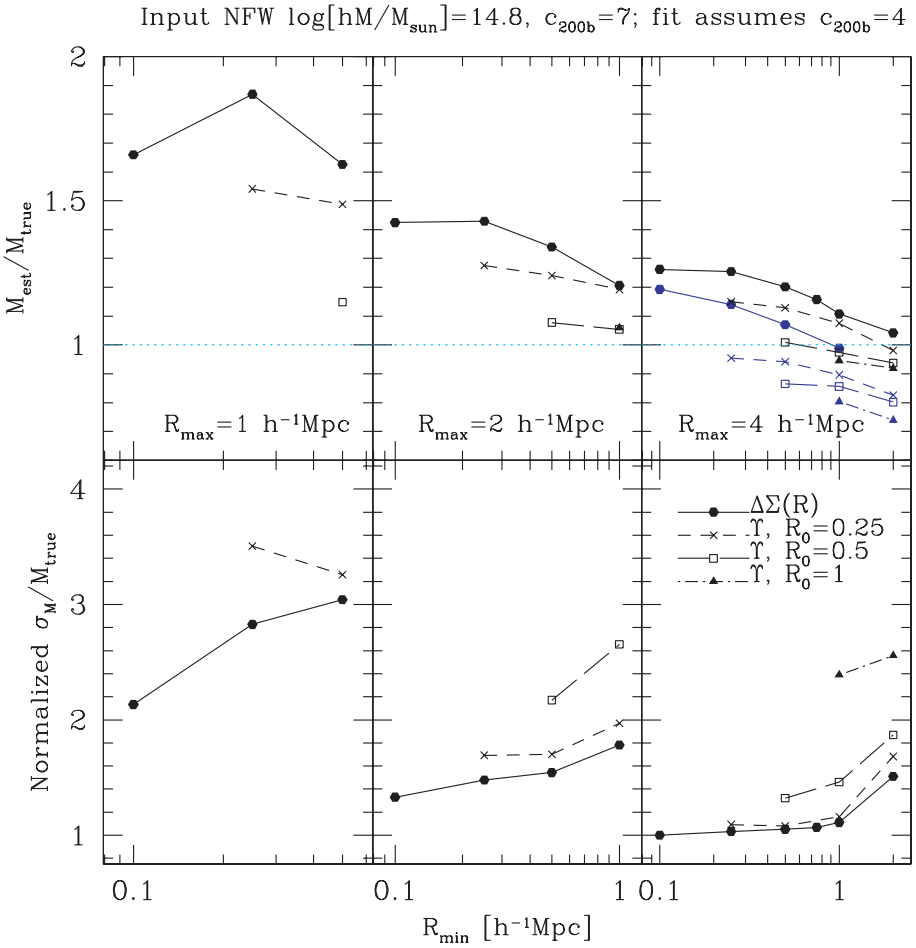


**Figure 2.** Results of parametric mass fits on noisy realizations of pure NFW profiles, with input $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.8$ and $c_{200b} = 7$, but $c_{200b} = 4$ assumed in the fits. The top and bottom rows show the ratio $M_{200b,\mathrm{est}}/M_{200b,\mathrm{true}}$ and the statistical error $\sigma$ $(M_{200b,\mathrm{est}})/M_{200b,\mathrm{true}}$, respectively. The latter is shown normalized to the minimum value of $\sigma$ $(M_{200b,\mathrm{est}})/M_{200b,\mathrm{true}} \sim 0.2$, which is obtained for the fit using the maximum information, $\Delta\Sigma(R)$ with $R_{\mathrm{min}} = 0.1$ and $R_{\mathrm{max}} = 4\,h^{-1}$ Mpc. The results are shown for various fitting methods (indicated with various line and point types shown on the plot) as a function of the minimum fit radius $R_{\mathrm{min}}$. From left to right, the panels show increasing $R_{\mathrm{max}}$ values of 1, 2 and $4\,h^{-1}$ Mpc. On the upper rightmost panel, the thin (blue) lines and points show the corresponding results for the $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.0$ profile.

profile with true $c_{200b} = 7$, as shown in Fig. 2, the fits to $\Delta\Sigma$ in the upper right panel with $R_{\mathrm{max}} = 4\,h^{-1}$ Mpc give ∼25 per cent overestimation of the mass for $R_{\mathrm{min}} \leq 0.5\,h^{-1}$ Mpc, improving to 3 per cent with $R_{\mathrm{min}} = 2\,h^{-1}$ Mpc (with, however, a doubling of the statistical error). The mass is overestimated in this case because for the majority of the radial range used for the fitting, the lensing signal for $c_{200b} = 4$ for this mass is below that for $c_{200b} = 7$ (Fig. 1), so the fitting routine compensates for the discrepancy by returning a higher mass. This trend of overestimated masses is decreased and eventually even reversed in sign for $\Upsilon(R; R_0)$ as we increase $R_0$, for reasons that are clear from Fig. 1. The reverse situation, with input $c_{200b} = 4$ and assumed $c_{200b} = 7$, leads to biases $M_{200b,\mathrm{est}}/M_{200b,\mathrm{true}}$ that are the inverse of the biases shown in Fig. 2, so we do not show this case in the figures. As shown, when using $\Upsilon(R; R_0)$ with $R_{\mathrm{min}} = R_0$, the statistical error increases over the minimum possible value from the $\Delta\Sigma(R)$ fits by factors of 1.14, 1.32 and 2.25 when using $R_0 = 0.25$, 0.5 and $1\,h^{-1}$ Mpc, respectively.

When fitting $\Upsilon(R; R_0)$ for all $R_0$ and $R_{\mathrm{min}}$, if we use a power law to fit for $\Delta\Sigma(R_0)$ (i.e. $q = 0$ in equation 15), then $M_{200b,\mathrm{est}}$ is consistently ∼3–5 per cent above $M_{200b,\mathrm{true}}$ even if the correct concentration is assumed in the fit. This overestimation of the mass occurs because the data are not consistent with a power law. Due to the trend of the signal with radius, the power-law fit tends to underestimate $\Delta\Sigma(R_0)$, thus overestimating $\Upsilon(R; R_0)$ and therefore $M_{200b,\mathrm{est}}$. However, we find that a full three-parameter fit significantly increases the noise, so we instead use a two-step procedure: we first fit with fixed $q = 0$ in equation (15) to get a mass, then we use the best-fitting signal to estimate $q$ at $R_0$, and use that fixed $q$ value for a second two-parameter fit for $\Delta\Sigma(R_0)$ which is used for a second fit to $\Upsilon(R; R_0)$ to get the mass. For the remainder of this work, we present results using that fitting procedure in order to best estimate the mass without increasing the noise too much.

Our criterion for a robust mass estimator on stacked cluster lensing data is that it should have systematic error that is relatively independent of the input $c_{200b}$ or the assumed $c_{200b}$ for the fit, at least when compared to the size of the statistical error. However, this robustness should not be achieved at the expense of too large an increase in the statistical error. As shown, the fits to $\Delta\Sigma(R)$ do not satisfy our robustness criterion, because assuming the wrong concentration can lead to a systematic error that is tens of per cent for reasonable $R_{\mathrm{min}}$. $\Upsilon(R; R_0)$ with $R_0 = 0.25\,h^{-1}$ Mpc improves somewhat on $\Delta\Sigma(R)$ in this regard, and for $R_{\mathrm{min}} = 1\,h^{-1}$ Mpc achieves a good combination of low systematic error and only a small increase in statistical error. $\Upsilon(R; R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc satisfies our

criterion for robustness when using $R_{\rm min} = R_0$ while increasing the error by about 20 per cent. A value of $R_0 = 1\,h^{-1}$ Mpc erases too much information and doubles the statistical errors. For individual cluster lensing, the criterion for a robust mass estimator may differ, since if one adds many more clusters then the statistical error may further decrease below the systematic error, so an even smaller systematic error is required.

The systematic errors shown here may be overly pessimistic for stacked data, given the wide variation in concentration that was allowed relative to what is seen in *N*-body simulations. However, several other systematics discussed in Sections 2.2 and 2.3 can mimic a change in concentration, such as baryonic effects. Thus, it is only reasonable that we should consider a broader range of concentrations than in the *N*-body simulations. When considering a narrower range, such as $4 < c_{200b} < 5$, the biases in the masses when fitting to $\Delta\Sigma$ with a fixed concentration are typically of the order of 10 per cent, or $\lesssim$2 per cent when fitting to $\Upsilon(R; R_0)$. For individual cluster lensing data, given the large lognormal scatter in concentration seen in simulations, these systematic errors we quote are not overly pessimistic. Furthermore, at this level of S/N, the fit $\chi^2$ values are perfectly acceptable even for the wrong value of concentration, so goodness-of-fit cannot be used to tell whether there is a systematic error.

In the upper right panel of Fig. 2, there are thin (blue) lines corresponding to a lower mass model that can be used to assess the mass-dependence of these systematic biases. As shown, the mass overestimation when fitting to $\Delta\Sigma(R)$ is not as severe for the lower mass cluster as for the higher mass cluster at fixed $R_{\rm min}$ (because the strongly concentration-dependent part of the inner profile has moved to smaller radii). The virial radius for this mass is about 1.85 smaller than for the higher mass model, suggesting that the choice of $R_0$ should be mass-dependent, with the optimal value of 15–25 per cent of the virial radius. In practice, this relation between the virial radius and $R_0$ could be achieved iteratively by choosing some default value of $R_0$, fitting with that value of $R_0$ and then using the resulting best-fitting mass to choose a more appropriate value of $R_0$ via

$$R_0 = (0.25h^{-1}\,{\rm Mpc})\left(\frac{M_{200b}}{10^{14}\,h^{-1}\,{\rm M}_\odot}\right)^{1/3}. \quad (20)$$

Here, we have assumed $\Omega_{\rm m} = 0.25$ and a spherical overdensity of $200\bar{\rho}$, and use comoving coordinations.

We also note that the fitted masses are weakly cosmology-dependent. For a fixed density profile, the mass that we estimate depends on the assumed $\Omega_{\rm m}$, with $M_{200b} \propto \Omega_{\rm m}^{-0.25}$ (we confirmed this scaling for the limited range of $0.2 \leq \Omega_{\rm m} \leq 0.3$). The $\Omega_{\rm m}$ dependence has two sources: first, we rescale the transverse separation and signal amplitude to account for the $\Omega_{\rm m}$ dependence of the distance measures used to convert $\theta$ and $\gamma_{\rm t}$ to $R$ and $\Delta\Sigma$, and second (and more significantly), the halo mass definition changes since we use a spherical overdensity of $200\overline{\rho}$. Thus, for higher $\Omega_{\rm m}$, the overdensity we use is larger, which reduces the mass and virial radius, also decreasing the concentration $c_{200b}$ since the scale radius is held fixed.

While stacked cluster lensing analyses from large surveys can provide cluster lensing data to tens of $h^{-1}$ Mpc, individual cluster lensing analyses that are not survey based typically have a limit of $R_{\rm max} = 1$–$2\,h^{-1}$ Mpc depending on the cluster redshift and telescope FOV. Consequently, we also explore the dependence of our results on the maximum scale used for the fits. Based on Fig. 1, we expect that the biases will be even higher in this case, since when restricting to smaller scales the differences between the lensing profiles $\Delta\Sigma(R)$ are more pronounced for the different values of $c_{200b}$.

The results of this test are shown only for the $\log_{10}[hM_{200b}/{\rm M}_\odot] = 14.8$ and $c_{200b} = 7$ profile, with assumed $c_{200b} = 4$, in the different columns of Fig. 2. As expected, when we decrease $R_{\rm max}$ (moving right to left across the figure), the systematic errors increase fairly drastically. For $R_{\rm max} = 1\,h^{-1}$ Mpc, the best we can achieve for the fitting methods tested here is with $\Upsilon(R; R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc, and even that method has a 25 per cent systematic error. For $R_{\rm max} = 2\,h^{-1}$ Mpc, $\Upsilon(R; R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc gives several per cent systematic errors for both $R_{\rm min} = 0.5$ and $1\,h^{-1}$ Mpc. It is clear that the existence of data to $R_{\rm max} = 4\,h^{-1}$ Mpc ($\approx 2r_{200b}$) is very helpful in decreasing the systematic and statistical errors.

These results suggest that the choice of mass estimator may depend on the maximum scale to which the lensing data can be measured for a given data set. If $1\,h^{-1}$ Mpc is the maximum scale for which data are available, then truly robust parametric measures of mass may be difficult to find; in the next section, we explore whether non-parametric measures may be better than parametric ones in this case. For larger values of $R_{\rm max}$, $\Upsilon(R; R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc seems adequate from the perspective of minimizing the combination of systematic and statistical error.

We next consider the effect of cluster centroiding errors, which were discussed in detail in Section 2.2. Note that our results here are more general than that particular systematic error, since several observational systematics in Section 2.3 have a similar form. We use the signals with $c_{200b} = 4$ and 7 for both $\log_{10}[hM_{200b}/{\rm M}_\odot] = 14.0$ and 14.8, and apply the offset model from Johnston et al. (2007) as described in the beginning of Section 5.1. It is important to note that this is only one example of how photometric errors in imaging data can cause centroiding errors for the cluster catalogue.

In Fig. 3, we show the results of the NFW mass fits to the profiles, with this offset distribution imposed on the data but ignored in the fit. Because Fig. 2 suggested that using $\Upsilon(R; R_0)$ with $R_0 = 1\,h^{-1}$ Mpc degrades the S/N unacceptably, we have only shown results for fits to $\Delta\Sigma(R)$ and for $\Upsilon(R; R_0)$ with $R_0 = 0.25$ and $0.5\,h^{-1}$ Mpc. As shown, for the higher mass model, for the input $c_{200b} = 4$ models, even when the correct $c_{200b}$ is assumed in the fit to $\Delta\Sigma$, the best-fitting masses are reduced by 5–25 per cent (lower mass) and by up to 7 per cent (higher mass) depending on $R_{\rm min}$. For the higher mass model, we find that $\Upsilon(R; R_0)$ with $R_0 = R_{\rm min} = 0.5\,h^{-1}$ Mpc gives fairly consistent results regardless of the input and assumed concentration. For the lower mass model, $\Upsilon(R; R_0)$ with $R_0 = R_{\rm min} = 0.25\,h^{-1}$ Mpc gives the most consistent results regardless of assumed $R_{\rm min}$. Moving to the right-hand column of this figure, for input $c_{200b} = 7$, we see that even with the correct assumed $c_{200b}$, fitting with $\Delta\Sigma(R)$ can lead to underestimated masses by up to 30 per cent (lower mass) or 10 per cent (higher mass) depending on $R_{\rm min}$. As for the input $c_{200b} = 4$ model, we find that the fitting technique and minimum scale that is most independent of assumed $c_{200b}$ is $\Upsilon(R; R_0)$ with $R_0 = R_{\rm min} = 0.5\,h^{-1}$ Mpc and $0.25\,h^{-1}$ Mpc for higher and lower mass scales, respectively. The ability of $\Upsilon(R; R_0)$ to robustly estimate masses even with these centroiding errors is a consequence of what we have noted in the bottom panel of Fig. 1, that the centroiding errors lead to biases in $\Upsilon(R; R_0)$ that change sign at some intermediate scale, so their effects approximately cancel out.

One important point raised by Fig. 3 is that the mass estimates using $c_{200b} = 4$ (assumed) are less affected by centroid offsets. This finding results from the fact that with a low concentration, the model already includes a relatively low level of mass in the inner cluster regions, and therefore is less affected than a higher concentration

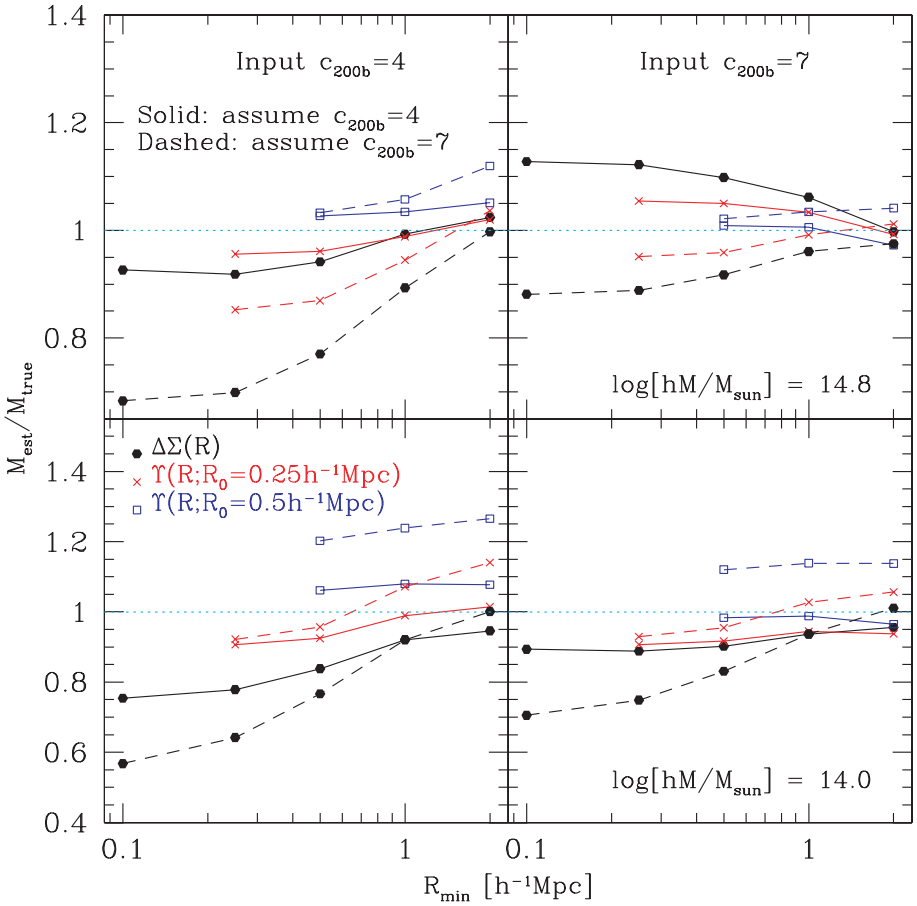


**Figure 3.** Here, we show the ratio $M_{200b,\rm est}/M_{200b,\rm true}$ for the pure $c_{200b} = 4$ (left column) and $c_{200b} = 7$ (right column) NFW models with $\log_{10}[hM_{200b}/{\rm M}_\odot] = 14.0$ (bottom row) and 14.8 (top row) after including effects of centroiding errors in the mock data. Results for the various fitting methods are shown as a function of the minimum fit radius $R_{\rm min}$, for fixed $R_{\rm max} = 4\,h^{-1}$ Mpc. The different point styles and colours as indicated on the plot show what type of fitting was done [$\Delta\Sigma(R)$ or $\Upsilon(R;R_0)$ with various $R_0$ values]; the different line types (solid versus dashed) indicate which value of $c_{200b}$ was assumed. The dotted horizontal lines indicate a ratio of 1, the ideal unbiased case.

halo. Thus, it may be advantageous to assume a concentration at the low end of the expected range when fittings to $\Upsilon(R;R_0)$ in scenarios involving possibly substantial offsets of the chosen BCG from the true cluster centre.

### 5.1.2 *Non-parametric modelling*

In this section, we use the same noisy realizations of theoretical cluster profiles as in the previous section, but we estimate masses using the aperture mass statistic $\zeta_c$. In this case, we begin with the NFW profile with $\log_{10}[hM_{200b}/{\rm M}_\odot] = 14.8$ and $c_{200b} = 7$. We try various options for the different aspects of this analysis:

(i) Varying $R_1$ (the radius below which we are trying to estimate the enclosed mass, using the shear above that radius) between three values: 0.275, 0.5 and 1.1 $h^{-1}$ Mpc.

(ii) Varying $R_{o1}$ between two values: 1.1 and 2.0 $h^{-1}$ Mpc.

(iii) Varying $R_{o2}$ between two values: 2 and 4 $h^{-1}$ Mpc (maintaining at all times the strict hierarchy $R_1 < R_{o1} < R_{o2}$).

(iv) Neglecting the second term in equation (12) as in Okabe et al. (2009), and estimating it using the best-fitting NFW profile with some assumed concentration, as in Hoekstra (2007). We do not test the case in which the integral from $R_{o1}$ to $R_{o2}$ may be done analytically, because often for individual cluster lensing studies this is not even possible since $R_{o2}$ is outside the FOV. With survey data or mosaic telescope data, the signal may indeed be measured to $R_{o2}$, but it is typically quite noisy on those large scales, so this procedure would introduce even more noise into the estimated masses.

(v) Assuming $c_{200b} = 4$ and 7 whenever a profile assumption is necessary: for the estimate of the second term in equation (12), and for the conversion from $M_{\rm 2D}(< R_1)$ to the 3D $M_{200b}$.

The procedure is as follows. We use the (noisy) realizations of the lensing signal for pure NFW profiles in logarithmic annular bins to estimate $\zeta_c$ using a given set of radii $(R_1, R_{o1}, R_{o2})$. Thus, we use the signal for $R_1 < R < R_{o1}$ to calculate the first term in equation (12) via direct summation over the noisy mock data in broad logarithmic bins in $R$. We also estimate the second term using the fits to $\Delta\Sigma(R)$ for $R_1 < R < 4\,h^{-1}$ Mpc for the assumed value of $c_{200b}$. To do so, we use the lensing profile for the best-fitting $M_{200b}$, determined to high precision as in the start of Section 5.1, and estimate the second term using direct summation over the numerically determined (non-noisy) profile in narrow logarithmic bins in $R$. Given $\zeta_c$ estimated with and without the second term, we then use our assumed $c_{200b}$ to convert the $M_{\rm 2D}(< R_1)$ to a 3D virial radius $M_{200b}$, which (at fixed $c_{200b}$) is a simple one-to-one mapping that can be determined via numerical integration.

In Table 3, we present the following, first without the correction term for the outer annulus and then with it: the accuracy in recovering $M_{\rm 2D}(< R_1)$, the accuracy in recovering $M_{200b}$ and the

**Table 3.** Results of tests of NFW mass recovery for $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.8$ and $c_{200b} = 7$ when using the aperture mass statistic $\zeta_c$.

| $R_1$ $h^{-1}$ Mpc | $R_{o1}$ $h^{-1}$ Mpc | $R_{o2}$ $h^{-1}$ Mpc | $M_{2D}/M_{2D,true}$ | $M_{200b}/M_{200b,true}$ Neglect second term | $\sigma_M^{(\zeta_c)}/\sigma_M^{(fit)}$ | $M_{2D}/M_{2D,true}$ | $M_{200b}/M_{200b,true}$ Estimate second term | $\sigma_M^{(\zeta_c)}/\sigma_M^{(fit)}$ |
|---|---|---|---|---|---|---|---|---|
| | | | | Assume $c_{200b} = 7$ | | | | |
| 0.275 | 1.1 | 2 | 0.83 | 0.74 | 1.18 | 1.00 | 1.00 | 1.40 |
| 0.275 | 1.1 | 4 | 0.83 | 0.74 | 1.18 | 1.00 | 1.00 | 1.46 |
| 0.275 | 2 | 4 | 0.95 | 0.90 | 1.43 | 1.00 | 1.00 | 1.54 |
| 0.5 | 1.1 | 2 | 0.66 | 0.58 | 0.87 | 1.00 | 1.00 | 1.20 |
| 0.5 | 1.1 | 4 | 0.66 | 0.58 | 0.87 | 1.00 | 1.00 | 1.24 |
| 0.5 | 2 | 4 | 0.90 | 0.86 | 1.12 | 1.00 | 1.00 | 1.17 |
| 1.1 | 2 | 4 | 0.62 | 0.55 | 0.71 | 1.00 | 1.00 | 1.01 |
| | | | | Assume $c_{200b} = 4$ | | | | |
| 0.275 | 1.1 | 2 | 0.83 | 1.26 | 1.60 | 1.01 | 1.75 | 1.95 |
| 0.275 | 1.1 | 4 | 0.83 | 1.26 | 1.60 | 1.02 | 1.80 | 2.00 |
| 0.275 | 2 | 4 | 0.95 | 1.62 | 2.01 | 1.02 | 1.81 | 2.24 |
| 0.5 | 1.1 | 2 | 0.66 | 0.74 | 0.95 | 1.01 | 1.35 | 1.37 |
| 0.5 | 1.1 | 4 | 0.66 | 0.74 | 0.95 | 1.02 | 1.48 | 1.44 |
| 0.5 | 2 | 4 | 0.90 | 1.21 | 1.15 | 1.02 | 1.50 | 1.40 |
| 1.1 | 2 | 4 | 0.62 | 0.72 | 0.68 | 1.04 | 1.19 | 1.08 |

statistical error on the recovered $M_{200b}$ relative to that from the fit to $\Delta\Sigma(R)$ using $R_1 < R < 4\,h^{-1}$ Mpc. These results are shown for both assumed concentration values, $c_{200b} = 4$ and 7, given the true profile with $\log_{10}[hM_{200b}/\mathrm{M}_\odot] = 14.8$ and $c_{200b} = 7$.

There are a few conclusions that can be drawn from this table. First, we begin with the idealized case in the top section of the table, where the assumed $c_{200b}$ is the same as the true one. In this case, we see that depending on the configuration of the three radii used to estimate $\zeta_c$, the projected mass may be underestimated by 5–40 per cent if the second term in equation (12) is ignored. This underestimate is propagated into an underestimate of the 3D $M_{200b}$ that ranges from 10 to 45 per cent. This underestimate due to ignoring the mass in the outer annulus is less important for $R_{o1} \gg R_1$ as it is for cases where the two radii are relatively close to each other. We also see that the statistical error on the inferred $M_{200b}$ from the aperture mass is typically comparable to that for the fits to $\Delta\Sigma$ using $R_1 < R < 4\,h^{-1}$ Mpc.

In this ideal case with the correct assumed $c_{200b}$, correcting for the second term in $\zeta_c$ using the best-fitting profile to $\Delta\Sigma(R)$ for $R_1 < R < 4\,h^{-1}$ Mpc leads to unbiased recovery of both $M_{2D}(< R_1)$ and $M_{200b}$; however, the statistical errors on $M_{200b}$ are larger than when fitting to $\Delta\Sigma(R)$ by typically tens of per cent. This higher level of noise is due to the noisy profile used to estimate the second term in $\zeta_c$.

Next, we consider the lower half of the table, in which we use a profile with $c_{200b} = 7$, and assume $c_{200b} = 4$. First, when we do not include the second term in equation (12), the projected masses are underestimated, as before. Second, when we include the second term in equation (12), the projected masses are all slightly overestimated (by several per cent), and the 3D $M_{200b}$ are overestimated by 20–80 per cent (depending on $R_1$, with smaller $R_1$ leading to larger biases). We can explain the slight overestimation of the $M_{2D}$ when including the second term in $\zeta_c$ by the fact that we do the correction using profiles with a low $c_{200b}$, which give too much mass in the outer regions from which the second term is derived. The significant overestimation of $M_{200b}$ arises because, when we assume too low a concentration, then we anticipate a profile with a low amount of mass on small scales, so the conversion factor from $M_{2D}(< R_1)$ to $M_{200b}$ is a large number. This effect will be worse for small $R_1$, since the difference between the lensing profiles for different concentrations is most significant there. If we allow a smaller variation, such as true $c_{200b} = 5$ and assumed $c_{200b} = 4$, then we find a 10–20 per cent effect on the 3D virial masses.

In general, the results for an input profile with $c_{200b} = 4$ can be understood as the inverse of the results given in Table 3. However, for a less concentrated profile, the bias in $M_{2D}$ due to neglect of mass in the outer annulus is more significant. For a lower mass halo and fixed transverse separation, the mass in the outer annulus is less important.

We next consider the effect of centroiding errors on the aperture mass. When using the two mass models, we find that the projected masses $M_{2D}$ are systematically suppressed by 10–14 per cent due to centroiding errors. The exact level of suppression depends slightly but not very strongly on the value of $R_1$ in the range we have considered, and this suppression is then propagated into a suppression of $M_{200b}$.

Because of the definition of $\zeta_c$, biases in the lensing signal calibration that can be expressed as a single scale-dependent factor enter linearly into the estimated masses in projection, $M_{2D} \propto \Delta\Sigma$. However, when using some model for the spherical density profile to estimate the mass within some radius defined in terms of a spherical overdensity, such as $M_{200b}$, the mass will scale even more strongly with $\Delta\Sigma$, because as the signal increases, the spherical overdensity radius moves outwards, thus including more mass in the total. The exact scaling of the enclosed mass within some spherical overdensity depends on the model used to define the appropriate radius, and on which overdensity is used, but typically the inferred $M_{200b} \propto \Delta\Sigma^{1.5}$.

One important point regarding the bias given in Table 3 due to the wrong assumed concentration [for converting $M_{2D}(< R_1)$ to $M_{200b}$] is that it has the same sign as the bias due to assumption of the wrong concentration when fitting to $\Delta\Sigma(R)$. Consequently, consistency of the $M_{200b}$ from the aperture mass calculation and the NFW fits to $\Delta\Sigma(R)$ does not tell us whether the assumed concentration is correct.

In summary, we have found that the aperture mass statistic $\zeta_c$ has a strong dependence on the assumed $c_{200b}$ when converting the extracted projected masses to 3D $M_{200b}$. An additional problem

is that a (much less concentration-dependent) correction must be used to properly correct for the term from the outer annulus $R_{o1} < R < R_{o2}$; otherwise, the projected masses can be underestimated by tens of per cent, an effect that is worse for more massive clusters. While less affected by centroiding errors than fits to $\Delta\Sigma(R)$ that use scales below $0.5\,h^{-1}$ Mpc, the aperture mass statistic can still be suppressed by roughly 10 per cent due to centroiding errors (or any of the other errors from Section 2.3 that have a similar form). Finally, it can be substantially noisier, typically by 50 per cent, than fits to $\Delta\Sigma(R)$ using the same scales [which means that it is noisier than fits to $\Upsilon(R; R_0)$].

In principle, these biases due to the concentration-dependence of the 2D to 3D conversion may be removed if the conversion from $M_{2D}(< R_1)$ to $M_{200b}$ is carried out using the best-fitting NFW profile from fits to both $c_{200b}$ and $M_{200b}$, as in Okabe et al. (2009). However, as will be shown in the next section, these fits tend to be substantially noisier due to the additional fit parameter, which will further amplify the noise on the recovered mass from $\zeta_c$. Thus, this approach is not very advantageous relative to the fits to $\Upsilon(R; R_0)$, which are similarly insensitive to the assumed concentration but are only slightly noisier than fits to $\Delta\Sigma(R)$.

### 5.2 Profiles from *N*-body simulations

In this section, we present the results of tests of mass estimation using cluster profiles measured from the simulations described in Section 3. The properties of these simulated cluster samples are summarized in Table 2. We use the signal from simulations for mass threshold samples selected by taking all clusters above some $M_{200b}$ such that $\overline{n} = 0.25$, 2 and $16 \times 10^{-6}$ $(h/\,\mathrm{Mpc})^3$, with the first of these samples shown in Fig. 4. The samples have mean masses $\langle M_{200b}\rangle = 7.36$, 3.95 and $1.55 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$, though the stitching to NFW profiles below certain scales as described in Section 3 increases the mass by several per cent. All comparisons between estimated $M_{200b,\mathrm{est}}$ and true $M_{200b,\mathrm{true}}$ take this small increase into account. The error bars shown in Fig. 4, which include cosmic variance, are estimated by dividing the eight simulation boxes each into 20 sub-volumes comparable in size to that of the maxBCG cluster sample, and finding the variance of the signal between the 160 total sub-volumes. We have only shown the case of stitching to NFW profiles with $c_{200b} = 5$ at $0.2\,h^{-1}$ Mpc in Fig. 4; when stitching to an NFW profile with $c_{200b} = 7$ at $1\,h^{-1}$ Mpc, the signal on smaller scales is steeper. In the former case, this resolution correction increases the mass by 1.5 per cent compared to the mass in the simulations; in the latter case, the correction is 6 per cent.

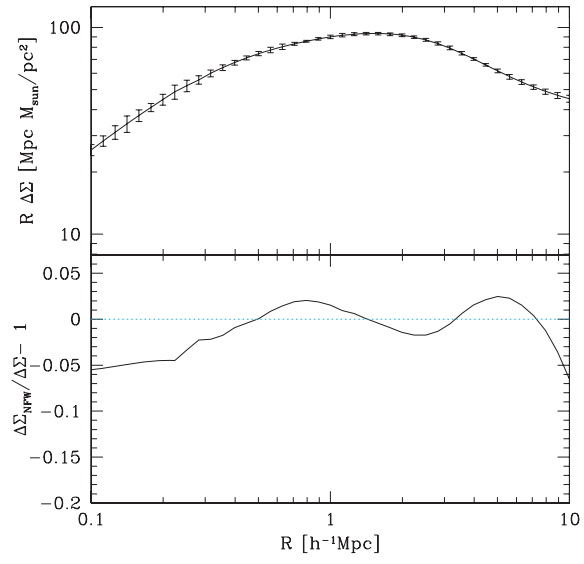


**Figure 4.** Top: lensing signal $R\,\Delta\Sigma(R)$ from simulations for the higher mass (lower number density) threshold sample described in the text. The solid lines with error bars show the signal stitched to an NFW profile with $c_{200b} = 5$ for $r < 0.2\,h^{-1}$ Mpc (to remove resolution effects). Bottom: ratio of the signal for the best-fitting NFW profile to the true simulation signal.

In the bottom panel of Fig. 4, we compare $\Delta\Sigma(R)$ from the simulations to that for the best-fitting NFW profile (determined by varying both $M_{200b}$ and $c_{200b}$ and fitting using $0.2 < R < 2\,h^{-1}$ Mpc). As shown, for most of the scales of interest, the deviations are less than 5 per cent. We see that the NFW profile overestimates the signal on $\sim$3–$8\,h^{-1}$ Mpc scales. This result is consistent with that from Clowe et al. (2004), who also find that on large scales the density profiles fall off faster than NFW. The effect is more significant when expressed in terms of the density profile $\rho(r)$. On the largest scales shown here, as $R$ approaches $10\,h^{-1}$ Mpc, the NFW profile signal starts to be too low, because the simulation includes contributions from LSS [again, this effect is more pronounced in $\rho(r)$ and appears at lower radii].

For the sections that follow, we have added realistic levels of shape noise to the signal, based on calculations of the lensing signal using the maxBCG cluster catalogue with similar number density samples.

#### *5.2.1 Parametric modelling*

We begin by showing the effects of parametric modelling of the lensing profiles from simulations. We use the three aforementioned mass threshold samples, with the two methods of connecting to NFW profiles (Section 3) to correct for resolution effects: $c_{200b} = 5$ at $r = 0.2\,h^{-1}$ Mpc, and $c_{200b} = 7$ at $r = 1\,h^{-1}$ Mpc. We then fit to $\Delta\Sigma(R)$ and $\Upsilon(R; R_0)$ with $R_0 = 0.25$ and $0.5\,h^{-1}$ Mpc, with varying $R_{\mathrm{min}}$ and $R_{\mathrm{max}}$, for our two extreme concentration values of $c_{200b} = 4$ and $c_{200b} = 7$. The fitting procedure is the same as for the analytic profiles in Section 5.1.1. Fig. 5 shows the results of these fits for the highest and lowest of the mass threshold samples.

The important point to consider in this plot is that we would like the output mass from a given estimator to be relatively insensitive to the form of the inner profile (represented by the two different connections to NFW profiles on small scales) and to the assumed concentration. Furthermore, we would like it to be only weakly dependent on the mass, assuming that corrections for systematic bias will be derived from simulations, but that strong mass dependence may be difficult to calibrate out correctly. Consequently, what we hope to see in an optimal estimator of cluster mass is that all the lines on a given panel (representing the results with different input profiles, assumed concentrations and masses) give very similar results; we do not want to use an estimator that has large scatter between the lines. So, for example, the lower left panel shows, as we already saw with pure NFW profiles in Section 5.1, that fitting $\Delta\Sigma(R)$ to NFW profiles with $R_{\mathrm{max}} = 1\,h^{-1}$ Mpc and a fixed concentration leads to very large systematic uncertainties, more than a factor of 2 total range in the best-fitting masses. As we increase $R_{\mathrm{max}}$, we become less sensitive to the inner details of the profile, so the scatter between the lines becomes less significant, but for $R_{\mathrm{min}} \leq 1\,h^{-1}$ Mpc they still cover a range of $\sim$40 per cent in mass even for $R_{\mathrm{max}} = 4\,h^{-1}$ Mpc, well outside the virial radius. For $R_{\mathrm{min}} = 2$ and $R_{\mathrm{max}} = 4\,h^{-1}$ Mpc, the systematic uncertainty is only $\sim$10 per cent;

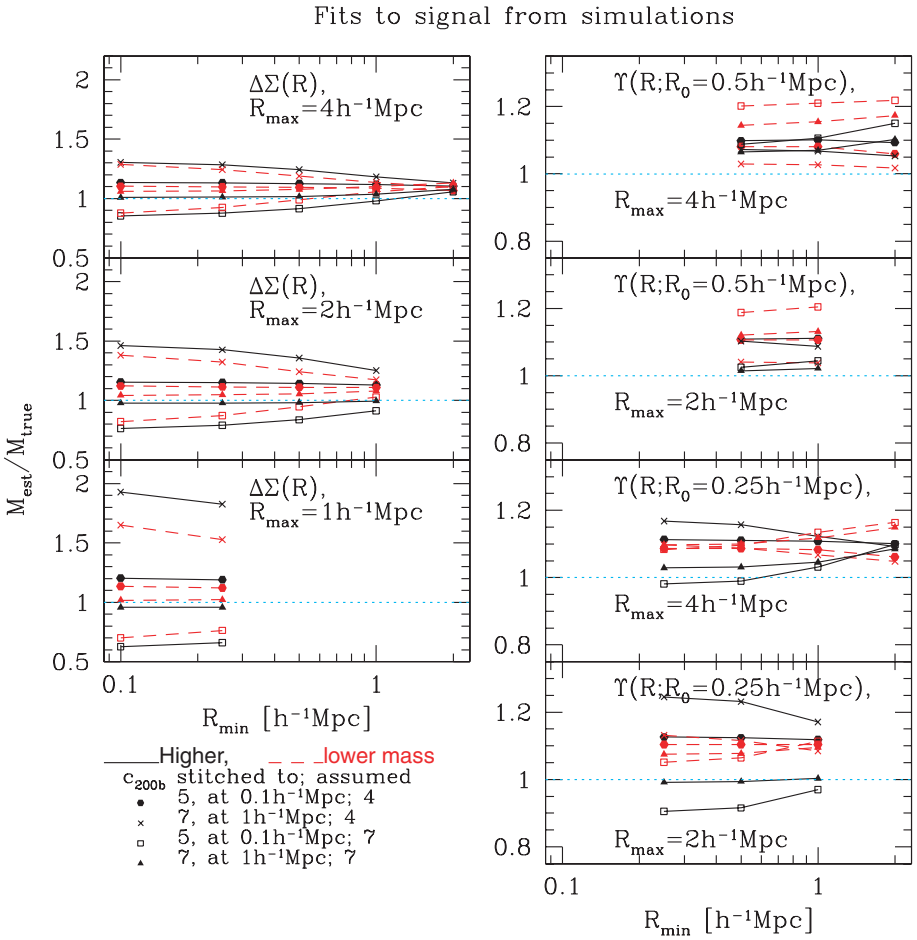


**Figure 5.** Results for $M_{200b,\mathrm{est}}/M_{200b,\mathrm{true}}$ as a function of the minimum fit radius, $R_{\mathrm{min}}$, from parametric fits to the lensing signal from simulations, for seven different combinations of observable [$\Delta\Sigma(R)$ or $\Upsilon(R;R_0)$] and $R_{\mathrm{max}}$ shown separately in each panel. As indicated in the legend, line colours and types are used to indicate the mass scale, whereas point styles are used to indicate the input signal (which NFW profile was used to correct for resolution effects) and assumed concentration in the fits, either $c_{200b} = 4$ or 7. The horizontal dotted line on each panel shows the ideal unbiased result. The vertical axis is the same for all panels in the left column, and spans a smaller range for all panels in the right column so that the details will be more visible.

however, the statistical error on the mass (not shown on this plot) has roughly doubled relative to the results with $R_{\mathrm{min}} \leq 0.5\,h^{-1}$ Mpc.

In contrast, we see that $\Upsilon(R; R_0 = 0.25$ and $0.5\,h^{-1}$ Mpc), in the right-hand panels in Fig. 5, performs quite well. The difference between the two mass threshold samples suggests that a larger $R_0 \sim R_{\mathrm{min}}$ is preferable for samples with larger halo masses, with minimal profile-related systematics for $R_0 = 0.5\,h^{-1}$ Mpc for the sample with a mass above $7 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$, and $R_0 = 0.25\,h^{-1}$ Mpc for the sample with a mass around $1.6 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ (and therefore smaller scale and virial radii). While the cluster mass is not known a priori, a preliminary fit with one choice of $R_0$ could be used to estimate an approximate mass, and then a new $R_0$ could be chosen to be around 1/4 to 1/5 of the virial radius, provided that this scale is reliable from the perspective of small-scale systematics (Section 2.3).

In all cases, $\Upsilon(R; R_0)$ does not converge to the true mean mass, for two reasons: (1) the lensing signal includes a small but non-negligible contribution due to LSS on the scales we have used, leading to an overestimation of $M_{200b,\mathrm{est}}$; and (2), even on scales where LSS is not important, the simulation profiles fall off faster than the NFW model, which somewhat counteracts the previous effect. Fortunately, since it is relatively insensitive to the inner details of the profile, the assumed concentration and the mass, this systematic positive bias in the masses can be calibrated out using simulations, whereas systematic uncertainty in $\Delta\Sigma(R)$-based mass estimates due to concentration assumptions and small-scale effects cannot be calibrated out in this way.

Some differences in these results from Section 5.1 can be attributed to the LSS in the simulations that was not put into the pure NFW profiles, and to the fact that the simulation profiles are not strictly NFW profiles. So, for example, in Fig. 2, the results for fitting to $\Delta\Sigma(R)$ converge to the true mass on large scales if the right concentration is assumed, whereas the fitting to $\Delta\Sigma(R)$ in simulations converges to a mass that is too high by 5 to 10 per cent when using the largest scales only.

As in Section 5.1, we point out that for a stacked cluster sample, the level of variation we have allowed in the assumed $c_{200b}$ is likely excessive from the standpoint of $N$-body simulations. However, given the systematic profile changes that may occur due to baryonic effects, centroiding errors and intrinsic alignments, the variation we have assumed is not entirely unreasonable. For fits to individual cluster lensing data, the variation we have assumed is quite reasonable, and possibly even an underestimate of the true variation, given the large lognormal scatter in cluster concentrations in $N$-body simulation plus these other systematics that change the profile on small scales.

We also estimate the effects of centroiding errors on the parametric mass recovery. As for the theoretical profiles, we use the model for centroiding errors given in Johnston et al. (2007), with offset fractions of 20 and 25 per cent for the lower and higher abundance thresholds, respectively.

Here, we describe how centroiding errors modify the curves that were shown in Fig. 5. As we have seen before, the offsets suppress masses estimated directly from $\Delta\Sigma(R)$, with larger biases when restricting to smaller scales. Furthermore, the profiles with more mass in the inner regions are more strongly affected. For example, the simulation signal stitched to NFW with $c_{200b} = 7$ at $1\,h^{-1}$ Mpc is more strongly affected than the signal stitched to $c_{200b} = 5$ at $0.2\,h^{-1}$ Mpc. Given that the former resulted in mass estimates that were above the masses estimated from the latter when fitting to $\Delta\Sigma(R)$ (without offsets, Fig. 5) by up to tens of per cent depending on the value of $R_{\rm max}$, the net effect of offsets is to lower all estimated masses while also reducing the difference between the curves, since those with the two stitched profiles now tend to agree more closely. For example, when using $R_{\rm max} = 1\,h^{-1}$ Mpc, the values of $M_{200b,{\rm est}}/M_{200b,{\rm true}}$ without including centroiding errors in the modelling range from 0.6 to 1.9 (a factor of 3). Centroiding errors in the input data reduce the range of $M_{200b,{\rm est}}/M_{200b,{\rm true}}$ to 0.4 to 0.9 (a factor of 2), where the main cause of this variation is the assumed value of $c_{200b}$ rather than the input profile. For $R_{\rm min} = 0.5$ and $R_{\rm max} = 4\,h^{-1}$ Mpc, $M_{200b,{\rm est}}/M_{200b,{\rm true}}$ ranges from 0.9 to 1.25 when we do not include centroiding errors, whereas when we include them, it ranges from 0.8 to 1. As we have seen before when using pure NFW profiles, $\Upsilon(R;R_0)$ with $R_{\rm min} = R_0 = 0.5\,h^{-1}$ Mpc is almost completely insensitive to this model for centroiding errors when using $R_{\rm max} = 4\,h^{-1}$ Mpc (masses are suppressed at the 10 per cent level with $R_{\rm max} = 2\,h^{-1}$ Mpc). This insensitivity to such systematics makes the ADSD statistic $\Upsilon(R;R_0)$ the optimum choice for parametric mass fitting on stacked clusters selected from imaging data, which is prone to centroiding errors of this variety.

In principle, explicit modelling of the offset distribution, as in Johnston et al. (2007), can remove its effects when fitting to $\Delta\Sigma(R)$. However, the exact results may be sensitive to the details of the centroiding model used and its accuracy when compared to the true distribution, which is not typically well known. For example, that paper uses mock simulations to estimate the centroiding error distribution, which means that this model is quite sensitive to the realism of the model for populating the simulation dark matter haloes with galaxies. Furthermore, the other systematic uncertainties associated with using $\Delta\Sigma(R)$ (e.g. sensitivity to baryonic effects and intrinsic alignments) remain, whereas their influence on $\Upsilon(R;R_0)$ is much smaller.

Another issue we consider is the effect of overall lensing signal calibration biases on the estimated masses. As a test, we use the signals from simulations multiplied by factors of 0.9 and 1.1, and refit for the masses. The results are used to estimate a power-law relation $M_{200b} \propto \Delta\Sigma^{\eta}$, and $\eta$ is determined for the different mass scales, stitched signals, assumed concentrations, fit method [$\Delta\Sigma(R)$ or $\Upsilon(R;R_0)$], and minimum and maximum fit radii. Note that $\eta$ is also dependent on the spherical overdensity used to define the profile, though we do not explore this effect in detail. A naive scaling of surface mass density with mass predicts $\eta = 1.5$, but other effects will modify this. The results of this test are shown in Fig. 6.

As shown, $\eta$ is a decreasing function of $R_{\rm min}$ and $R_{\rm max}$. When fitting to $\Delta\Sigma(R)$, $\eta$ does not depend on the details of the inner profile, and is larger for higher masses and lower assumed $c_{200b}$, with the dependence on $c_{200b}$ being the more significant dependence. For example, when fitting to $\Delta\Sigma(R)$ for the lower mass sample from simulations stitched to an NFW profile with $c_{200b} = 5$ at $0.2\,h^{-1}$ Mpc, using $0.5 < R < 4\,h^{-1}$ Mpc for the fits, we find that $M_{200b} \propto \Delta\Sigma^{1.42}$. In contrast, when fitting to $\Upsilon(R;R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc, the trends in the fitting mass with calibration are stronger for a given combination of ($R_{\rm min}$, $R_{\rm max}$). Here, we see that there is minimal dependence on mass, and some small dependence on the details of the profile and the assumed concentration. For the same case considered when fitting to $\Delta\Sigma(R)$, we find $\eta = 1.75$, an increase of 23 per cent. Consequently, systematic errors in $\Upsilon(R;R_0)$ due to miscalibration of the lensing signal are larger than systematic errors in $\Delta\Sigma(R)$ (assuming that other aspects of the fit, such as $R_{\rm min}$ and $R_{\rm max}$, are similar).

Next, we briefly discuss the effects of allowing both $c_{200b}$ and $M_{200b}$ to vary, rather than fixing $c_{200b}$, as in Okabe et al. (2009). While this procedure has the disadvantage of increasing the statistical errors on the mass, it does allow for improved mass recovery. Our results suggest that with NFW fits to $\Delta\Sigma(R)$ with $0.5 < R < 4\,h^{-1}$ Mpc, the degeneracy between $M_{200b}$ and $c_{200b}$ is such that $M_{200b} \propto c_{200b}^{-1/3}$. This result explains the magnitude of the deviations from the true mass when the concentration is fixed to a value that is not consistent with the best-fitting concentration (though, again, the deviations in concentration we have tested are not sufficiently bad that the fit $\chi^2$ values reveal a clear discrepancy). In contrast, the exponent on that scaling between $M_{200b}$ and $c_{200b}$ is far closer to zero when fitting to $\Upsilon(R;R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc using the same scales: $M_{200b} \propto c_{200b}^{0.05}$. This degeneracy becomes more striking when the fits are restricted to smaller scales, e.g. $M_{200b} \propto c_{200b}^{-1}$ when fitting to $\Delta\Sigma(R)$ using $0.1 < R < 1\,h^{-1}$ Mpc.

When fitting the simulation signals with both $M_{200b}$ and $c_{200b}$ as free parameters, we find that even when centroiding errors are included in the data, the fits are able to recover the masses for both mass scales and inner profiles, for several types of fits that we attempted [using $\Delta\Sigma(R)$ from $0.5 < R < 4\,h^{-1}$ Mpc, from $0.1 < R < 1\,h^{-1}$ Mpc, and using $\Upsilon(R;R_0)$ with $R_0 = 0.5$ and $0.5 < R < 4\,h^{-1}$ Mpc]. When using data from $0.5 < R < 4\,h^{-1}$ Mpc, the deviations of the signal in simulations from an NFW profile led to best-fitting masses that are 5 per cent higher than the true masses; when fitting from $0.1 < R < 1\,h^{-1}$ Mpc, the best-fitting masses are only $\sim$2 per cent higher than the true ones, because the deviations from NFW are not as striking on those scales, and the LSS term is also more negligible. The mass estimates tend to be noisier in this case, and the concentrations that are recovered are highly suppressed relative to the true concentrations when centroiding errors are included (e.g. from best-fitting $c_{200b} \sim 5$ and $\sim$6.5 without centroiding errors, down to 3 and 3.5 with centroiding errors). We find that these two-parameter fits for mass and concentration lead to statistical errors in the masses that are larger than the errors in one-parameter fits by approximately 45 per cent. This increase is larger than the increase when fitting to $\Upsilon(R;R_0)$ (14 or 32 per cent, for $R_0 = 0.25$ or $0.5\,h^{-1}$ Mpc), and $\Upsilon(R;R_0)$ has the additional advantage of removing the impact of small-scale systematics, which would still be present when fitting $\Delta\Sigma(R)$ to the two-parameter model.

Finally, when fitting with free $M_{200b}$ and $c_{200b}$, the dependence of the best-fitting masses on the lensing signal calibration is reduced. For example, when fitting to $\Delta\Sigma(R)$ using $0.5 < R < 4\,h^{-1}$ Mpc and fixed concentration, we had found previously that $M_{200b} \propto \Delta\Sigma^{1.42}$. When the concentration is allowed to vary, that exponent becomes $\eta = 1.25$. This change results from the fact that if the signal increases, the assumed mass and therefore $r_{200b}$ increase as well, so for a fixed scale radius determined from the data, the concentration

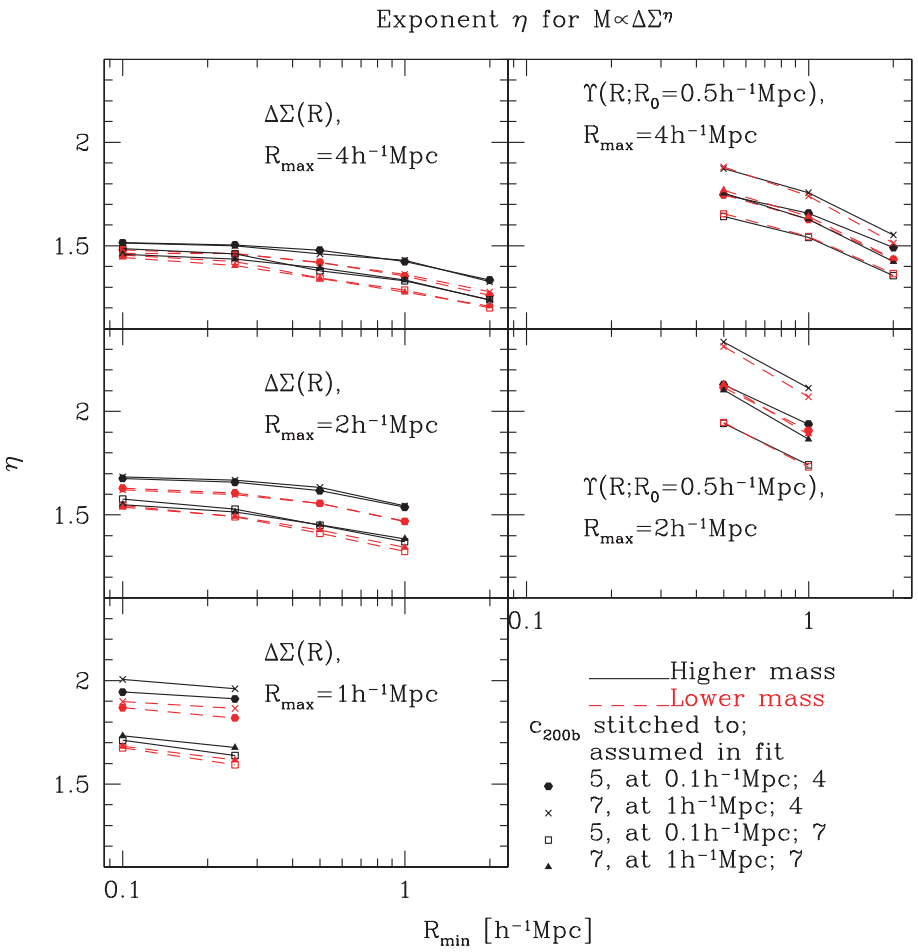


**Figure 6.** Results for $\eta$, the scaling of the estimated $M_{200b}$ with the lensing signal calibration, as a function of the minimum fit radius, $R_{\rm min}$, from parametric fits to the lensing signal from simulations, for five different combinations of observable [$\Delta\Sigma(R)$ or $\Upsilon(R;R_0)$] and $R_{\rm max}$ shown separately in each panel. We only show $\eta$ for the highest and lowest mass threshold sample.

would naturally tend to increase. When fitting to $\Upsilon(R;R_0)$, $\eta$ is not affected by whether or not the concentration is fixed.

In summary, we have found that $\Upsilon(R;R_0)$ is the optimal statistic for parametric mass modelling given its insensitivity to the profile at small scales, with $R_0 = 0.25$–$0.5\,h^{-1}$ Mpc for the cluster masses used here, giving a reasonable compromise in reducing systematic error while retaining reasonable S/N on the recovered masses for the case discussed here, but in general the choice of $R_0$ will depend on the specific application one has in mind and on the scales to which the data can be considered relatively systematics-free. This statistic tends to slightly overestimate the mass due to the combination of two competing effects: the profile deviation from NFW on large scales and the neglected LSS contribution to the lensing signal. However, these effects are only very weakly dependent on the details of the profile, the mass and the cosmology, making them easy to calibrate out at the few per cent level using $N$-body simulations. This result is in stark contrast to the effect of small-scale systematics on the masses estimated from $\Delta\Sigma(R)$ (e.g. varying concentrations, and deviations from NFW due to intrinsic alignments and baryonic effects), which lead to larger systematic uncertainties in the recovered masses. These conclusions hold in cases where the NFW concentration is fixed. If it is allowed to vary, then the statistical errors will increase more than when using $\Upsilon(R;R_0)$ with a reasonable $R_0$, but systematic errors decrease, provided that the systematic errors in the lensing signal appear reasonably similar to a change in NFW concentration, which is not the case for several of the small-scale systematics in Section 2.3.

### 5.3 Example application with data

Here, we consider the maxBCG cluster lensing data in six scaled richness bins ($12 \leq N_{200} \leq 79$), which was previously used in Mandelbaum et al. (2008a) for joint estimation of the concentration–mass relation and the mass–richness relation. Here, we use several examples of fixed concentration–mass relations and several of the fitting methods considered in the previous sections to estimate the mass–richness relation, always with $\Omega_{\rm m} = 0.25$. This estimation as follows:

(i) We generate 200 bootstrap-resampled data sets to estimate the noise in the data. For this bootstrap procedure, the data are divided into 200 regions on the sky which are bootstrapped (rather than bootstrapping the individual lenses). More details on this procedure is given in Mandelbaum et al. (2008a).

(ii) For each data set, we separately fit the data in each richness bin for $M_{200b}$ assuming some $c_{200b}(M_{200b})$ relation and fit method, for each richness bin. The choice of fit method includes specifying the statistic to fit and the range of transverse separations to use. Thus, given logarithmic bins in transverse separation denoted by $i$ ($R_i$), data set $j$, richness bin $k$, statistic for a given fit method $\ell$

[denoted by $\Xi$ for $\Xi = \Delta\Sigma$ or $\Upsilon(R; R_0)$] and $c_{200b}(M_{200b})$ relation $m$, we use the Levenberg–Marquardt algorithm to separately minimize the $j \times k \times \ell \times m$ values of $\chi^2$ defined as follows:

$$\chi^2_{jk\ell m} = \sum_i \frac{\left(\Xi^{(\rm data)}_{jk\ell}(R_i) - \Xi^{(\rm model)}_{\ell m}(R_i)\right)^2}{\sigma^2(\Xi_{k\ell}(R_i))} \quad (21)$$

where we use $i$ such that $R_{\rm min,\ell} \leq R_i \leq R_{\rm max,\ell}$. The result of this procedure is a matrix with $j \times k \times \ell \times m$ values of $M_{200b}$, where in practice we use $j = 200$, $k = 6$, $\ell = 3$ and $m = 3$. The fit methods and concentration–mass relations are described in detail below.

(iii) The set of $k$ $M_{200b}$ values for a given data set ($j$), fit method ($\ell$) and concentration–mass relation ($m$) are used to fit for a power-law relation between scaled richness and halo mass:

$$M^{(\rm model)}_{200b} = \left[\left(M_{200b,20} \times 10^{14}\right)\ h^{-1}\,{\rm M}_\odot\right]\left(\frac{N_{200}}{20}\right)^{\gamma}. \quad (22)$$

This fit has two parameters: an amplitude $M_{200b,20}$ that is the mass at our pivot richness of $N_{200} = 20$ in units of $10^{14}\,h^{-1}\,{\rm M}_\odot$, and an exponent $\gamma$. We find the best-fitting values of $M_{200b,20}$ and $\gamma$ for each ($j$, $\ell$, $m$) by minimizing

$$\chi^2_{j\ell m} = \sum_k \frac{\left(M^{(\rm data)}_{200b,jk\ell m} - M^{(\rm model)}_{200b}(N_{200,k})\right)^2}{\sigma^2(M_{200b,\ell m})}. \quad (23)$$

The result is a matrix with $j \times \ell \times m$ values of $M_{200b,20}$ and $\gamma$.

(iv) We use the list of $j$ power-law fits for each bootstrap-resampled data set to estimate the mean and variance of $M_{200b,20}$ and $\gamma$ for a given combination of fit method $\ell$ and concentration–mass relation $m$.

We include $m = 3$ concentration–mass relations in our tests: a power law with

$$c_{200b} = 5\left(\frac{M_{200b}}{10^{14}\,h^{-1}\,{\rm M}_\odot}\right)^{-0.1}, \quad (24)$$

consistent with Mandelbaum et al. (2008a); a constant $c_{200b} = 4$ and a constant $c_{200b} = 7$. We examine the results for $\ell = 3$ fit methods: an extreme one assuming a small aperture for the cluster data, using $\Delta\Sigma(R)$ from 0.1 to 1 $h^{-1}$ Mpc; using $\Delta\Sigma(R)$ from 0.5 to 4 $h^{-1}$ Mpc and using $\Upsilon(R; R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc from 0.5 to 4 $h^{-1}$ Mpc, given its good performance on theoretical profiles and simulations in the previous sections. We consider the fit results first without and then with correction factors derived from Fig. 5. While we only have simulation-based correction factors for samples with three mean masses (which are threshold samples, not discrete mass bins as in this example) and two concentrations, we interpolate those results to derive approximate corrections for all the fits done in this section.

The final type of correction that we apply is a calibration factor that reduces the lensing signal calibration from Mandelbaum et al. (2008a) and Reyes et al. (2008) by 6 per cent. The reason for this correction is that for 30 per cent of the spectroscopic training set presented in Mandelbaum et al. (2008b) for calibration of photometric redshifts that are used to estimate the lensing signal, an incorrect photometric calibration was used when computing the photometric redshifts. We emphasize that this incorrect calibration was only used for the KPHOTOZ photometric redshifts, not for any other photometric redshift sample, and thus the lensing signal calibrations that are quoted for other photometric redshift methods in that paper are correct. As a result of this error, the calibrations from KPHOTOZ which were used for the data in Mandelbaum et al. (2008a) and Reyes et al. (2008) that we analyse here were 6 per cent too high, so we now apply a correction to the signal. We then present the results for the best-fitting masses after application of both the signal calibration correction and the simulation-based correction factors due to the mass estimation method.

Fig. 7 shows the observed signal for the lowest and highest richness bins for $0.1 < R < 4\,h^{-1}$ Mpc, and the theoretical signal from the fits. This theoretical signal is derived by taking the best-fitting mass–richness relation, evaluating it at the mean richness of the bins that are shown, and using the resulting mass and assumed concentration to define the theoretical signal. The fits did not only use the data shown on the plot, because the requirement of a power-law mass–richness relation means that the theoretical signal at the richness bins shown was also influenced by the data in all other bins.

For reference, given a best-fitting mass–richness relation from equation (22) with $M_{200b,20} = 1.55$ and $\gamma = 1.15$ (which is a typical value given the scatter between the results in Table 4), the combination with equation (24) gives a concentration–richness relation of

$$c_{200b} = 4.78\left(\frac{N_{200}}{20}\right)^{-0.115} \quad (25)$$

Thus, within our richness range of $12 \leq N_{200} \leq 79$, the concentrations vary from 5 to 4 as we move from the lowest to the highest richnesses. When we instead fix $c_{200b} = 4$ independent of mass, we lower the concentrations at the low $N_{200}$ end of the sample by 20 per cent, without changing the concentrations at the very high mass end. When we fix $c_{200b} = 7$, then we raise all the concentrations by a very significant amount, from ∼40 per cent increases at the low mass end to 75 per cent at the high-mass end. The results for the three concentration–mass relations and fitting methods are given in Table 4.

We begin by discussing the first fit method, using $\Delta\Sigma(R)$ from $0.1 < R < 1\,h^{-1}$ Mpc. As we have noticed in previous examples with the theoretical profiles and simulations, the results using these scales are highly sensitive to the assumed concentration–mass relation. We see that changing the assumed concentration among our three options leads to 50 per cent variation of the amplitude $M_{200b,20}$, significantly larger than the statistical errors on this parameter, when we do not impose corrections from the simulations. The exponent $\gamma$ undergoes 20 per cent changes, which are roughly consistent with the size of the $1\sigma$ statistical errors. The changes in this exponent can be easily understood as follows. First, if we change from the power-law concentration–mass relation in equation (24) to fixed $c_{200b} = 4$, then we are lowering the assumed concentration for all but the highest mass haloes. This means that, due to the typical concentration–mass anti-correlation when fitting $\Delta\Sigma$, the best-fitting masses should increase at the lower mass end. As a result, the best-fitting mass–richness relation becomes less steep. When we change to use a higher concentration $c_{200b} = 7$, then due to this concentration–mass anti-correlation, the best-fitting masses are significantly suppressed (which explains the large change in $M_{200b,20}$). Furthermore, this suppression is stronger at the higher mass end, where the difference between $c_{200b} = 7$ and equation (24) is most pronounced. This trend will tend to suppress $\gamma$, as is seen in the table.

When we impose corrections from the simulations to the results from the first fit method, we find that the variation in $M_{200b,20}$ and $\gamma$ when we change the assumed concentration is significantly reduced. However, there is still 30 per cent level variation, which may be ascribed to the fact that the scales that are used in this fit are quite prone to systematics such as intrinsic alignments and centroiding errors,

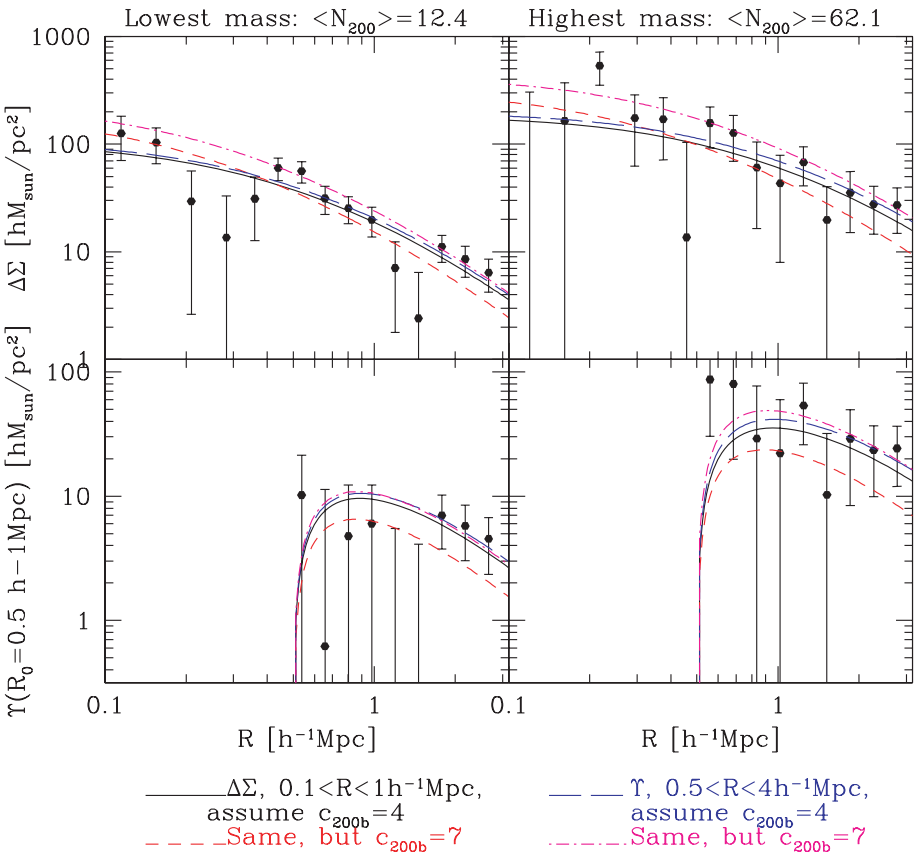


**Figure 7.** Observed lensing signal from Mandelbaum et al. (2008a) for stacked maxBCG clusters, presented as $\Delta\Sigma(R)$ and $\Upsilon(R;R_0)$ with $R_0 = 0.5\,h^{-1}$ Mpc in the top and bottom panels, respectively. We show the lowest (left) and highest (right) richness bins out of the six used for the analysis. In addition to the data with bootstrap error bars, we also show four theoretical signals labelled on the plot. Two of them were derived by fitting $\Delta\Sigma(R)$ using $R_{\rm min} = 0.1$ and $R_{\rm max} = 1\,h^{-1}$ Mpc with different assumed concentrations; the other two, by fitting $\Upsilon(R;R_0)$ with $R_0 = R_{\rm min} = 0.5$ and $R_{\rm max} = 4\,h^{-1}$ Mpc. The 6 per cent calibration correction described in the text has been applied. Because we required a power-law relationship between mass and richness, the best-fitting signals shown for these two bins were influenced by the data in the other richness bins (not shown).

**Table 4.** Results of power-law fits for a mass–richness relation using stacked maxBCG cluster lensing data, using three fit methods and three concentration–mass relations. First present the best-fitting masses; then, include corrections for the bias on the mass estimation from simulations (Fig. 5); finally, with both the simulation corrections and a 6 per cent decrease of the amplitude on the lensing signal, as described in the text.

| Fit method | $c_{200b}(M_{200b})$ | $M_{200b,20}$ | $\gamma$ | $M_{200b,20}$ | $\gamma$ | $M_{200b,20}$ | $\gamma$ |
|---|---|---|---|---|---|---|---|
| | | No correction | | Simulation correction | | Sim. and photo-$z$ corrections | |
| $\Delta\Sigma(R)$, | equation (24) | $1.64 \pm 0.20$ | $1.24 \pm 0.35$ | 1.31 | 1.10 | $1.19 \pm 0.10$ | $1.10 \pm 0.28$ |
| $0.1 < R < 1\,h^{-1}$ Mpc | $c_{200b} = 4$ | $1.58 \pm 0.15$ | $1.07 \pm 0.26$ | 1.14 | 1.01 | $1.04 \pm 0.09$ | $1.01 \pm 0.24$ |
| | $c_{200b} = 7$ | $1.01 \pm 0.08$ | $0.93 \pm 0.22$ | 1.16 | 0.98 | $1.06 \pm 0.09$ | $0.98 \pm 0.24$ |
| $\Delta\Sigma(R)$, | equation (24) | $1.72 \pm 0.13$ | $1.18 \pm 0.18$ | 1.56 | 1.14 | $1.44 \pm 0.10$ | $1.14 \pm 0.17$ |
| $0.5 < R < 4\,h^{-1}$ Mpc | $c_{200b} = 4$ | $1.70 \pm 0.12$ | $1.14 \pm 0.16$ | 1.51 | 1.11 | $1.40 \pm 0.10$ | $1.10 \pm 0.16$ |
| | $c_{200b} = 7$ | $1.52 \pm 0.10$ | $1.06 \pm 0.15$ | 1.46 | 1.11 | $1.35 \pm 0.09$ | $1.11 \pm 0.16$ |
| $\Upsilon(R;R_0)$, | equation (24) | $1.79 \pm 0.18$ | $1.20 \pm 0.24$ | 1.67 | 1.20 | $1.50 \pm 0.16$ | $1.21 \pm 0.24$ |
| $R_0 = 0.5\,h^{-1}$ Mpc, | $c_{200b} = 4$ | $1.81 \pm 0.18$ | $1.18 \pm 0.23$ | 1.75 | 1.16 | $1.56 \pm 0.16$ | $1.17 \pm 0.23$ |
| $0.5 < R < 4\,h^{-1}$ Mpc | $c_{200b} = 7$ | $2.02 \pm 0.19$ | $1.11 \pm 0.21$ | 1.73 | 1.17 | $1.56 \pm 0.16$ | $1.17 \pm 0.23$ |

which will affect the fits with different assumed concentrations in different ways. The simulation corrections can only correct for the fitting methods' different responses to a theoretical cluster lensing profile, not for their different responses to additional systematics that may be present in the data.

When we fit using $\Delta\Sigma(R)$ from $0.5 < R < 4\,h^{-1}$ Mpc, we find smaller variations in the (uncorrected) amplitude $M_{200b,20}$ of the mass–richness relation when we change the concentration–mass relation, at most 13 per cent, which is still problematic since it is close to twice the $1\sigma$ statistical error. (However, note that the fit $\chi^2$ are not sufficiently different to rule out any of these three models; the lensing data only weakly constrain the concentration.) The trends in $\gamma$ with $c_{200b}(M_{200b})$ have the same sign as when fitting using $\Delta\Sigma(R)$ from $0.1 < R < 1\,h^{-1}$ Mpc, but are less pronounced (11 per cent variation, slightly smaller than the $1\sigma$ statistical error). Because of the longer range in transverse separation, the statistical errors on the

fit parameters have become smaller, though we do not fully benefit from this fact due to the systematic uncertainties. We also note that for a given concentration–mass relation, such as equation (24), the amplitude $M_{200b,20}$ is increased by 4 per cent relative to the previous results. This increase may be due to systematics that decrease the signal on scales below $0.5\,h^{-1}$ Mpc, such as intrinsic alignments or centroiding errors. The fact that $\gamma$ has decreased relative to the $0.1 < R < 1\,h^{-1}$ Mpc results suggests that the change in masses is more significant at lower richness than at higher richness.

When we impose corrections from the simulations in Fig. 5 to the results of this second fit, we find that the total range of $M_{200b,20}$ and $\gamma$ values is quite small, roughly 7 and 3 per cent, respectively. This finding is encouraging: it suggests that we may be converging to a result that is more robust to small-scale systematics. Since the typical corrected mass from this fit method is 25 per cent higher than that for the fits using $0.1 < R < 1\,h^{-1}$ Mpc, we conclude that the fits that use those smaller scales may be significantly influenced by small-scale systematics.

Finally, we consider the results of fits to $\Upsilon(R;R_0)$ with $R_0 = 0.5$ using $0.5 < R < 4\,h^{-1}$ Mpc. First, we see that the statistical errors on fit parameters are larger than when using $\Delta\Sigma(R)$ for the same scales [50 per cent larger, comparable to or smaller than the errors when using $\Delta\Sigma(R)$ from $0.1 < R < 1\,h^{-1}$ Mpc]. This trend in the errors may seem inconsistent with the results in the simulations, which suggested ∼30 per cent increase in mass estimation statistical errors. However, the 50 per cent increase is for the power-law amplitude that comes from using six mass bins. On each individual mass bin, the mass uncertainties increase by 30 per cent when using $\Upsilon(R;R_0)$ with $R_0 = R_{\rm min} = 0.5$ and $R_{\rm max} = 4\,h^{-1}$ Mpc relative to $\Delta\Sigma(R)$ with $R_{\rm min} = 0.5$ and $R_{\rm max} = 4\,h^{-1}$ Mpc. Thus, the mass increase we see here for individual mass bins is consistent with that in the simulation. Second, the variation in the uncorrected $M_{200b,20}$ when we change the concentration–mass relation is 11 per cent, comparable to the $1\sigma$ errors, though we emphasize again that the variations in concentration that we have allowed are relatively extreme compared to what is seen in simulations. The variation in $\gamma$ is 7 per cent, more than a factor of 2 smaller than the statistical error. The sense of the change in $M_{200b,20}$ when changing $c_{200b}(M_{200b})$ is the opposite as when fitting to $\Delta\Sigma(R)$, as we have seen before in the simulations.

When we use the simulation results to correct these final fits that use $\Upsilon(R;R_0)$, we see that the corrections again reduce the spread in the best-fitting $M_{200b,20}$ and $\gamma$ values when we use different concentration–mass relations. The residual 4 per cent variation in both fit parameters is well below the statistical error. We note that the typical mass $M_{200b,20}$ at richness $N_{200} = 20$ has increased by 10 per cent relative to the fits using $\Delta\Sigma(R)$ on the same exact scales, even after the imposition of the correction from simulations in Fig. 5. We suggest that this change may result from low-level residual contamination of $\Delta\Sigma(R)$ due to systematics such as centroiding errors even for $R > R_{\rm min} = 0.5\,h^{-1}$ Mpc. Such contamination can, as we have shown, bias fits to $\Delta\Sigma(R)$ while not affecting fits to $\Upsilon(R;R_0)$. Thus, we adopt our mass normalization at the pivot richness $N_{200} = 20$ as $M_{200b,20}/(10^{14}\,h^{-1}\,{\rm M}_\odot) = 1.54 \pm 0.16$ (stat.) $\pm$ 0.06 (sys.), the mean of the values from the fits to $\Upsilon(R;R_0)$ with the different concentration–mass relations. This systematic error results from an uncertainty of 0.03 due to uncertainties in the mass estimation due to both the assumed and true profile, added in quadrature with the lensing signal calibration uncertainty of 0.05.

We now compare these results against the $M_{200b}(N_{200})$ relations determined in several previous papers. First, we compare against that from Mandelbaum et al. (2008a), which used these data to fit for a concentration–mass and mass–richness relation. Given that the best-fitting concentration–mass relation in that paper was quite similar to our equation (24), and that the fits in that paper used $\Delta\Sigma(R)$ from $0.5 < R < 3\,h^{-1}$ Mpc, we expect quite similar results to the results in this paper using $c_{200b}(M_{200b})$ from equation (24) and $\Delta\Sigma(R)$ from $0.5 < R < 4\,h^{-1}$ Mpc. In that paper, we found $M_{200b,20} = 1.55$ and $\gamma = 1.14$. The mass normalization is quite similar to what we quote here, because (a) in Mandelbaum et al. (2008a) the masses were reduced by approximately 10 per cent due to small-scale systematics [from the use of $\Delta\Sigma(R)$ rather than $\Upsilon(R;R_0)$], but (b) the lensing signal amplitude was too high by 6 per cent, as explained above, which raised the best-fitting mass by $1.06^{1.4}$, a 9 per cent difference.

Reyes et al. (2008) used the maxBCG cluster lensing data to estimate a mass–richness relation. That work used fits to $\Delta\Sigma(R)$ from 0.5 to $4\,h^{-1}$ Mpc assuming equation (24) for the concentration–mass relation, with the same source shape measurements, shear calibration and source redshift distribution calibration as in this paper. However, the richness range used in that paper was slightly different, since it used the entire public catalogue from the minimum $N_{200} = 10$ to the maximum-scaled richness. Furthermore, the binning into richness bins within the range that is shared by this work and that one was different. Finally, as for Mandelbaum et al. (2008a), they explicitly modelled the halo–halo term using the same halo model formalism and assumed mass–bias relation. Their result was a best-fitting mass-richness power-law with $M_{200b,20} = 1.42$ and $\gamma = 1.16$. Thus, the calibration is 8 per cent lower than the value we have adopted here, but this could be attributed to differences in richness ranges.

Finally, we compare against the fits to the maxBCG catalogue cluster lensing signal in Johnston et al. (2007). The differences in procedure compared to this paper are numerous. First, the richness range is different, because they use a non-public version of the catalogue that extends down to $N_{200} = 3$. They fit to $\Delta\Sigma(R)$ using $0.05 \leq R \leq 30\,h^{-1}$ Mpc, and allow the halo concentration and the amplitude of the LSS term to vary. They also use a model for BCG centroiding errors based on mock catalogues, and incorporate this model into their fitting routine to correct for the tendency of centroid errors to suppress the estimated masses. They explicitly include lognormal scatter on the mass–richness relation (with a strong prior in the fits). Finally, while they use the same galaxy shape measurements, they use different photometric redshifts, which we have shown in Mandelbaum et al. (2008b) leads to a calibration bias in the lensing signal of −15 per cent. Since we have found that the fitted masses when assuming an NFW profile scale like $\Delta\Sigma^{1.4}$, this bias in $\Delta\Sigma$ corresponds to a 20 per cent suppression of the masses. Thus, while they find $M_{200b,20} = 1.2$ and $\gamma = 1.3$ (for a spherical overdensity of $180\overline{\rho}$, which should only differ from our definition by several per cent), we compare against a corrected value of $M_{200b,20} = 1.5$. This result is within a few per cent of our value of $M_{200b,20} = 1.54$ that we have adopted here. Given the different richness range (which also contributes to the different value of $\gamma$) and the many other differences in fit procedure, the three per cent discrepancy is not of concern, and is comparable to our quoted systematic uncertainty.

## 6 CONCLUSIONS

In this paper, we have assessed the degree to which certain systematic errors in lensing measurement and methods of mass estimation can bias weak lensing cluster mass estimates. In brief, the challenges we considered included the following.

(i) Lensing calibration bias, which leads to changes in the mass $\propto \Delta\Sigma^{\eta}$ for $\eta$ typically in the range 1.4–2 depending on the radial range and fit method used for the parametric NFW mass fits [lower for $\Delta\Sigma(R)$ than for $\Upsilon(R;R_0)$], or $\propto \Delta\Sigma$ for the non-parametric mass estimates within a fixed physical aperture (or a steeper scaling when estimated the mass within some spherical overdensity radius) using $\zeta_c$ (Sections 2.3.1, 5.1.2 and 5.2.1).

(ii) Offsets of the identified BCG from the minimum of the cluster potential well (Section 2.2.3) were incorporated into the lensing profiles using a model from mock catalogues presented in Johnston et al. (2007). This model includes the effects of photometric errors in selecting the wrong BCG, and is therefore an overly conservative estimate in cases where the BCG can be unambiguously identified for all clusters or where X-ray data can precisely locate the cluster centre.

(iii) The effect of differences between an assumed NFW concentration and the true NFW concentration were studied using pure NFW lensing signals.

(iv) Differences in the halo profile relative to a pure NFW profile were studied using fits to density profiles from $N$-body simulations.

(v) The effects of mass from structures other than the cluster itself on the lensing signal were also studied using the signal from simulations, since we have not included only the mass that is virialized when computing $\Delta\Sigma$ in the simulations.

(vi) Contamination of the source sample by cluster member galaxies, intrinsic alignments of those member galaxies and baryonic effects on the halo density profile were considered to be included among the previous tests, namely changes in NFW concentration (in Section 5.1), changes in the inner region of the profile using variations of the $N$-body simulation outputs (in Section 5.2) and centroid offsets that modify the signal only in the inner regions of the cluster.

When fitting a parametric model (in our case, the NFW profile) to $\Delta\Sigma(R)$, with fixed concentration, we find that the uncertainties due to unknown true concentration plus changes in the lensing profile due to small-scale systematics yield systematic errors that range from a factor of 2 in mass (when only using small scales in the fits, e.g. 0.1–1 $h^{-1}$ Mpc) to tens of per cent (when using $R > 0.5\,h^{-1}$ Mpc) to several per cent (for $R > 2\,h^{-1}$ Mpc, which yields stable mass estimates but large statistical errors, and which may not be available for individual cluster lensing analyses due to limited telescope FOV). This level of systematic error occurred when allowing a relatively broad variation in concentration ($4 < c_{200b} < 7$), given the disagreement between simulations on the concentration–mass relation at high masses, the large lognormal scatter in this relation and other systematics such as baryonic effects discussed in Sections 2.2, 2.3 and 2.6. When using a narrower range in concentration, the systematic errors decreased comparably, but are still unacceptably large relative to what is needed for precise cosmological parameter constraints.

The addition of centroiding errors to the list of systematics we considered led to uniform suppression of the mass estimates of the order of tens of per cent (for $R_{\rm min} = 0.1\,h^{-1}$ Mpc). To completely avoid this suppression while fitting to $\Delta\Sigma(R)$ and ignoring the possibility of centroiding errors, we found it necessary to restrict the fits to $R_{\rm min} > 1\,h^{-1}$ Mpc. Generally, the addition of larger scales, out to $\sim 2r_{200b}$, is useful in minimizing the effects of small-scale systematics; going beyond that can lead to excessive contribution from LSS, which will bias the mass estimates if it is not modelled accurately. Allowing a variation in concentration in the fits is another way to reduce systematic error, at the expense of statistical errors that are increased by 45 per cent, but this scheme is not helpful when dealing with systematics that have a radial profile that does not mimic a change in concentration. $\Upsilon(R;R_0)$ is still more reliable at removing the impact of small-scale systematics on the lensing signal.

The aperture mass statistic $\zeta_c$ led to accurate estimates of projected masses, provided that either (a) the mass in the outer annulus was estimated rather than ignored, or (b) the mass in the outer annulus was ignored, but $R_{o1} \gg R_1$ (i.e. a large range of transverse separations was included in the first integral in equation 12). For many applications, such as the halo mass function, the quantity of interest is the 3D virial mass, for which a density profile must be assumed to do the conversion from the 2D projected mass within $R_1$. We found that uncertainty in the true density profile led to tens of per cent level biases in the 3D virial masses. The effect of centroiding errors was to uniformly suppress the aperture masses by ∼10–20 per cent depending on the halo mass, degree of centroiding errors and transverse separations used for the analysis; these biases were then propagated into the 3D enclosed mass estimates. The aperture mass-based estimates of the cluster virial mass were substantially noisier than fits to $\Delta\Sigma(R)$ using the same range of scales.

Finally, the new statistic we introduce here, $\Upsilon(R;R_0)$, removes the effect of small scales from the lensing signal, and gave superior performance over $\Delta\Sigma(R)$ when fitting an NFW profile to the cluster lensing signal. This statement is true not just for the basic tests with pure NFW profiles and profiles from simulations, but also when including the effects of centroiding errors. The increases in statistical error on the mass can be ∼40 per cent relative to fitting to $\Delta\Sigma(R)$ over the same scales. The residual systematic uncertainties after removal of an overall offset in the masses is of the order of several per cent, when fitting from $0.5 < R < 4\,h^{-1}$ Mpc, as demonstrated using SDSS maxBCG data. The effects of $\Upsilon(R;R_0)$ in decreasing systematic error are less dramatic when only small scales ($\leq 2\,h^{-1}$ Mpc) are used for the mass estimates; however, the residual systematics of the order of 10 per cent are still at least a factor of 2 smaller than when fitting to $\Delta\Sigma(R)$.

These conclusions also apply for individual cluster lensing analyses; however, we caution that in that case, we expect additional uncertainties in the true halo profile due to contamination by cluster member galaxies, the lognormal scatter in concentration at fixed mass, mergers, substructure, triaxiality and projection effects (Sections 2.2 and 2.3), so the systematic errors will tend to be larger than for stacked analyses using the same mass estimation method.

Next, we will briefly discuss the implications of our findings about mass estimation methods for several previously published cluster lensing studies. We begin with Okabe et al. (2009), which contains an analysis of circularly averaged cluster lensing data for 30 individual clusters by comparison with spherical models. They begin with direct fitting of the tangential shear profile to parametric models, including the NFW profile. These fits allow all model parameters to vary; for example, the NFW fits have two parameters, a mass and a concentration, unlike the cases we have considered here with a fixed concentration. Consequently, the estimated masses from the NFW model fits are unlikely to be strongly biased due to modelling assumptions, since the concentration is not fixed. However, they may still have some systematic bias due to the NFW profile not describing cluster profiles well, due to deviations in individual cluster profiles due to substructure, mergers and triaxiality, and possibly significant biases due to small-scale systematics such as contamination by cluster member galaxies and centroiding errors.

Okabe et al. (2009) also use the aperture mass statistic $\zeta_c$ to estimate $M_{\rm 2D}$, while neglecting the second term in equation (12)

and choosing the outer annulus such that it does not contain any significant structures. As we have seen here, the aperture mass statistic when including both terms properly leads to quite accurate projected mass estimates, or can yield results that are accurate at the several per cent level even without the second term provided that $R_1 \ll R_{o1}$. Given the scales that are accessible with the Subaru Suprime-Cam, and the typical cluster redshifts, we should compare against the top portion of Table 3, the rows with $(R_1, R_{o1}, R_{o2}) = (0.275, 1.1, 2)$ and $(0.5, 1.1, 2)\,h^{-1}$ Mpc. Those results suggest that for the most massive clusters, neglect of the second term may cause 15–20 per cent suppression of the projected masses. We find that the suppression is reduced to 5–10 per cent for more typical cluster masses of $10^{14}\,h^{-1}\,M_\odot$. Furthermore, as we have already seen, effects that suppress the signal in the inner cluster regions, such as centroiding errors and contamination by cluster member galaxies, can suppress the aperture masses at the ∼10 per cent level.

Hoekstra (2007) contains an analysis of cluster weak lensing data for 20 individual clusters. This work utilizes parametric mass estimates from the tangential shear distortion averaged in annuli, fitting to an NFW profile with fixed concentration–mass relation from *N*-body simulations using $0.25 < R < 1.5\,h^{-1}$ Mpc. In this case, we can assess systematic uncertainties as being somewhere between the results for $(R_{\rm min}, R_{\rm max}) = (0.25, 1)$ and $(0.25, 2)\,h^{-1}$ Mpc in Fig. 5. The figure suggests that uncertainties due to differences between the assumed and the true profile lead to ∼50 per cent variations in the estimated cluster halo masses. This variation may be manifested as significant noisiness in the mass estimates for a given true mass, as well as a mean bias if the true profiles (with the imposition of systematics such as contamination by cluster member galaxies) differ from the NFW profile with that assumed concentration–mass relation. This problem is in addition to other uncertainties in individual cluster mass estimates noted previously, such as LSS (for which they explicitly increase their error bars) and triaxiality.

Hoekstra (2007) also use the aperture mass statistic to estimate projected masses, $M_{\rm 2D}$, while estimating the second term in equation (12) due to the outer annulus using the best-fitting NFW model. In that case, we note that while Hoekstra (2007) do not miss mass by excluding the second term in the aperture mass calculation, their conversion from $M_{\rm 2D}(< R_1)$ to virial radii using spherical overdensities that can define the mass function will be strongly concentration-dependent. While Hoekstra (2007) claim that the fact that the masses from the fits to $\Delta\Sigma(R)$ and from the aperture mass calculation are consistent shows that their fitting procedure is unbiased, as discussed in Section 5.1.2 this claim is not true. The fact that Vikhlinin et al. (2009) use the cluster mass estimates from this work to calibrate their mass function constraints is therefore of concern, because of the possible biases due to these systematics in the signal and the large systematic scatter that we have found.

In summary, we believe that weak lensing is the best observational technique to robustly estimate cluster virial masses (regardless of their dynamical state) at the level required for precision cosmology. Given the small statistical errors of recent cluster abundance analyses, the cosmological constraints are already dominated by the systematic precision of the cluster mass determination (Vikhlinin et al. 2009). As we argue in this paper, current methods are inadequate for this purpose because they rely on the information from the inner parts of the cluster, which can be contaminated or modified due to a variety of effects discussed in this paper, and because they do not use numerical *N*-body simulations to calibrate their results. Our results suggest eliminating lensing information from scales below $R_0$ (for which we suggest the range $0.2 < R_0 < 0.5\,h^{-1}$ Mpc or about 15–25 per cent of the virial radius, as determined via an iterative procedure). Our proposed statistic for parametric estimates of cluster mass, the ADSD $\Upsilon(R; R_0)$, achieves this by construction, and is consequently more robust to many different systematics and to the details of the model to which the data are fitted, all of which are more problematic in the inner parts of the cluster. Use of $\Upsilon(R; R_0)$ to estimate cluster masses allows systematic errors to be reduced to the several per cent level, which is up to a factor of 10 smaller than when fitting to the lensing signal $\Delta\Sigma(R)$ itself, suggesting that for current and future data sets, $\Upsilon(R; R_0)$ should be the statistic of choice for parametric mass fitting to cluster weak lensing data. While we have focused on clusters in this paper, similar concerns about accurately determining the halo mass would arise also for smaller haloes. For these, the stellar component from the galaxy (and possibly the associated redistribution of the dark matter) would modify the mass distribution relative to predictions from pure *N*-body simulations in the inner parts, suggesting that eliminating the inner halo information by using $\Upsilon(R; R_0)$ could lead to more accurate mass determination of group and galaxy-type haloes as well.

## ACKNOWLEDGMENTS

We thank the anonymous referee for many useful comments. RM was supported for the duration of this project by NASA through Hubble Fellowship grant #*HST*-HF-01199.02-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. US is partly supported by the Swiss National Foundation under contract 200021-116696/1, Packard Foundation and WCU grant R32-2008-000-10130-0. TB acknowledges support by a grant of the German Academic Foundation during the initial phase of this project. RES acknowledges support from a Marie Curie Reintegration grant and the SNF.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

## REFERENCES

Abate A., Wittman D., Margoniner V. E., Bridle S. L., Gee P., Tyson J. A., Dell'Antonio I. P., 2009, ApJ, 702, 603
Abazajian K. et al., 2003, AJ, 126, 2081
Abazajian K. et al., 2004, AJ, 128, 502

Abazajian K. et al., 2005, AJ, 129, 1755
Abazajian K. N. et al., 2009, ApJS, 182, 543
Adelman-McCarthy J. K. et al., 2006, ApJS, 162, 38
Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634
Adelman-McCarthy J. K. et al., 2008, ApJS, 175, 297
Agustsson I., Brainerd T. G., 2006, ApJ, 644, L25
Baldauf T., Smith R. E., Seljak U., Mandelbaum R., 2010, Phys. Rev. D, 81, 063531
Barkana R., Loeb A., 2010, MNRAS, in press (doi:10.1111/j.1365-2966.2010.16587.x) (arXiv:0907.1102)
Becker M. R. et al., 2007, ApJ, 669, 905
Bildfell C., Hoekstra H., Babul A., Mahdavi A., 2008, MNRAS, 389, 1637
Biviano A., Girardi M., 2003, ApJ, 585, 205
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Borgani S., Kravtsov A., 2009, preprint (arXiv:0906.4370)
Bridle S. et al., 2009, Ann. Appl. Stat., 3, 6
Broadhurst T., Takada M., Umetsu K., Kong X., Arimoto N., Chiba M., Futamase T., 2005, ApJ, 619, L143
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Buote D. A., Gastaldello F., Humphrey P. J., Zappacosta L., Bullock J. S., Brighenti F., Mathews W. G., 2007, ApJ, 664, 123
Clowe D., Luppino G. A., Kaiser N., Henry J. P., Gioia I. M., 1998, ApJ, 497, L61
Clowe D., DeLucia G., King L., 2004, MNRAS, 350, 1038
Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Corless V. L., King L. J., 2007, MNRAS, 380, 149
Corless V. L., King L. J., 2009, MNRAS, 396, 315
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Diaferio A., Geller M. J., Rines K. J., 2005, ApJ, 628, L97
Dodelson S., 2004, Phys. Rev. D, 70, 023008
Dolag K., Bartelmann M., Perrotta F., Baccigalupi C., Moscardini L., Meneghetti M., Tormen G., 2004, A&A, 416, 853
Einasto J., 1965, Trudy Inst. Astrofiz. Alma-Ata, 5, 87
Eisenstein D. J. et al., 2001, AJ, 122, 2267
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Fahlman G., Kaiser N., Squires G., Woods D., 1994, ApJ, 437, 56
Faltenbacher A., Li C., Mao S., van den Bosch F. C., Yang X., Jing Y. P., Pasquali A., Mo H. J., 2007, ApJ, 662, L71
Fedeli C., Bartelmann M., Meneghetti M., Moscardini L., 2007, A&A, 473, 715
Finkbeiner D. P. et al., 2004, AJ, 128, 2577
Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, MNRAS, 387, 536
Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148
Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16
Gunn J. E. et al., 1998, AJ, 116, 3040
Guzik J., Seljak U., 2002, MNRAS, 335, 311
Heymans C. et al., 2006, MNRAS, 368, 1323
Hirata C., Seljak U., 2003, MNRAS, 343, 459
Hirata C. M. et al., 2004, MNRAS, 353, 529
Hirata C. M., Mandelbaum R., Ishak M., Seljak U., Nichol R., Pimbblet K. A., Ross N. P., Wake D., 2007, MNRAS, 381, 1197
Ho S., Lin Y.-T., Spergel D., Hirata C. M., 2009, ApJ, 697, 1358
Hoekstra H., 2001, A&A, 370, 743
Hoekstra H., 2003, MNRAS, 339, 1155
Hoekstra H., 2007, MNRAS, 379, 317
Hogg D. W., Finkbeiner D. P., Schlegel D. J., Gunn J. E., 2001, AJ, 122, 2129
Ivezić Ž. et al., 2004, Astron. Nachr., 325, 583
Johnston D. E. et al., 2007, preprint (arXiv:0709.1159)
Katgert P., Biviano A., Mazure A., 2004, ApJ, 600, 657
King L. J., Schneider P., Springel V., 2001, A&A, 378, 748
Kleinheinrich M. et al., 2005, A&A, 439, 513
Koester B. P. et al., 2007a, ApJ, 660, 221
Koester B. P. et al., 2007b, ApJ, 660, 239
Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128
Levenberg K., 1944, Q. Appl. Math., 2, 164
Limousin M. et al., 2007, ApJ, 668, 643
Lupton R. H., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., Yasuda N., 2001, in Harnden F. R., Jr, Primini F. A., Payne H. E., eds, ASP Conf. Ser. Vol. 238. Astronomical Data Analysis Software and Systems X. Astron. Soc. Pac., San Francisco, p. 269
Mandelbaum R. et al., 2005a, MNRAS, 361, 1287
Mandelbaum R., Tasitsiomi A., Seljak U., Kravtsov A. V., Wechsler R. H., 2005b, MNRAS, 362, 1451
Mandelbaum R., Hirata C. M., Ishak M., Seljak U., Brinkmann J., 2006a, MNRAS, 367, 611
Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006b, MNRAS, 372, 758
Mandelbaum R., Seljak U., Hirata C. M., 2008a, J. Cosmol. Astro-Part. Phys., 8, 6
Mandelbaum R. et al., 2008b, MNRAS, 386, 781
Mantz A., Allen S. W., Ebeling H., Rapetti D., 2008, MNRAS, 387, 1179
Marian L., Smith R. E., Bernstein G. M., 2009, ApJ, 698, L33
Marian L., Smith R. E., Bernstein G. M., 2010, ApJ, 709, 286
Marquardt D., 1963, SIAM J. Appl. Math., 11, 431
Massey R. et al., 2007, MNRAS, 376, 13
Merritt D., Graham A. W., Moore B., Diemand J., Terzić B., 2006, AJ, 132, 2685
Metzler C. A., White M., Norman M., Loken C., 1999, ApJ, 520, L9
Metzler C. A., White M., Loken C., 2001, ApJ, 547, 560
Milgrom M., 1983, ApJ, 270, 365
Naab T., Johansson P. H., Ostriker J. P., Efstathiou G., 2007, ApJ, 658, 710
Nagai D., Kravtsov A. V., Vikhlinin A., 2007, ApJ, 668, 1
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Neto A. F. et al., 2007, MNRAS, 381, 1450
Okabe N., Takada M., Umetsu K., Futamase T., Smith G. P., 2009, preprint (arXiv:0903.1103)
Pedersen K., Dahle H., 2007, ApJ, 667, 26
Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, AJ, 125, 1559
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes in C: The Art of Scientific Computing, 2nd edn. Cambridge University Press, Cambridge
Reyes R., Mandelbaum R., Hirata C., Bahcall N., Seljak U., 2008, MNRAS, 390, 1157
Richards G. T. et al., 2002, AJ, 123, 2945
Rines K., Diaferio A., 2006, AJ, 132, 1275
Rines K., Geller M. J., Kurtz M. J., Diaferio A., 2003, AJ, 126, 2152
Rines K., Diaferio A., Natarajan P., 2007, ApJ, 657, 183
Rozo E. et al., 2010, ApJ, 708, 645
Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19
Salucci P., Lapi A., Tonini C., Gentile G., Yegorova I., Klein U., 2007, MNRAS, 378, 41
Schmidt R. W., Allen S. W., 2007, MNRAS, 379, 209
Schmidt F., Rozo E., Dodelson S., Hui L., Sheldon E., 2009, Phys. Rev. Lett., 103, 051301
Schuecker P., Finoguenov A., Miniati F., Böhringer H., Briel U. G., 2004, A&A, 426, 387
Scoccimarro R., 1998, MNRAS, 299, 1097
Seljak U., 2000, MNRAS, 318, 203
Sheldon E. S. et al., 2004, AJ, 127, 2544
Siverd R. J., Ryden B. S., Gaudi B. S., 2009, preprint (arXiv:0903.2264)
Smith R. E., 2009, MNRAS, 400, 851
Smith J. A. et al., 2002, AJ, 123, 2121
Spergel D. N. et al., 2003, ApJS, 148, 175
Spergel D. N. et al., 2007, ApJS, 170, 377
Springel V., 2005, MNRAS, 364, 1105
Stoughton C. et al., 2002, AJ, 123, 485
Strauss M. A. et al., 2002, AJ, 124, 1810
Tucker D. L. et al., 2006, Astrono. Nachr., 327, 821

van den Bosch F. C., Weinmann S. M., Yang X., Mo H. J., Li C., Jing Y. P., 2005, MNRAS, 361, 1203
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691
Vikhlinin A. et al., 2009, ApJ, 692, 1060
Wright C. O., Brainerd T. G., 2000, ApJ, 534, 34
Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., Weinmann S. M., Meneghetti M., 2006, MNRAS, 373, 1159
York D. G. et al., 2000, AJ, 120, 1579
Zentner A. R., Rudd D. H., Hu W., 2008, Phys. Rev. D, 77, 043507
Zhang Y.-Y., Finoguenov A., Böhringer H., Kneib J.-P., Smith G. P., Kneissl R., Okabe N., Dahle H., 2008, A&A, 482, 451
Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2003, ApJ, 597, L9
Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, ApJ, 707, 354

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