The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses

Hoekstra, H.; Herbonnet, R.T.L.; Muzzin, A.V.; Babul, A.; Mahdavi, A.; Viola, M.; Cacciato, M.

Citation
Hoekstra, H., Herbonnet, R. T. L., Muzzin, A. V., Babul, A., Mahdavi, A., Viola, M., & Cacciato, M. (2015). The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses. *Monthly Notices Of The Royal Astronomical Society*, *449*(1), 685-714. Retrieved from https://hdl.handle.net/1887/49369

Version: Not Applicable (or Unknown)
License:
Downloaded from: https://hdl.handle.net/1887/49369

Note: To cite this publication please use the final published version (if applicable).

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **449,** 685–714 (2015) doi:10.1093/mnras/stv275

The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses⋆

Henk Hoekstra,[1]† Ricardo Herbonnet,[1] Adam Muzzin,[1] Arif Babul,[2] Andi Mahdavi,[3] Massimo Viola[1] and Marcello Cacciato[1]

[1]*Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, the Netherlands*
[2]*Department of Physics and Astronomy, University of Victoria, 3800 Finnerty Rd, Victoria, BC V8P 5C2, Canada*
[3]*Department of Physics and Astronomy, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA*

Accepted 2015 February 8. Received 2015 February 6; in original form 2014 October 14

ABSTRACT

Masses of clusters of galaxies from weak gravitational lensing analyses of ever larger samples are increasingly used as the reference to which baryonic scaling relations are compared. In this paper we revisit the analysis of a sample of 50 clusters studied as part of the Canadian Cluster Comparison Project. We examine the key sources of systematic error in cluster masses. We quantify the robustness of our shape measurements and calibrate our algorithm empirically using extensive image simulations. The source redshift distribution is revised using the latest state-of-the-art photometric redshift catalogues that include new deep near-infrared observations. None the less we find that the uncertainty in the determination of photometric redshifts is the largest source of systematic error for our mass estimates. We use our updated masses to determine b, the bias in the hydrostatic mass, for the clusters detected by *Planck*. Our results suggest $1 - b = 0.76 \pm 0.05\,(\mathrm{stat}) \pm 0.06\,(\mathrm{syst})$, which does not resolve the tension with the measurements from the primary cosmic microwave background.

Key words: cosmology: observations – dark matter.

1 INTRODUCTION

The observed number density of clusters of galaxies as a function of mass and redshift depends sensitively on the expansion history of the Universe and the initial conditions of the density fluctuations. Comparison with predictions from a model of structure formation can thus constrain cosmological parameters, such as the mean density Ω_m and the normalization of the matter power spectrum σ_8 (e.g. Bahcall & Fan 1998; Henry 2000; Reiprich & Böhringer 2002; Henry et al. 2009), or the dark energy equation-of-state w (e.g. Vikhlinin et al. 2009; Mantz et al. 2010, 2015). For a recent review see Allen, Evrard & Mantz (2011).

The fact that the observations do not provide actual cluster counts as a function of mass, but rather the number density of objects with certain observational properties, such as the number of red galaxies or the X-ray flux within a given aperture, complicates a direct comparison with predictions: the cosmological interpretation requires knowledge of the selection function and the scaling relation between the observable and the underlying mass. Furthermore, scaling relations typically have intrinsic scatter that also needs to be measured, or at least accounted for.

One way forward is to simulate the observable properties of clusters, but the complex non-linear physics involved limits the fidelity of such approaches, at least for the moment. Therefore direct estimates of the clusters masses are needed. This can be achieved through dynamical analyses, such as the measurement of the motion of cluster members, or by measuring the temperature of the hot intracluster medium (ICM). However, in both cases the cluster is assumed to be in equilibrium, which is generally not a valid assumption. For instance, simulations suggest that hydrostatic X-ray masses are biased low (e.g. Rasia et al. 2006; Nagai, Vikhlinin & Kravtsov 2007; Lau, Kravtsov & Nagai 2009).

A more direct probe of the (dark) matter distribution would be preferable, which is provided by the gravitational lensing distortion of background galaxies: the gravitational potential of the cluster perturbs the paths of photons emitted by these distant galaxies, resulting in a slight, but measurable, coherent distortion. This in turn provides a direct measurement of the gravitational tidal field, which can be used to directly infer the projected mass distribution. Note, however, that the comparison to baryonic tracers does typically depend on the assumed geometry of the cluster. For a recent review of the use of gravitational lensing to study cluster masses and density profiles, we refer the reader to Hoekstra et al. (2013).

⋆Based on observations from the Canada–France–Hawaii Telescope, which is operated by the National Research Council of Canada, le Centre National de la Recherche Scientifique and the University of Hawaii.

†E-mail: hoekstra@strw.leidenuniv.nl

© 2015 The Authors
Published by Oxford University Press on behalf of the Royal Astronomical Society

The sizes of cluster samples are increasing rapidly thanks to wide-angle surveys at various wavelengths, especially at millimetre wavelengths thanks to the inverse-Compton scattering of cosmic microwave background (CMB) photons off hot electrons in the ICM, the Sunyaev–Zel'dovich effect (SZE; Sunyaev & Zeldovich 1972). In fact, the lack of calibrated scaling relations is currently the dominant limitation of blind searches that exploit this effect, such as those carried out using the South Pole Telescope (Reichardt et al. 2013) or the Atacama Cosmology Telescope (Hasselfield et al. 2013). The importance of accurate mass calibration is furthermore highlighted by the tension between the cosmological parameters determined from the primary CMB measured by *Planck* (Planck Collaboration XVI 2014b) and those inferred from the cluster counts (Planck Collaboration XX 2014c).

Fortunately it is not necessary to determine masses for all clusters, but instead it is sufficient to calibrate the appropriate scaling relation and its scatter. However, doing so still requires substantial samples of clusters for which weak lensing masses need to be determined. Even for the most massive clusters the uncertainty in the projected mass is ∼10 per cent. The triaxial nature of cluster haloes, however, leads to an additional intrinsic scatter of ∼15–20 per cent (e.g. Corless & King 2007; Meneghetti et al. 2010; Becker & Kravtsov 2011). Hence to calibrate the normalization of a scaling relation to a few per cent requires a sample of 50 or more clusters.

To examine the relation between the baryonic properties of clusters and the underlying matter distribution the Canadian Cluster Comparison Project (CCCP) started with the study of archival observations of 20 clusters of galaxies, described in Hoekstra (2007) and Mahdavi et al. (2008). This sample was augmented by observations of an additional 30 clusters with $0.15 < z < 0.55$ with the Canada–France–Hawaii Telescope (CFHT). A detailed description of the sample can be found in Hoekstra et al. (2012, hereafter H12) and Mahdavi et al. (2013). The comparison to the X-ray properties, presented in Mahdavi et al. (2008) and Mahdavi et al. (2013) confirmed the predictions from numerical simulations that the hydrostatic mass estimates are biased low.

Other groups have carried out similar studies. The Local Cluster Substructure Survey (LoCuSS) used the Subaru telescope to carry out a weak lensing study of 50 clusters, with the most recent results presented in Okabe et al. (2013). A thorough analysis of a sample of 51 clusters was presented by Weighing the Giants (WtG; Applegate et al. 2014; von der Linden et al. 2014a). For a large fraction of the clusters the latter study also obtained photometric redshifts for the sources (Kelly et al. 2014). Most recently Umetsu et al. (2014) presented results for a sample of 20 massive clusters, 17 of which were observed by WtG. In general there is significant overlap as these studies all target massive well-known clusters of galaxies. This was exploited by Applegate et al. (2014) who compared the masses from the various studies. Although they find an excellent correlation with the results from H12, the CCCP masses are on average ∼20 per cent lower than their estimates. This is much larger than the statistical uncertainties and warrants further investigation. This is the aim of this paper.

A correct interpretation of the inferred weak lensing signal relies on accurate shape measurements and knowledge of the redshifts of the sources used in the analysis. The former has been examined quite extensively over the past decade, for instance in several blind studies using simulated images (Heymans et al. 2006; Massey et al. 2007; Bridle et al. 2010; Kitching et al. 2012; Mandelbaum et al. 2014a). The results of such simulations have been used to quantify the biases in shape measurements, but the sensitivity of the calibration to the input of the simulations has not been investigated in much detail. However, thanks to an improved understanding of the sources of bias, and how they propagate (e.g. Massey et al. 2013; Miller et al. 2013; Semboloni et al. 2013; Viola, Kitching & Joachimi 2014), it has become evident that a correct interpretation of these simulations depends critically on how well they match the specific observations under consideration. In Section 2, we examine the importance of the fidelity of the image simulations. To calibrate our method, we create an extensive set of images, varying a number of input parameters.

Another important source of uncertainty is the redshift distribution of the sources. In Section 3, we present our photometric redshift estimates based on measurements in 29 bands in the Cosmic Evolution Survey (COSMOS) field (Capak et al. 2007; Scoville et al. 2007) including new deep observations in five near-infrared (NIR) bands from UltraVISTA (McCracken et al. 2012). We also revisit the issue of contamination by cluster members. We present new weak lensing mass estimates in Section 4 and use these in Section 5 to calibrate the hydrostatic masses used by Planck Collaboration XX (2014c) to infer cosmological parameters. Throughout the paper we assume a cosmology with $\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\,h_{70}$ km s^{-1} Mpc^{-1}.

2 CALIBRATION OF SHAPE MEASUREMENTS

The measurement of the shapes of small, faint galaxies is one of two critical steps in order to derive accurate cluster masses from weak gravitational lensing, the other step involving knowledge of the source redshift distribution. We discuss the latter in Section 3 and focus first on the algorithms used to measure galaxy shapes. Most studies to date have focused on the correction for the blurring by the point spread function (PSF), which leads to rounder images (due to the size of the PSF) and preferred orientations (if the PSF is anisotropic). An incomplete correction for the former leads to a multiplicative bias μ and a residual in the latter to an additive bias c; the observed shear and true shear are thus related by (e.g. Heymans et al. 2006)

$$\gamma_i^{\rm obs} = (1+\mu)\gamma_i^{\rm true} + c, \quad (1)$$

where we implicitly assumed that the biases are the same for both shear components. For cosmic shear studies the additive bias is a major source of concern because the (residual) PSF introduces power on relevant scales (e.g. Hoekstra 2004). For cluster lensing the additive bias is less important because the measurement of cluster masses involves the azimuthally averaged tangential shear and PSF patterns largely average out for our data. We study the residual additive bias in Appendix A and find that we can indeed ignore the residuals arising from PSF anisotropy in our analysis.

One approach to recover the true galaxy shape is to assume a suitable model for the galaxy light distribution, which is subsequently sheared, convolved with the PSF and pixellated. The model parameters are varied until a best fit to the data is obtained. This has the advantage that the detrimental effects of the PSF (and other instrumental biases) can be incorporated into a Bayesian framework (e.g. Miller et al. 2013; Bernstein & Armstrong 2014). The challenge, however, is to use a model that provides a good description of the galaxies, while having a limited number of parameters in order to avoid overfitting. A model that is too rigid will lead to model bias (e.g. Bernstein 2010), whereas a model that is too flexible tends to fit noise in the images (e.g. Kacprzak et al. 2012). Furthermore, accurate priors for the size and ellipticity distributions (and any other parameter entering the model) are required to obtain an unbiased estimate for the shear.

An alternative approach, which we use here, involves measuring the moments of the galaxy images, which are subsequently corrected for the PSF. The shapes can be quantified by the polarization

$$e_1 = \frac{I_{11} - I_{22}}{I_{11} + I_{22}}, \text{ and } e_2 = \frac{2I_{12}}{I_{11} + I_{22}}, \quad (2)$$

where the quadrupole moments I_{ij} are given by

$$I_{ij} = \frac{1}{I_0} \int \mathrm{d}^2\boldsymbol{x}\, x_i x_j\, W(\boldsymbol{x}) f(\boldsymbol{x}), \quad (3)$$

where $f(\boldsymbol{x})$ is the observed galaxy image, $W(\boldsymbol{x})$ a suitable weight function to suppress the noise and I_0 the weighted monopole moment. In the case of unweighted moments, I_0 corresponds to the flux and the correction for the PSF is straightforward as the PSF corrected moments are given by[1]

$$I_{ij}^{\mathrm{true}} = I_{ij}^{\mathrm{obs}} - I_{ij}^{\mathrm{PSF}}, \quad (4)$$

i.e. one only needs to subtract the moments of the PSF from the observed moments. The result provides an unbiased estimate of the polarization. However, the change in polarization δe_i due to a shear $\delta\gamma_i$ depends on the unsheared shape e_{int}: it is more difficult to change the shape of an object that is already elongated. This response is quantified by the polarizability P^γ, defined such that $\delta e_i = P^\gamma \delta\gamma_i$. As the shear is obtained from an ensemble of galaxies, an unbiased estimate thus requires knowledge of the intrinsic ellipticity distribution (e.g. Viola et al. 2014).

Unfortunately real data contain noise and thus unweighted moments are not practical. To suppress the effects of noise a weight function needs to be chosen, ideally matched to the size and shape of the galaxy image. However, as discussed in e.g. Massey et al. (2013) and Semboloni et al. (2013), this complicates matters as the correction for the PSF now involves higher order moments, which themselves are affected by noise. Limiting the expansion in moments is similar to the model bias in fitting methods.

Recent studies using simulated data have shown that multiplicative biases depend strongly on the signal-to-noise ratio (SNR; Bridle et al. 2010; Kitching et al. 2012; Miller et al. 2013) with some hints already present in the dependence of the bias on magnitude in Massey et al. (2007). As the origin of this bias is now better understood, it has also become clear that the performance of a particular algorithm will depend on the data it is applied to. Hence the performance evaluation, such as the determination of the bias that one wishes to correct for, depends on the input of the simulations: if the input does not match the actual data, the inferred bias may be different from the actual value. Although SNR is the most critical parameter, the bias may also depend on the galaxy profile, or the size and ellipticity distributions (e.g. Kacprzak et al. 2012; Melchior & Viola 2012; Miller et al. 2013; Viola et al. 2014). Unless the fidelity of the simulation can be somehow guaranteed, the sensitivity of a method to the input parameters needs to be quantified and the uncertainties propagated.

In this paper we focus on the commonly used KSB method developed by Kaiser, Squires & Broadhurst (1995) and Luppino & Kaiser (1997) with corrections provided in Hoekstra et al. (1998) and Hoekstra, Franx & Kuijken (2000). It was used to determine masses for the CCCP sample in Hoekstra (2007) and H12; we refer the interested reader to these papers for more details. The object detection is done using the hierarchical peak finder described in Kaiser et al. (1995), which gives an estimate for r_{g}, the Gaussian scale radius of the object. This value is used to compute the weighted moments, which are corrected following Hoekstra et al. (1998). In addition we also compute σ_e, the uncertainty in the polarization, which is approximately $\propto 1/\nu$, where ν is the SNR of the detection (Hoekstra et al. 2000). This allows us to down-weight the noisy galaxies and we therefore estimate the average shear for an ensemble of galaxies as

$$\langle\gamma_i\rangle = \frac{\sum w_i e_i/\tilde{P}^\gamma}{\sum w_i}, \text{ with } w_i = \frac{1}{\langle\epsilon^2\rangle + \left(\frac{\sigma_e}{\tilde{P}^\gamma}\right)^2}, \quad (5)$$

where $\langle\epsilon^2\rangle$ is the intrinsic variance of the galaxy ellipticity components. This is the dominant source of uncertainty for the shear for bright objects, and we adopt a value of $\langle\epsilon^2\rangle^{1/2} = 0.25$ (Hoekstra et al. 2000). In our image simulations we vary the input ellipticity distribution (see Section 2.1), which in principle would require adjusting the value for $\langle\epsilon^2\rangle^{1/2}$ accordingly to optimally weight objects. However, for simplicity we keep it fixed when we quantify the multiplicative bias.

2.1 Input galaxy properties

To populate our image simulations, we use a sample of galaxies for which morphological parameters were measured from resolved *F*606*W* images from the GEMS survey (Rix et al. 2004). These galaxies were modelled as single Sérsic models with GALFIT (Peng et al. 2002) and for our study we use the half-light radius, magnitude and Sérsic index n. We only consider galaxies fainter than $m_r = 20$ because bright objects might cause unrealistic features in the simulated images. Excluding these does not impact our results as we do not use them in our source sample anyway: to measure the lensing signal we use galaxies with $22 < m_r < 25$.

The resulting number density of galaxies as a function of apparent magnitude m_r is presented in Fig. 1 (solid black histogram). The results suggest a power law for the counts, where the flattening for

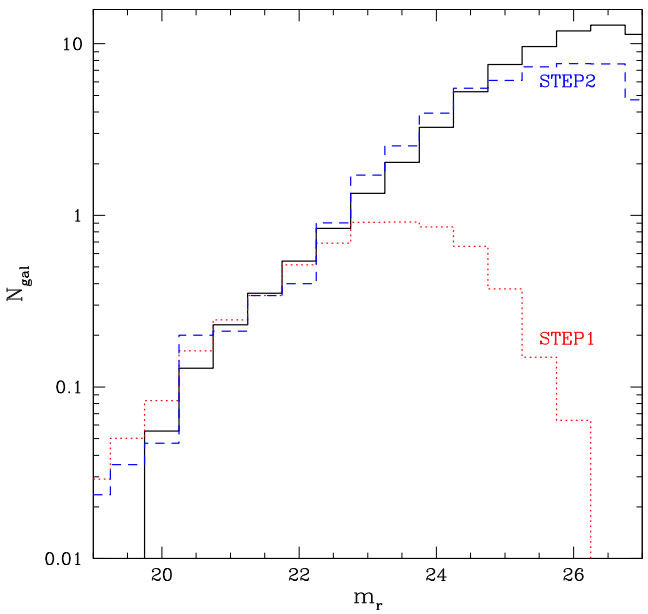

Figure 1. Number counts as a function of r-band magnitude for our image simulations (solid black histogram), STEP1 (dotted red histogram) and STEP2 (dashed blue histogram). The counts were normalized in the range $20 < m_r < 22$ and we adjusted the magnitudes of the STEP simulations for the use of different filters.

[1] We assume that the measurement is centred on the location where the dipoles vanish.

$m_r > 25.5$ is caused by incompleteness of the input catalogue. In principle, faint unresolved galaxies can affect shape measurements of brighter galaxies through modulation of the background noise and blending. In Section 2.3 we therefore examine the need to include fainter galaxies in the simulations.

H12 based their assessment of the accuracy of the shape measurements on the results from the Shear TEsting Programmes (STEP; Heymans et al. 2006; Massey et al. 2007). These were blind challenges with the aim to benchmark the performance of shape measurement algorithms, especially for cosmic shear studies. In both cases the implementation used by H12 performed well, with an average multiplicative bias of ∼2 per cent. As a consequence, H12 ignored the multiplicative bias in their mass estimates.

STEP1 (Heymans et al. 2006) simulated CFHT observations in the *I* band with an integration time of 3600 s, which should be quite comparable to our data (a total integration time of 1 h in the r' band using CFHT). The red dotted histogram in Fig. 1 shows the galaxy number counts that were used as input for STEP1 (converted to *r* band assuming a mean galaxy colour of $r - i = 1$). The counts are normalized such that the sum is the same for all three examples in the range $20 < m_r < 22$. Comparison to the GEMS catalogue shows that STEP1 lacks the faint galaxies that are present in real observations, even if they are unresolved. As we will show in Section 2.3, this leads to a significant underestimate of the multiplicative bias for the actual CCCP data.

STEP2 (Massey et al. 2007) simulated images that would be obtained with an exposure time of 40 min in good conditions with SuprimeCam on Subaru. Given the larger aperture and throughput of Subaru compared to CFHT, this corresponds to a total exposure time that is ∼4 times longer than the CCCP data. The input galaxy images were based on a `shapelet` decomposition of resolved galaxy images from the *Hubble Space Telescope* (*HST*) COSMOS survey (Scoville et al. 2007), as described in Massey et al. (2004). As a result the simulations should better capture the complex morphologies of real galaxies. The number counts, shown by the blue dashed histogram in Fig. 1, match the GEMS input counts much better than STEP1, although incompleteness occurs at $m_r \sim 24.5$.

In addition to the magnitudes, the GEMS catalogue provides values for the Sérsic index of the galaxies, as well as their half-light radii and ellipticities. The use of Sérsic profiles to describe the galaxies may limit the fidelity of the simulations (see Kacprzak et al. 2014, for a study of the biases that may arise). We examine the bias as a function of Sérsic index in Section 2.3 and find that our shape measurement algorithm is not particularly sensitive to the profile, especially when compared to other sources of bias. We therefore expect that the difference with using realistic galaxy morphologies is small. One of the aims of the third gravitational lensing accuracy testing challenge[2] (GREAT3; Mandelbaum et al. 2014b) is to compare the results of shape measurement methods on postage stamps of actual (PSF corrected) *HST* observations and the corresponding Sérsic fits.

The different parameters describing the galaxies are jointly sampled from the GEMS catalogue to account for their intrinsic correlations (e.g. brighter galaxies are on average larger). However, we do not use the ellipticities provided by Rix et al. (2004), because of concerns that these do not match our data, as discussed in Section 2.2. Instead we use a parametric description, which allows us to investigate the role of the ellipticity distribution. We assign ellipticities[3] ϵ that are drawn from a Rayleigh distribution given by

$$P(\epsilon, \epsilon_0) = \frac{\epsilon}{\epsilon_0^2} \mathrm{e}^{-\epsilon^2/2\epsilon_0^2}, \quad (6)$$

where the value of ϵ_0 determines the width of the distribution, as well as the average $\langle\epsilon\rangle = \epsilon_0\sqrt{\pi/2}$. We need to truncate the distribution because the ellipticity cannot exceed unity, and also because galaxy discs have a finite thickness. We therefore set $P(\epsilon, \epsilon_0) = 0$ if $\epsilon > 0.9$. We assume that the ellipticity distribution is independent of other galaxy properties, whereas e.g. van Uitert et al. (2011) did observe different distributions for early- and late-type galaxies. Given the accuracy we require here, we find that this assumption does not impact our results. The ellipticity distribution of the GEMS catalogue matches that of $\epsilon_0 = 0.35$ for moderate ellipticities (Melchior & Viola 2012) which is indicated by the black curve in Fig. 2. We also show input ellipticity distributions for $\epsilon_0 = 0.1$ (red dotted curve), and $\epsilon_0 = 0.5$ (blue dashed curve). For comparison we also show the input ellipticity distribution used by STEP1 (Heymans et al. 2006), which peaks at very low ellipticities.

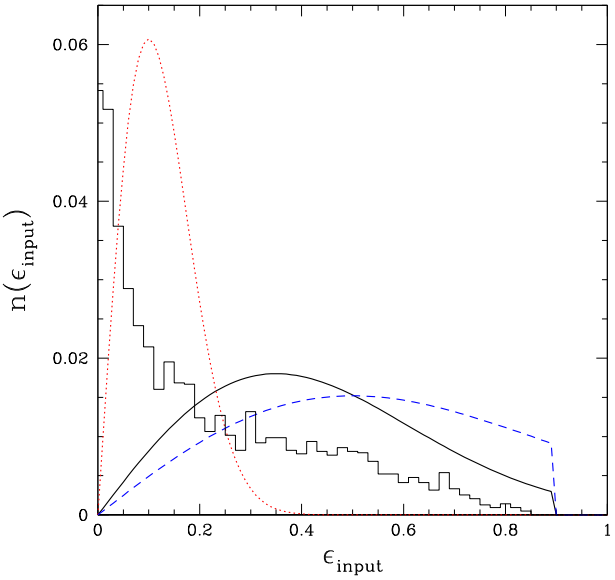

Figure 2. Input ellipticity distributions, described by a Rayleigh distribution truncated at $\epsilon = 0.9$, for $\epsilon_0 = 0.35$ (thick black curve), $\epsilon_0 = 0.1$ (red dotted curve) and $\epsilon_0 = 0.5$ (blue dashed curve). For reference we also show a histogram of ellipticities from one of the STEP1 simulations.

2.2 Description of the simulations

To create the images we use GALSIM (Rowe et al. 2014), a publicly available code that was developed for GREAT3 (Mandelbaum et al. 2014b). The main input is a list of galaxies with a position, flux, half-light radius, Sercic index and ellipticity, from which sheared images are computed. We limit the sample to objects with $0.5 < n < 4.2$ because of limitations of the version of GALSIM we used. To create the simulated images we draw objects from the GEMS catalogue. Given the limited number of galaxies observed by GEMS, objects typically appear multiple times in the simulation, but with a different ellipticity and orientation.

[2] http://www.great3challenge.info

[3] The ellipticity is defined as $(a - b)/(a + b)$, with a and b the major and minor axes, respectively. The polarization for such a galaxy would be $\sim(a^2 - b^2)/(a^2 + b^2)$

The weak lensing analysis of CCCP Megacam data is done on stacks with a total integration time of 1 h each. For each cluster two such stacks are observed, which are merged at the catalogue stage (see H12 for details). Our image simulations therefore assume the same noise level as observed in these data. To simulate the observed data we also need to provide a realistic PSF, for which we use a circular Moffat profile with a full width at half-maximum (FWHM) = 0.67 arcsec and $\beta = 3.5$. This resembles our observations of Abell 1835, which are typical for the CCCP sample, and we also adopt the noise level observed in these data. We include a low number of stars to measure the PSF in the images. In Appendix B, we quantify the impact of realistic star densities. We find that the observed star densities in the CCCP data are sufficiently low that they do not impact the results.

We create pairs of images where the galaxies are rotated by 90 degrees in the second image to reduce the noise due to the intrinsic ellipticity distribution (see e.g. Massey et al. 2007): by construction the mean intrinsic ellipticity when both are combined is zero. We analyse the images separately and thus, due to noise in the images, this is no longer exactly true, especially for faint galaxies. The input shears typically range from −0.06 to 0.06 in steps of 0.01 (for both components), yielding 169 image pairs for each ellipticity distribution. Each image has a size of 10 000 by 10 000 pixels, with a pixel scale of 0.185 arcsec, the same as our MegaCam data. This results in a sample of $\sim 10^7$ galaxies with $20 < m_r < 25$ for each value of ϵ_0. To examine the dependence on seeing and PSF anisotropy we create somewhat smaller sets, consisting of 49 pairs of images.

We analyse these images in the same way as the CCCP data. Fig. 3 shows the distribution of observed polarizations and half-light radii from the actual data (solid histograms) and simulated data (dotted and dashed histograms). We reproduce the magnitude distribution (not shown) and the size distribution for galaxies fainter than $m_r = 22$. As shown in the top-right panel of Fig. 3, there are many simulated bright galaxies that have half-light radii that are large compared to our CCCP data. For these galaxies the polarizations are significantly smaller than the distribution with $\epsilon_0 = 0.35$ found by Melchior & Viola (2012). Although the input catalogue is based on *HST* data, we suspect that the use of GALFIT may give too much weight to the outer regions of the galaxies, which are down-weighted in moment-based methods. This highlights the difficulty in establishing the input ellipticity distribution, which remains rather uncertain. For the main sample of sources, with $22 < m_r < 25$, we find a good match if we adopt $\epsilon_0 = 0.25$. We take this value as our reference in the remainder of this paper. We will conservatively assume that $0.15 < \epsilon_0 < 0.3$ when we estimate systematic uncertainties in the empirical bias correction.

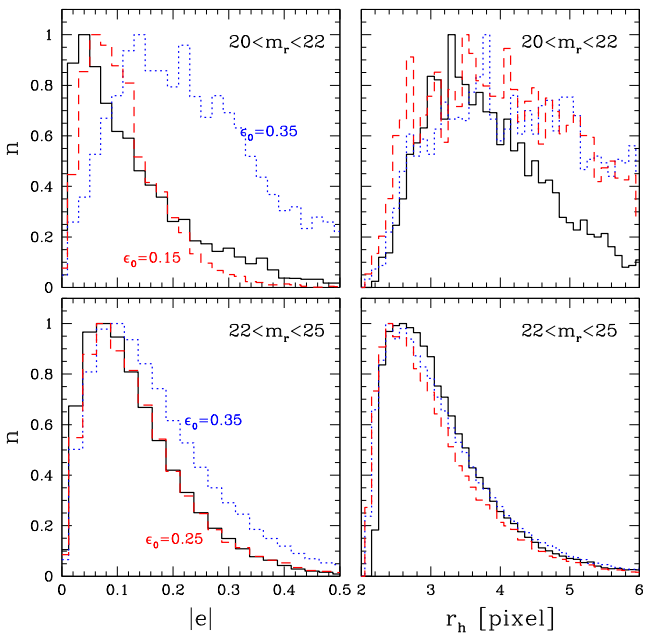

Figure 3. Comparison of the simulations and the actual CCCP data for Abell 1835 (black lines) for bright and faint sources. The left-hand panels show the observed polarizations, i.e. uncorrected for the PSF. We find that the ellipticity distribution for bright galaxies ($20 < m_r < 22$) is best approximated if we take $\epsilon_0 = 0.15$, whereas the ellipticity distribution for fainter galaxies ($22 < m_r < 25$) requires a higher value of $\epsilon_0 = 0.25$. For reference we show $\epsilon_0 = 0.35$ which is ruled out by the data. As shown in the right-hand panels, the simulations match the observed distribution of half-light radii fairly well.

2.3 Multiplicative bias as a function input parameters

As the underlying ellipticity distribution remains uncertain, we start by examining the average bias of KSB as a function of ϵ_0 for galaxies with $20 < m_r < 25$, i.e. the range in magnitude of the sources used in the CCCP analysis by H12. To detect objects we use the hierarchical peak finder described in Kaiser et al. (1995), which is the default algorithm in our analysis. The main difference between the various implementations of the KSB algorithm is the way the shear polarizability P^γ is estimated. As the observed values are noisy, H12 used a parametric fit to average values as a function of size for different magnitude bins (also see Heymans et al. 2006, for a concise description). The red squares in Fig. 4 show the bias as a function of ϵ_0 for this implementation of KSB. The bias changes by 0.04, which corresponds to a relative change of ∼40 per cent, over

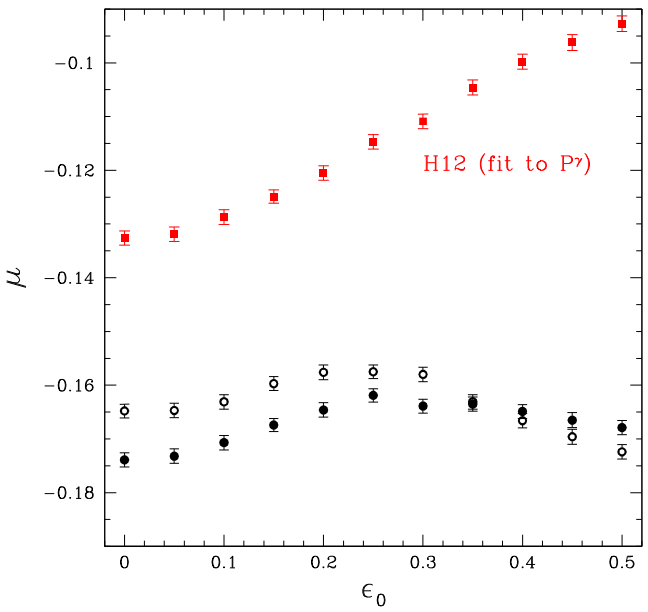

Figure 4. Multiplicative bias for sources with $20 < m_r < 25$ as a function of ϵ_0 (see text). The red squares show the results for the implementation of KSB used in H12. The black points show the results when we use the observed estimate for P^γ for individual galaxies. Although the bias is larger in the latter case, it depends only weakly on ϵ_0, especially in the relevant range of 0.15–0.3. The open circles indicate the bias if SExtractor is used instead of HFINDPEAKS.

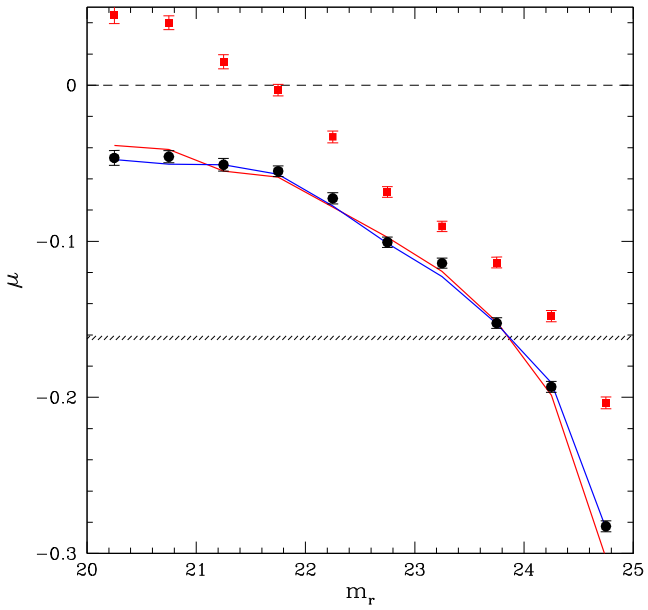

Figure 5. Multiplicative bias as a function of apparent magnitude for the simulated CCCP data. The measurements are averages for bins with a width of 0.5 mag. The black points correspond to an ellipticity distribution with $\epsilon_0 = 0.25$. The red and blue lines are for $\epsilon_0 = 0.15$ and 0.3, respectively. The red squares indicate the bias if we follow the procedure to evaluate P^γ used in H12. For comparison, the hatched region indicates the 68 per cent confidence interval for the average bias for galaxies with $20 < m_r < 25$ (for $\epsilon_0 = 0.25$).

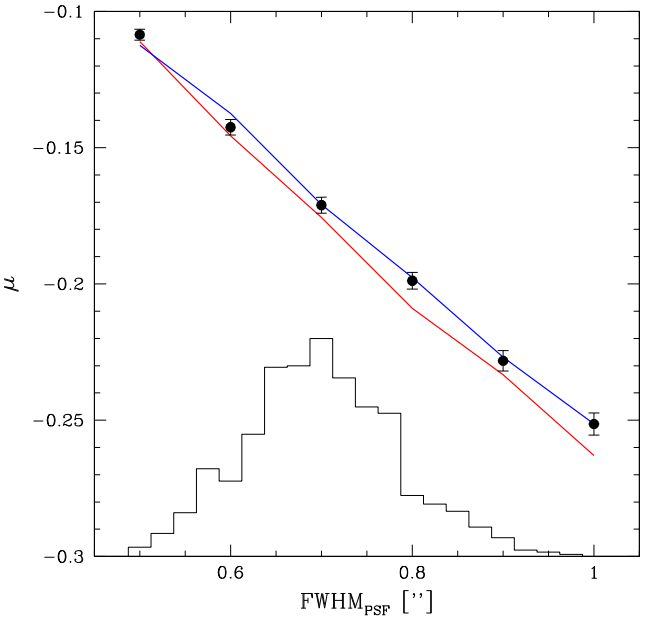

Figure 6. Multiplicative bias as a function of seeing for galaxies with $20 < m_r < 25$ for an ellipticity distribution with $\epsilon_0 = 0$ (red line), the reference case with $\epsilon_0 = 0.25$ (black points) and an extreme case with $\epsilon_0 = 0.5$ (blue line). The histogram shows the distribution of PSF sizes of the CCCP data measured for each chip. The image quality is typically best in the inner regions of the field of view, which are most relevant for the mass estimates.

the rather extreme range in ellipticity distribution. For $\epsilon_0 = 0.25$ we find a bias of $\mu \sim -0.115$, which is much larger than the value reported in Heymans et al. (2006) and Massey et al. (2007). Consequently we cannot ignore the multiplicative bias, as was done in H12.

The black points show the results if we use the measured value of P^γ for each galaxy. In this case the bias is larger ($\mu \sim -0.165$ for $\epsilon_0 = 0.25$), but also less sensitive to the ellipticity distribution. For this reason, as well as simplicity, we adopt this implementation as our reference. We note that the full chain of detection and shape analysis needs to be simulated. This is highlighted by the open black points, which indicate the bias if we use SExtractor (Bertin & Arnouts 1996) to detect objects and use the value for `FLUX_RADIUS` to compute the corresponding value for r_g: the observed bias is affected at the per cent level.[4]

Fig. 5 shows that the bias increases quickly for fainter galaxies, irrespective of the ellipticity distribution. This is also the case when we consider the implementation used by H12 (red squares). A strong dependence of the bias on the SNR was already observed in Bridle et al. (2010) and Kitching et al. (2012). The lack of faint galaxies in STEP1 is the main reason that a small bias was observed in Heymans et al. (2006). When we restrict the analysis to the magnitude range simulated by STEP1 we reproduce the small bias for the implementation used for that paper. Note that STEP2 simulated data that are deeper than our CCCP data. The implementation used by H12 gives smaller biases when considering the full range in magnitude, but overcorrects bright galaxies (i.e. $\mu > 0$). It appears that the choice of the fitting function partly compensated for the bias due to noise.

The SNR is also affected by the PSF size: the larger the PSF, the lower the SNR as the flux is spread over more pixels. The seeing also determines how well galaxies are resolved, which impacts the bias as well (see Appendix C). Fig. 6 shows the value of μ for galaxies with $20 < m_r < 25$ as a function of seeing for $\epsilon_0 = 0$ (red line), $\epsilon_0 = 0.25$ (black points) and an extreme case with $\epsilon_0 = 0.5$ (blue line). Note that we keep the range in apparent magnitude the same. The results demonstrate the importance of good image quality: the bias more than doubles from -0.11 to -0.25 as the seeing deteriorates from 0.5 to 1arcsec.

The number of faint galaxies increases rapidly (cf. Fig. 1), which results in source galaxies blending with fainter ones. Even if a faint galaxy is not detectable, it will impact the noise level, effectively introducing correlated noise that affects the local background determination. Both of these will modify the multiplicative bias in a way that that can only be quantified through simulations. In Bridle et al. (2010), Kitching et al. (2012) and Mandelbaum et al. (2014a) only postage stamps of isolated galaxies were analysed, and thus the effects of blending and faint galaxies were not included. Fig. 7 shows that this is an important effect, and cannot be neglected. To obtain these results we create images where we include galaxies down to a limiting magnitude $m_{\rm lim}$. The input GEMS catalogue is incomplete for $m_r > 25.5$ (see Fig. 1) and we augment the catalogue by duplicating the fainter galaxies such that the input counts follow the power-law relation seen at brighter magnitudes. At the faintest magnitudes these galaxies are unresolved in the simulated ground-based data, and hence the details of their structural properties are not critical.

The black points in Fig. 7 show that the bias for galaxies with $20 < m_r < 25$ increases until $m_{\rm lim} > 26.5$; in general we find that

[4] In the process of making this comparison we discovered that SExtractor (we used version 2.5.0) incorrectly assigns objects `FLAG=16` if they are elongated horizontally. This problem can be avoided by adopting a value for `MEMORY_BUFSIZE` larger than the image dimensions. We note that Gruen et al. (2014) discovered the same problem and reported on this.

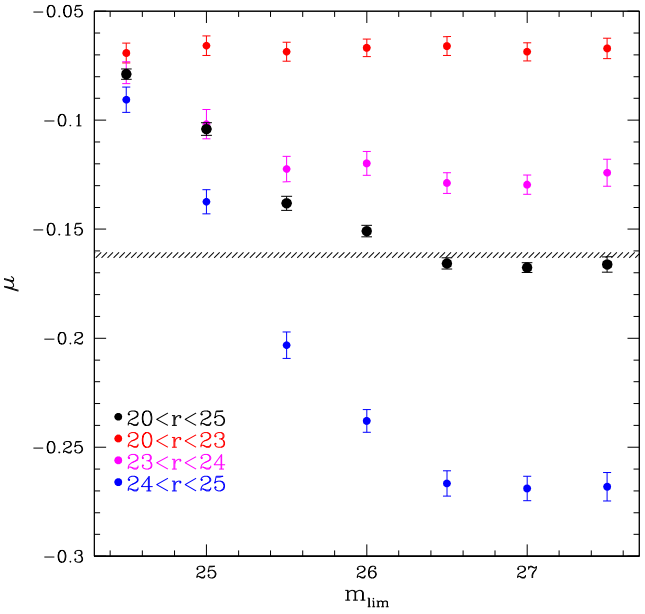

Figure 7. Multiplicative bias for an ellipticity distribution with $\epsilon_0 = 0.25$ where only galaxies with magnitudes brighter than $m_{\rm lim}$ are included in the simulation. The black points show the bias for $20 < m_r < 25$. For comparison, the hatched region indicates the 68 per cent confidence interval for the average bias for galaxies using the GEMS input catalogue. Irrespective of the magnitude range, the bias converges when $m_{\rm lim}$ is 1.5 mag fainter than the magnitude limit of the source sample.

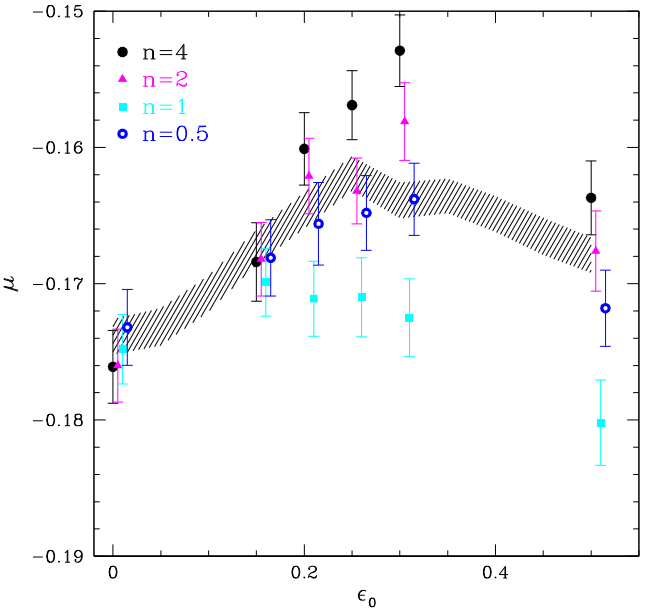

Figure 8. Multiplicative bias as a function of input ellipticity distribution for four different Sérsic indices (as indicated) for galaxies with $20 < m_r < 25$. The hatched region indicates the 68 per cent confidence region when the distribution of Sérsic indices from GEMS is used. The bias depends on the value of the Sérsic index, although we note that the results cannot be compared directly, as explained in the text.

the bias converges if we include sources that are 1.5 mag fainter than the magnitude limit of the sample of sources used to measure the weak lensing signal. This also appears to be true if we consider narrow bins in magnitude, such as the bin with $24 < m_r < 25$ for which the bias is large, but converges for $m_r > 26.5$. The dominant contribution of these faint galaxies is to act as a source of correlated noise, affecting the shape measurements of brighter galaxies. These results demonstrate that it is important to ensure that the input catalogue used for image simulations contains a sufficient number of galaxies fainter than the magnitude limit one is interested in.

For comparison the hatched area in Fig. 7 indicates the 68 per cent confidence region for the bias we obtain when we use the GEMS input catalogue, without introducing additional faint galaxies to account for incompleteness (for $m_r > 25.5$). Comparison with the black points suggests that the input catalogue is sufficient for the interpretation of the CCCP data, and we use it to compute the corrections in Section 2.4.

Finally we examine whether the bias depends on the assumed distribution of Sérsic indices. To do so, we create images where all galaxies have the same Sérsic index n, while keeping the other parameters the same. The results are presented in Fig. 8 for different values of ϵ_0. The bias depends on the value of n, although the range is small ($\sim$0.02 for $\epsilon_0 = 0.3$), with the results for $n = 1$ (corresponding to exponential profiles) most discrepant. We note, however, that the half-light radii were kept to the values listed in the GEMS catalogue, which can lead to small changes in the SNR, complicating a direct comparison. Given the small variation in μ, and the fact that the observed distribution of Sérsic indices is well-constrained, we assume that the uncertainty in this distribution can be ignored. Hence, the dominant uncertainty in our bias estimate arises from the uncertainty in the ellipticity distribution.

2.4 Empirical correction

If the simulated data resemble the actual observations sufficiently well, the average bias for the source sample could be used to adjust the cluster masses accordingly. Note that one would still have to determine the bias as a function of seeing. An additional complication arises, however, because we lack redshifts for the sources: the bias depends on the fluxes and sizes of the sources. As a result the bias may be redshift dependent, which is not captured by the image simulations as the same shear is applied to all sources. In reality the amplitude of the shear signal depends on the geometry of the lens–source configuration, quantified by the critical surface density(e.g. Bartelmann & Schneider 2001)

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l} D_{\rm ls}}, \tag{7}$$

where $D_{\rm s}$, $D_{\rm l}$ and $D_{\rm ls}$ are the angular diameter distances between the observer and the source, the observer and the lens and the lens and source, respectively. The sensitivity to the source redshift distribution is quantified by the ratio $\beta = D_{\rm ls}/D_{\rm s}$. The average shear for an ensemble of galaxies is proportional to $\langle(1+\mu)\beta\rangle$. If photometric redshifts for the individual sources are available, the redshift dependence of the lensing signal can be accounted for on an object-by-object basis and an average correction for the multiplicative bias is possible. An alternative route, which we take here, is to compute the multiplicative bias using the observed properties of individual galaxies. The correction, however, will still depend on the intrinsic ellipticity distribution of the sources.

We assume that the bias is only a function of SNR (e.g. Kacprzak et al. 2012; Melchior & Viola 2012) and the size relative to that of the PSF (e.g. Massey et al. 2013). We quantify the latter by the

parameter $\mathcal{R}$, defined as

$$\mathcal{R}^2 = \frac{r^2_{\rm h,*}}{r^2_{\rm h,gal} - r^2_{\rm h,*}}, \quad (8)$$

where $r_{\rm h,*}$ denotes the half-light radius of the PSF and $r_{\rm h,\,gal}$ that of the observed galaxy. Despite being a simple prescription, we show in Section 2.5 that this captures the dependence on PSF size quite well. As a proxy for the SNR we take $\nu = 1/\sigma_e$, the reciprocal of the uncertainty in the polarization (Hoekstra et al. 2000). We refer the interested reader to Appendix C for more details about our empirical correction, which is given by

$$\mu(\nu) = \frac{b(\nu)}{1 + \alpha(\epsilon_0)\mathcal{R}}. \quad (9)$$

The dependence on the resolution parameter $\mathcal{R}$ is described by a single parameter α that is a function of ϵ_0 only. We require three free parameters to describe dependence of the bias on ν: $b(\nu) = b_0 + b_1/\sqrt{(\nu)} + b_2/\nu$, with fit parameters that vary smoothly with ϵ_0. The best-fitting parameters as a function of ϵ_0 are listed in Table C1. Although this parametrization does not describe the simulated data perfectly, and obvious improvements can be suggested, we prefer our choice as it provides a sufficiently accurate correction, with a relatively small number of parameters. Including more parameters did not improve the robustness of the correction.

We find that our parametrization of the bias does not perform well for galaxies with large observed sizes ($r_{\rm h} > 5$ pixels; see e.g. Fig. C1). As discussed in more detail in Section 4 we find that the recovered lensing signal for these galaxies is biased low (see Fig. 14). Closer investigation of the simulated data shows that most of these galaxies are intrinsically small and faint. In some of the cases the sizes are increased by noise in the images, but a large fraction is blended with other galaxies. The large increase in galaxy density in clusters of galaxies is expected to exacerbate this problem, which is not captured by our simulations (which are representative of the field). To minimize the impact of blending, we include only galaxies with $r_{\rm h} < 5$ pixels in the lensing analysis. This size cut is applied to the tests presented below, as well as our actual measurements.

2.5 Testing the empirical correction

To quantify how well the correction works when we apply it to the simulated data, we first examine the residual bias $\mu_{\rm cor}$ as a function of ϵ_0. As explained in more detail in Section 3.1, we restrict the source sample to galaxies with $22 < m_r < 25$ to allow for a better correction for the contamination by cluster members. In addition we apply a size cut, requiring that $r_{\rm h} < 5$ pixels. This is motivated by our image simulations where we found that the correction for large galaxies is biased, because they are blended or too faint to have their shapes measured reliably. We therefore limit the discussion of the performance of the empirical correction to this range in apparent magnitude and galaxy size. The results are presented in Fig. 9, which shows that for the range of interest for ϵ_0 (indicated by the hatched region) $|\mu_{\rm cor}| < 0.005$.

As the intrinsic ellipticity distribution remains uncertain, it is useful to examine the bias that is introduced when an incorrect value for ϵ_0 is used for the empirical correction. If we use the parameters corresponding to $\epsilon_0 = 0.25$ to correct the measurements from other input distributions we find that $\mu_{\rm cor}$ is still small, as indicated by the black line in Fig. 9. Our empirical correction is quite robust against the uncertainty in the input ellipticity distribution (if we take $\epsilon_0 = 0.25$). As discussed in Appendix C the parametrization

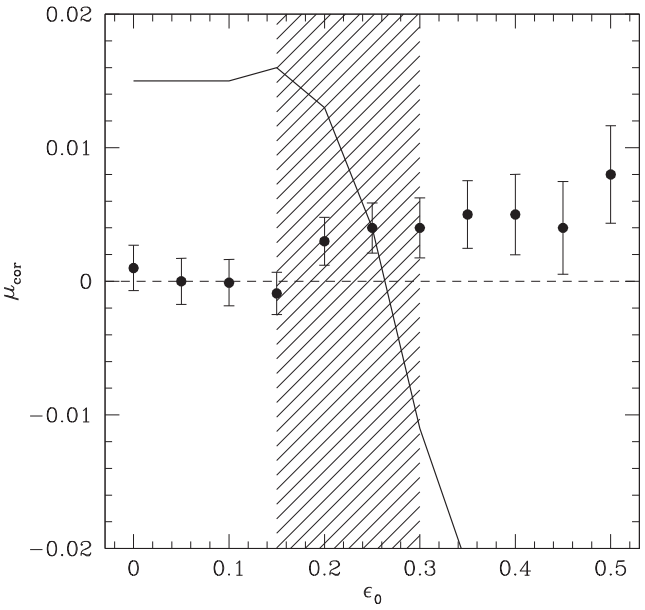

Figure 9. Residual multiplicative bias as a function of input ellipticity distribution for galaxies with $22 < m_r < 25$ and $r_{\rm h} < 5$ pixels. The black points indicate the results when the correct value for ϵ_0 is used in the correction. The line indicates the residual bias if we assume $\epsilon_0 = 0.25$ in the correction, instead of the correct value for the simulated distribution. Adopting a value $\epsilon_0 = 0.25$ for the correction results in $|\mu_{\rm cor}| < 0.015$ over the expected range in actual ϵ_0 values (indicated by the hatched region).

for the size dependence of the bias is not accurate for large galaxies, which are typically bright. This is indeed reflected in the residual bias as a function of apparent magnitude: we observe $\mu_{\rm cor} \sim 0.02$ for $m_r < 22$, with a bias $\sim$0 for galaxies with $m_r > 22$.

The empirical correction was determined for a particular PSF and integration time. Although it is in principle possible to create simulated data sets for each set of observing conditions, a useful correction scheme should be more generally applicable. As discussed in Appendix D, we also simulated data from the second Red-sequence Cluster Survey (RCS2). These data are shallower, but the results presented in Fig. D1 indicate that the correction works fairly well for these shallower data. This suggests that the modelling of the SNR-dependence is adequate.

More interesting is whether our approach to quantify how well galaxies are resolved, i.e. the choice of $\mathcal{R}$, can be used for a range of seeing values. To this end we correct the set of images used to study the seeing dependence of the bias (see Fig. 6). The results for galaxies with $22 < m_r < 25$ are presented in Fig. 10, which shows $\mu_{\rm cor}$ as a function of the FHWM of the PSF. Even for a FWHM of 1 arcsec the bias is reduced significantly. None the less the residual bias can still be substantial. However, as is shown by the seeing histogram, the CCCP data span a relatively narrow range, and the mean bias for the full sample is $\langle\mu_{\rm cor}\rangle = -0.001$. Furthermore, the largest FWHM values occur on chips far away from the cluster location. We therefore ignore the seeing dependence, as the residual bias is still much smaller than the statistical uncertainties for individual clusters, and the ensemble average bias very small.

We conclude that our empirical correction is adequate to determine cluster masses for the CCCP sample. Based on the residuals we assign a systematic uncertainty of 2 per cent in the cluster masses due to the uncertainty in the input ellipticity distribution, the limited

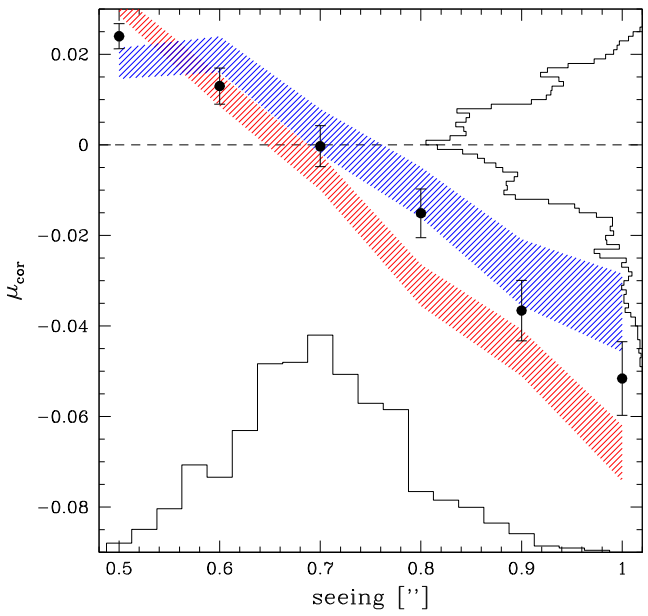

Figure 10. Residual multiplicative bias as a function of seeing for sources with $22 < m_r < 25$. The black points show the results for $\epsilon_0 = 0.25$. The red (blue) hatched regions indicate the 68 per cent confidence region for the bias if we use the parameters for $\epsilon_0 = 0.25$ to correct the simulations with input distribution with $\epsilon_0 = 0.15$ ($\epsilon_0 = 0.3$) instead. The bottom histogram shows the seeing distribution for each chip in CCCP, the side histogram shows the corresponding distribution of residual bias, with $\langle \mu_{cor} \rangle = -0.001$.

exploration of the role of morphology, and the variation in image quality.

3 SOURCE REDSHIFT DISTRIBUTION

We lack the colour information to derive photometric redshifts for the individual sources, as opposed to e.g. Applegate et al. (2014) and Umetsu et al. (2014). Fortunately it is sufficient to know the average source redshift distribution, which we discuss in more detail in Section 3.2. We note, however, that photometric redshifts enable an optimal weighting of the sources. In particular photometric redshifts allow for a separation of source galaxies and cluster members. The latter are unlensed, as are foreground galaxies, and thus dilute the observed lensing signal by a factor $f_{\rm contam}(r) = 1 + n_{\rm cl}(r)/n_{\rm fld}$, where $n_{\rm cl}(r)$ is the number density of cluster galaxies and $n_{\rm fld}$ the number density of field galaxies. This correction is especially important at small distances from the cluster centre. As we describe in Section 3.1, we can correct for the reduction in signal by quantifying the level of contamination. This assumes that the orientations of the cluster members are random, which is supported by observations (Sifón et al. 2015).

In doing so, we assume that the change in counts is solely caused by contamination by cluster members. However, gravitational lensing not only changes the galaxy shapes, but also magnifies sources. As a result the background sources appear brighter, leading to an increase in the observed counts. On the other hand, the actual volume is reduced, because the observed solid angle corresponds to a smaller solid angle behind the cluster. Consequently, the net change depends on the number density of background galaxies as a function of apparent magnitude (see e.g. Section 3.4 in Mellier 1999). In our case we observe a slope $\mathrm{d}\log N_{\rm gal}/\mathrm{d}M \sim 0.38$–$0.4$ for galaxies with $22 < m_r < 24$. This is somewhat steeper than the slope of $\sim$0.33 observed by Hogg et al. (1997) in the R band. In either case, the net effect is minimal: even for $\kappa = 0.1$ the change in observed counts is 1–3 per cent. Hence, it is safe to assume that the excess counts are solely caused by contamination by cluster members. Note, however, that the source redshift distribution is somewhat changed, as we do see intrinsically fainter galaxies. We verified that the resulting change in mean redshift can be safely neglected.

3.1 Contamination by cluster members

To reduce contamination by cluster members, H12 used their limited colour information to identify and remove galaxies on the red sequence. As shown in Hoekstra (2007), this does lower the contamination, but only by $\sim$30 per cent as many faint cluster members are blue. Furthermore H12 assumed that the excess number density of galaxies can be described as $f_{\rm contam} \propto r^{-1}$, with the amplitude determined for each cluster. The analysis presented here differs from H12 in several ways. Rather than applying a colour cut, we restrict the magnitude range of the sources. Furthermore we use a more flexible model to quantify the radial dependence of the excess counts. We also correct the excess source counts for the obscuration by cluster members (Simet & Mandelbaum 2014). Finally, as described below, rather than considering the excess counts, we account for the weight provided by the uncertainty in the shape measurement.

H12 used sources as bright as $m_r \sim 20$, for which the level of contamination is high. This is demonstrated by Fig. 11, which shows the corresponding correction factor as a function of distance from the cluster for different bins of apparent magnitude. For the Megacam data the counts are normalized to the average number density at radii larger than $r_{\rm max} = 4\,h_{70}^{-1}$ Mpc, i.e. we assume that the level of contamination can be ignored at those large radii. This is supported by comparison of the observed counts to those in blank fields. In addition, we used predictions based on the halo model described in Cacciato et al. (2013) to estimate the expected level of contamination due to neighbouring structures. In line with our comparison of the blank field galaxy counts, we find that the contribution from local structures can indeed be ignored. The field of view of the CFH12k data is smaller and we estimate the background level using the number density at radii larger than $r_{\rm max} = 3\,h_{70}^{-1}$ Mpc.

The contamination is much higher for bright galaxies: such galaxies are rare in the field, but much more common in clusters. For the brightest bin ($20 < m_r < 21$) the cluster members outnumber source galaxies 3 to 1 in the inner $\sim 1\,h_{70}^{-1}$ Mpc. Such a large level of contamination is difficult to model reliably and for this reason we decided to increase the bright limit of the source sample to $m_r = 22$ from the typical value of $m_r = 20$ used in H12. This leads to a reduction in excess counts that is comparable to excluding the galaxies on the red sequence. Conveniently, the empirical correction for the multiplicative bias in the shape measurement also performs better for galaxies with $22 < m_r < 25$. Furthermore, the lensing signal is higher for the fainter galaxies. The shapes of brighter galaxies are measured better and consequently given more weight in the lensing analysis (see equation 5). Rather than correcting for the excess counts, which effectively assumes that the weight is uniform, we compute the excess weight as a function of radius. This is a minor correction, which increases the masses of the parametric Navarro–Frenk–White (NFW) fits by $\sim$2–3 per cent (see Section 4.1).

The other important change we make is that we allow the radial profile of the excess weight to vary from cluster to cluster by introducing a core. The simple $1/r$ profile used by H12 is not a good description for all clusters or magnitude bins. Investigation of the ensemble averaged residuals suggest that it leads to an overestimation of the contamination in the inner $\sim 500\,h_{70}^{-1}$ kpc and an

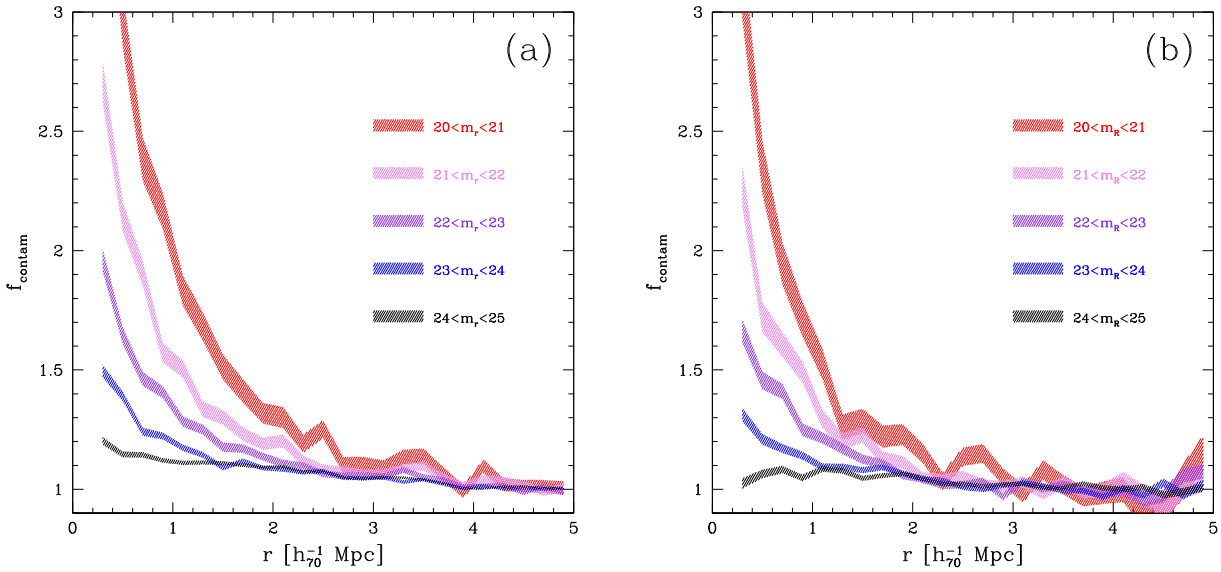

Figure 11. The ensemble averaged correction factor for contamination by cluster members as a function of distance for different bins of apparent magnitude. Panel (a) shows the results for clusters observed with Megacam, whereas panel (b) is for the CFH12k observations.

underestimation by 4–5 per cent at larger radii because the model attempts to compensate for the poor fit in the cluster cores. To describe the excess counts we now fit

$$f_{\rm contam}(r) = 1 + n_0 \left(\frac{1}{r + r_{\rm c}} - \frac{1}{r_{\rm max} + r_{\rm c}} \right), \qquad (10)$$

to each cluster, where we take $r_{\rm max} = 4\,h_{70}^{-1}$ Mpc for the Megacam data and $r_{\rm max} = 3\,h_{70}^{-1}$ Mpc for the CFH12k data. The core radius $r_{\rm c}$ is a free parameter that we fit for each cluster separately. Fig. 12 shows the ensemble averaged residual contamination for sources with $22 < m_{r/R_C} < 25$ as a function of radius, suggesting that the systematic uncertainty for the ensemble of clusters is at most a few per cent. Note that the residuals may be larger for individual clusters, resulting in increased scatter. However, the results presented in Fig. 12 suggest that residual contamination will have a minimal impact on the normalization of scaling relations derived from CCCP measurements and we adopt a systematic uncertainty in the mass of 2 per cent as a result of the imperfect correction for contamination by cluster members.

The observed counts are biased low in the inner regions because the presence of bright cluster members affects our ability to detect and analyse sources. Although Simet & Mandelbaum (2014) showed that this is an important source of bias for the measurement of magnification, it may also lead to a small bias in our estimate of the dilution of the lensing signal. We simulated the impact of this and as described in more detail in Appendix E we find that the impact is indeed small, boosting the masses from the parametric NFW fits by 1–2 per cent (see Section 4.1). The aperture masses, which are discussed in Section 4.2 are not affected, because they rely on estimates of the lensing signal at large radii where the density of cluster members is low.

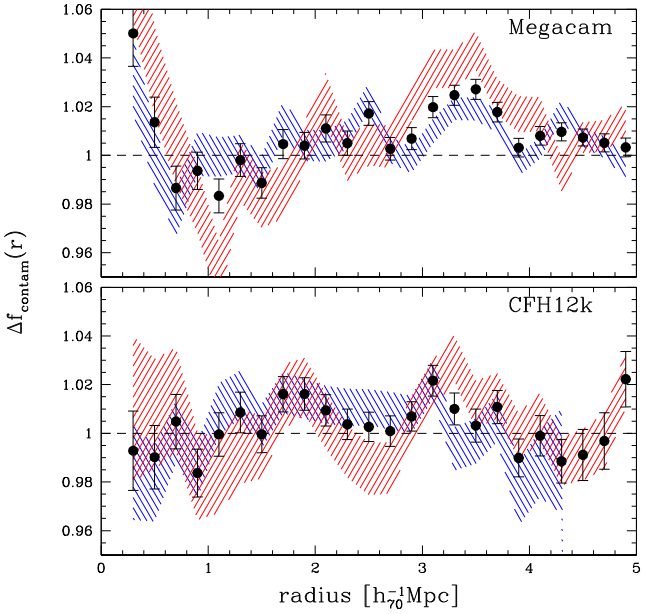

Figure 12. Plot of the ensemble averaged residual contamination as a function of distance to the cluster centre. The top panel shows the results for sources with $22 < m_r < 25$ based on the Megacam data. The hatched regions indicate the 68 per cent confidence intervals for the residuals for clusters with $z < 0.25$ (blue) and $z > 0.25$ (red). The bottom panel shows the same, but now for CFH12k data and sources with $22 < m_{R_C} < 25$. For both data sets the residual contamination on the scales of interest ($> 0.5\,h_{70}^{-1}$ Mpc) is at most 2 per cent.

3.2 Photometric redshift catalogue

The weak lensing analysis of the initial sample of 20 clusters observed with the CFH12k camera in Hoekstra (2007) used the photometric redshift distributions derived for the HDF North and South using the available deep multiwavelength data (Fernández-Soto, Lanzetta & Yahil 1999). However, the area covered is small, leading to concerns whether the redshift distributions are representative. For this reason H12 used the photometric redshift distribution from Ilbert et al. (2006), which is based on the four CFHT Legacy Survey (CFHTLS) Deep fields (each field covers 1 deg^2). However, Ilbert et al. (2006) derived photometric redshifts using observations in five optical filters (*ugriz*). Although these data are very deep, the lack of NIR data is a concern for high-redshift galaxies. Good quality NIR data are essential for this purpose because at $z > 1.5$

the Balmer and 4000 Åbreak features in galaxy spectral energy distributions, which are the strongest features for determining a photometric redshift, are redshifted into the NIR.

For our analysis we use data from the COSMOS (Scoville et al. 2007) which observed a single field covering 2 deg^2 with *HST*, and for which extensive multiwavelength and spectroscopic data are available. Ilbert et al. (2009) present photometric redshifts for these data based on measurements in 30 bands (hereafter referred to as COSMOS-30). This redshift distribution was used in Applegate et al. (2014) to determine their 'colour-cut' masses. At the time of the analysis by Ilbert et al. (2009), deep NIR data were not available. This situation has changed thanks to the UltraVISTA survey, an ESO public survey performing deep imaging of the COSMOS field in five NIR filters (see McCracken et al. 2012, for details). The UltraVISTA data are a significant upgrade to the available NIR imaging in the COSMOS field, and therefore allow for a marked improvement in the quality of photometric redshifts for galaxies at $z > 1.5$. Two public NIR-selected catalogues have been produced using the UltraVISTA data; one where galaxies were selected in the K_s band (Muzzin et al. 2013), and one where objects were selected using a co-added χ^2 image of the NIR bands (Ilbert et al. 2013). Both of these catalogues provide photometric redshifts and stellar masses for the galaxies.

Galaxies have a wide range of optical–NIR colours, and therefore NIR-selected samples of galaxies are typically quite different from *r*-selected samples, particularly for the high-redshift end of the distribution. Consequently we cannot simply use the available photometric redshift catalogues, because our source galaxies are selected from the deep CCCP *r*-band imaging. Furthermore, many of the sources that are of interest for the lensing analysis may be missing because the UltraVISTA data ($K_s = 23.9$ AB) are 1.5 mag shallower than the CCCP optical imaging.

In order to construct a representative photometric redshift distribution for the sources, we created a new r^+-selected catalogue of the COSMOS/UltraVISTA field using the Subaru imaging of the field from Capak et al. (2007). The Subaru r^+ imaging has good image quality (FWHM ~ 0.8 arcsec) and reaches a 5σ depth of ~ 26.5 AB. This is approximately a full magnitude deeper than the CCCP imaging, and therefore provides a complete sample of galaxies to $m_r \sim 25.5$ with well-measured spectral energy distributions, a prerequisite for calculating photometric redshifts. The r^+-selected catalogue was constructed in the identical manner as the K_s-selected catalogue described in Muzzin et al. (2013) and we refer the reader to that paper for complete details of the catalogue construction. In brief, the catalogue consists of photometry in 29 photometric bands ranging from 0.15-24 μm and incorporates the available *GALEX* (Martin et al. 2005), Subaru (Capak et al. 2007), UltraVISTA (McCracken et al. 2012) and *Spitzer* data (Sanders et al. 2007). Images were PSF-matched and photometry was performed in fixed 2.1 arcsec diameter apertures.

Photometric redshifts for all galaxies were calculated using the EAZY photometric redshift code (Brammer, van Dokkum & Coppi 2008), which determines photometric redshifts using linear combinations of multiple templates as well as a template error function to account for data/template mismatch. EAZY is well-tested and performs well amongst the best of publicly-available photometric redshift codes (e.g. Hildebrandt et al. 2010). The photometric redshifts were further refined by determining small offsets to the photometric zero-points using the $\sim$ 5000 spectroscopic redshifts available in the COSMOS field from the zCOSMOS-10k sample (Lilly et al. 2009). The process is iterative, and the final photometric catalogue contains photometric redshifts accurate to $\Delta z/(1+z) = 0.01$, with a catastrophic outlier fraction of $\sim$1 per cent. This estimate of the accuracy is based on the zCOSMOS spectroscopic redshifts, which are primarily bright galaxies at $z < 1.5$. At $z > 1.5$, it is more difficult to assess how accurate the photometric redshifts are, due to the lack of a spectroscopic calibration sample. Small numbers of spectroscopic redshifts are available from various NIR spectroscopic surveys, and those suggest that the accuracy at $z > 1.5$ for bright galaxies is only slightly worse, of order a few per cent. We have made the full r^+-selected catalogue publicly available on the K_s-selected catalogue website.[5] Also included is a simplified version of that catalogue that can be used for a quick calculation of photometric redshift distributions for future lensing analyses.

The left-hand panel in Fig. 13 shows the resulting redshift distribution for galaxies with $20 < m_r < 25$ using our *r*-band selected photometric redshift catalogue. The red dashed histogram shows the redshift distribution of the galaxies matched to the NIR-selected catalogue from Ilbert et al. (2013), which makes use of the $\sim$ 3000 unpublished zCOSMOS-deep spectroscopic redshifts. For reference, the blue histogram shows the corresponding redshift distribution for the CFHTLS Deep fields from Ilbert et al. (2006), which was used by H12. The inset panel shows a direct comparison of the photometric redshifts derived here and those from Ilbert et al. (2013) who used a different algorithm and measured the photometry independently. The overall agreement is remarkably good for the galaxies in common; only for $m_r > 24$ are some of the galaxies assigned a high redshift in the catalogue from Ilbert et al. (2013) and a low redshift by EAZY. These represent 2 per cent of the total sample of sources and about 20 per cent of the galaxies for which Ilbert et al. (2013) find $z > 2$. The excellent agreement for most of the galaxies is a demonstration of the quality of the data, both in terms of depth and wavelength coverage.

The impact of the differences in source redshift distributions on the cluster-mass estimates is quantified in the right-hand panel of Fig. 13, which shows the lensing efficiency $\langle\beta\rangle$ as a function of apparent magnitude. Our results are indicated by the black points, whereas the red dashed curve corresponds to the redshift distribution from Ilbert et al. (2013). For $m_r < 23$ the agreement is very good, whereas the higher number of $z > 2.5$ galaxies in the Ilbert et al. (2013) catalogue results in a higher value for $\langle\beta\rangle$, and thus a lower mass. We list $\langle\beta\rangle$ for the sources in Table 1; column 7 lists the values for our analysis of the COSMOS and UltraVISTA data, and column 8 lists the results using the results from Ilbert et al. (2013). To determine cluster masses we use $\langle\beta\rangle_{\rm used}$ provided in column 9, which is the average of the values obtained for the two redshift distributions. The value for $\langle\beta^2\rangle_{\rm used}$, which is a measure of the width of the distribution (see Hoekstra et al. 2000; Hoekstra 2007, for details) is also an average of the two estimates.[6] These numbers include a size cut similar to the one applied to our lensing data. The red line in Fig. 14 shows how $\langle\beta\rangle$ depends on the observed half-light radius; we find that the dependence on size is very small, and we therefore conclude that the size cuts do not introduce a significant bias.

The significant differences in redshift distribution demonstrate that the lack of reliable photometric redshift estimates remains a key source of error. The unique range in wavelength and quality

[5] http://www.strw.leidenuniv.nl/galaxyevolution/ULTRAVISTA/

[6] The values listed here are corrected for an error in the calculation of $\langle\beta^2\rangle$ that reduced the sensitivity to the convergence in our previous work.

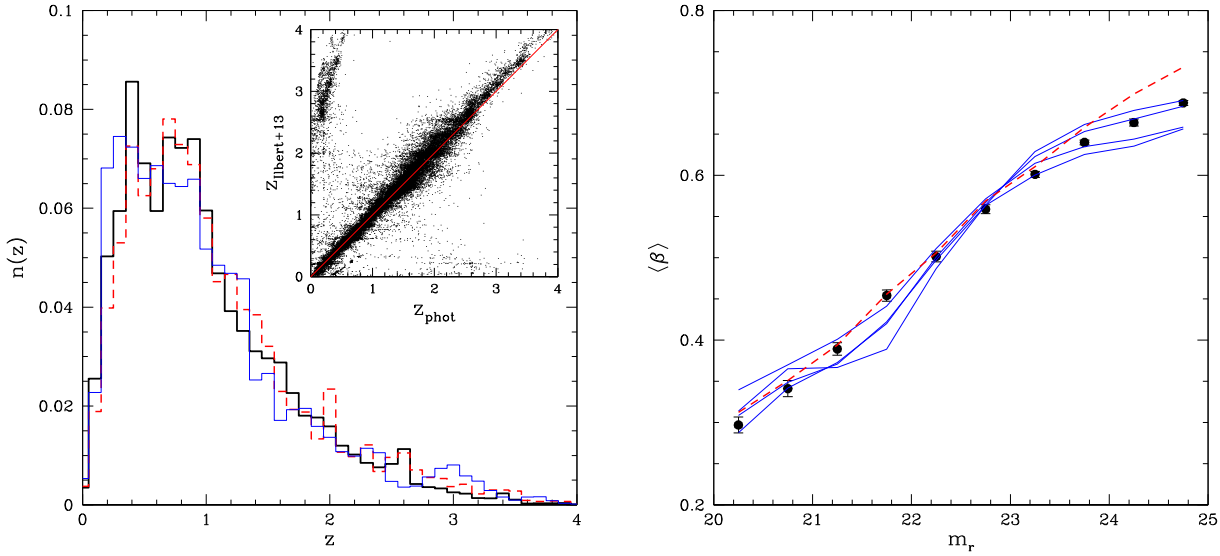

Figure 13. Left-hand panel: the redshift distribution of galaxies with $20 < m_r < 25$ based on the COSMOS + UltraVISTA photometric redshift analysis. The solid black histogram indicates the results from our analysis, whereas the red dashed histogram corresponds to the results from Ilbert et al. (2013). The inset shows the comparison of the photometric redshifts for the galaxies in common. The blue histogram shows the redshift distribution used in H12. Right-hand panel: value of $\langle\beta\rangle$ as a function of m_r for a cluster at $z = 0.2$. The points with error bars are our measurements, which agree well with the red dashed line (Ilbert et al. 2013) for $m_r < 23$. The four blue lines are the results for the CFHTLS Deep fields studied by Ilbert et al. (2006) and used in H12.

of the COSMOS+UltraVISTA data are a major step forward, but without complete spectroscopic coverage, the uncertainty at the highest redshifts remains. Furthermore, cosmic variance can still be important for a single field (see e.g. Hoekstra et al. 2011, for estimates). Assigning a systematic uncertainty remains difficult, but we use the difference between our photometric redshifts and those from Ilbert et al. (2013, indicated as I13 in Table 1) as an estimate for the systematic uncertainty.

The blue lines in the right-hand panel of Fig. 13 show $\langle\beta\rangle$ for the four CFHTLS Deep fields (Ilbert et al. 2006). The main difference occurs at $m_r \sim 21.5$, although the average is also lower than the new estimates for faint galaxies. We use the variation between the fields to estimate the contribution of cosmic variance for our redshift distribution. This is conservative, because the COSMOS field itself is larger than each of the CFHTLS Deep fields. The systematic uncertainty $\delta\beta$ listed in Table 1 is the sum of the dispersion measured from the four fields and $|\beta_{\rm I13} - \beta|/2$. These amount to ∼3 per cent for clusters at $z = 0.2$, increasing to ∼8 per cent for $z = 0.55$.

Variation in the actual source redshift distribution behind a cluster leads to an increase in the statistical uncertainty. The impact of this was studied by Hoekstra et al. (2011) using simulations. Their findings suggest that the lack of redshift information for the individual sources increases the statistical uncertainty in the mass by ∼3 per cent for a cluster at $z = 0.2$ and by ∼10 per cent at $z = 0.6$.

4 UPDATED MASS ESTIMATES

We use our new insights in the shape measurement bias and the source redshift distribution to update the mass estimates provided by H12. The only change we made to the original shape measurements is a corrected estimate for σ_e, the uncertainty in the polarization[7] because we use this quantity in our estimate of the multiplicative bias. This correction also affects the weight defined by equation (5), although the impact is very minor. The star–galaxy separation is somewhat more conservative: previously we included faint galaxies with sizes comparable to the PSF in the lensing analysis, although they were severely down-weighted in practice. As the correction scheme requires sizes larger than the PSF, and the fact that these objects do not contribute much to the average signal, we now only select objects larger than the PSF in the object catalogue. The star selection itself was not changed, and consequently we use the same PSF model parameters.

We apply our empirical correction given by equation (9) to the object catalogues using the observed values of $r_{\rm h}$ and ν and recompute the tangential shear profile as a function of radius, taking the cluster centres used in H12 (table 1 of that paper). Based on our image simulations we also apply a size cut $r_{\rm h} < 5$ pixels. The solid points in Fig. 14 show the lensing signal, averaged over the full ensemble of clusters, as a function of the observed half-light radius of the sources. The signal is quantified by the Einstein radius $r_{\rm E}$ obtained from a fit to the measurements between $0.5{-}4\,h_{70}^{-1}$ Mpc. We correct each bin for the mean β, which does not vary significantly with source size (indicated by the red line), suggesting that our results are not sensitive to the size cuts. The resulting value of $r_{\rm E}/\beta$ should not depend on the source size.

The observations support our finding from the image simulations that the signal is biased low for galaxies with $r_{\rm h} > 5$ pixels. For reference we also show the results if we do not apply our empirical correction for multiplicative bias (open points). The corrected results are consistent with a constant signal, but the uncorrected measurements are bias low for the smallest sources.

We update the correction for Galactic extinction using the Schlafly & Finkbeiner (2011) recalibration of the Schlegel, Finkbeiner & Davis (1998) infrared-based dust map. This reduces the correction for clusters in highly extinguished regions. To ensure a more robust correction for the contamination by cluster members we do not apply a colour cut, but instead limit the source sample to galaxies with $22 < m_{r/R_C} < 25$. In the case of the CFH12k data,

[7] The older version of the code included an incorrect treatment of the Poisson noise.

Table 1. Basic information for the CCCP clusters.

(1)	(2) Name	(3) z	(4) RA (J2000.0)	(5) Dec. (J2000.0)	(6) mag (h_{70}^{-1}kpc)	(7) $\langle\beta\rangle$	(8) $\langle\beta\rangle_{\rm I13}$	(9) $\langle\beta\rangle_{\rm used}$	(10) $\delta\beta$	(11) $\langle\beta^2\rangle_{\rm used}$
1	Abell 68	0.255	$00^{\rm h}37^{\rm m}06\overset{\rm s}{.}9$	$+09^\circ09'24''$	22–25	0.553	0.579	0.566	0.021	0.375
2	Abell 209	0.206	$01^{\rm h}31^{\rm m}52\overset{\rm s}{.}5$	$-13^\circ36'40''$	22–25	0.625	0.649	0.637	0.019	0.453
3	Abell 267	0.230	$01^{\rm h}52^{\rm m}42\overset{\rm s}{.}0$	$+01^\circ00'26''$	22–25	0.586	0.610	0.598	0.019	0.409
4	Abell 370	0.375	$02^{\rm h}39^{\rm m}52\overset{\rm s}{.}7$	$-01^\circ34'18''$	22–25	0.414	0.442	0.428	0.026	0.244
5	Abell 383	0.187	$02^{\rm h}48^{\rm m}03\overset{\rm s}{.}4$	$-03^\circ31'44''$	22–24.5	0.636	0.654	0.645	0.014	0.462
6	Abell 963	0.206	$10^{\rm h}17^{\rm m}03\overset{\rm s}{.}8$	$+39^\circ02'51''$	22–25	0.621	0.644	0.632	0.018	0.448
7	Abell 1689	0.183	$13^{\rm h}11^{\rm m}30\overset{\rm s}{.}0$	$-01^\circ20'30''$	22–24.5	0.647	0.666	0.656	0.015	0.475
8	Abell 1763	0.223	$13^{\rm h}35^{\rm m}20\overset{\rm s}{.}1$	$+41^\circ00'04''$	22–25	0.590	0.613	0.601	0.018	0.412
9	Abell 2218	0.176	$16^{\rm h}35^{\rm m}48\overset{\rm s}{.}8$	$+66^\circ12'51''$	22–24.5	0.646	0.662	0.654	0.012	0.471
10	Abell 2219	0.226	$16^{\rm h}40^{\rm m}19\overset{\rm s}{.}9$	$+46^\circ42'41''$	22–25	0.596	0.621	0.609	0.020	0.421
11	Abell 2390	0.228	$21^{\rm h}53^{\rm m}36\overset{\rm s}{.}8$	$+17^\circ41'44''$	22–25	0.597	0.624	0.611	0.021	0.423
12	MS 0015.9+1609	0.547	$00^{\rm h}18^{\rm m}33\overset{\rm s}{.}5$	$+16^\circ26'16''$	22–25	0.277	0.304	0.291	0.025	0.138
13	MS 0906.5+1110	0.170	$09^{\rm h}09^{\rm m}12\overset{\rm s}{.}6$	$+10^\circ58'28''$	22–25	0.678	0.700	0.689	0.016	0.515
14	MS 1224.7+2007	0.326	$12^{\rm h}27^{\rm m}13\overset{\rm s}{.}5$	$+19^\circ50'56''$	22–25	0.465	0.492	0.479	0.024	0.289
15	MS 1231.3+1542	0.235	$12^{\rm h}33^{\rm m}55\overset{\rm s}{.}4$	$+15^\circ25'58''$	22–25	0.587	0.614	0.600	0.021	0.412
16	MS 1358.4+6245	0.329	$13^{\rm h}59^{\rm m}50\overset{\rm s}{.}6$	$+62^\circ31'05''$	22–25	0.466	0.494	0.480	0.025	0.290
17	MS 1455.0+2232	0.257	$14^{\rm h}57^{\rm m}15\overset{\rm s}{.}1$	$+22^\circ20'35''$	22–25	0.564	0.594	0.579	0.024	0.388
18	MS 1512.4+3647	0.373	$15^{\rm h}14^{\rm m}22\overset{\rm s}{.}5$	$+36^\circ36'21''$	22–25	0.427	0.458	0.442	0.027	0.256
19	MS 1621.5+2640	0.428	$16^{\rm h}23^{\rm m}35\overset{\rm s}{.}5$	$+26^\circ34'14''$	22–25	0.373	0.404	0.389	0.027	0.211
20	CL0024.0+1652	0.390	$00^{\rm h}26^{\rm m}35\overset{\rm s}{.}6$	$+17^\circ09'44''$	22–25	0.393	0.420	0.407	0.025	0.226
21	Abell 115N	0.197	$00^{\rm h}55^{\rm m}50\overset{\rm s}{.}6$	$+26^\circ24'38''$	22–25	0.645	0.670	0.658	0.019	0.478
	Abell 115S	0.197	$00^{\rm h}56^{\rm m}00\overset{\rm s}{.}3$	$+26^\circ20'33''$	22–25	0.645	0.670	0.658	0.019	0.478
22	Abell 222	0.213	$01^{\rm h}37^{\rm m}34\overset{\rm s}{.}0$	$-12^\circ59'29''$	22–25	0.620	0.645	0.633	0.020	0.449
23	Abell 223N	0.207	$01^{\rm h}38^{\rm m}02\overset{\rm s}{.}3$	$-12^\circ45'20''$	22–25	0.629	0.653	0.641	0.019	0.459
	Abell 223S	0.207	$01^{\rm h}37^{\rm m}56\overset{\rm s}{.}0$	$-12^\circ49'10''$	22–25	0.629	0.653	0.641	0.019	0.459
24	Abell 520	0.199	$04^{\rm h}54^{\rm m}10\overset{\rm s}{.}1$	$+02^\circ55'18''$	22–25	0.642	0.667	0.655	0.019	0.475
25	Abell 521	0.253	$04^{\rm h}54^{\rm m}06\overset{\rm s}{.}9$	$-10^\circ13'25''$	22–25	0.559	0.583	0.571	0.021	0.381
26	Abell 586	0.171	$07^{\rm h}32^{\rm m}20\overset{\rm s}{.}3$	$+31^\circ38'01''$	22–25	0.668	0.687	0.678	0.014	0.501
27	Abell 611	0.288	$08^{\rm h}00^{\rm m}56\overset{\rm s}{.}8$	$+36^\circ03'24''$	22–25	0.512	0.536	0.524	0.022	0.332
28	Abell 697	0.282	$08^{\rm h}42^{\rm m}57\overset{\rm s}{.}6$	$+36^\circ21'59''$	22–25	0.532	0.559	0.545	0.023	0.354
29	Abell 851	0.407	$09^{\rm h}42^{\rm m}57\overset{\rm s}{.}5$	$+46^\circ58'50''$	22–25	0.391	0.418	0.405	0.026	0.224
30	Abell 959	0.286	$10^{\rm h}17^{\rm m}36\overset{\rm s}{.}0$	$+59^\circ34'02''$	22–25	0.528	0.556	0.542	0.024	0.350
31	Abell 1234	0.166	$11^{\rm h}22^{\rm m}30\overset{\rm s}{.}0$	$+21^\circ24'22''$	22–25	0.693	0.714	0.703	0.015	0.534
32	Abell 1246	0.190	$11^{\rm h}23^{\rm m}58\overset{\rm s}{.}8$	$+21^\circ28'50''$	22–25	0.651	0.673	0.662	0.017	0.483
33	Abell 1758	0.279	$13^{\rm h}32^{\rm m}43\overset{\rm s}{.}5$	$+50^\circ32'38''$	22–25	0.539	0.567	0.553	0.024	0.361
34	Abell 1835	0.253	$14^{\rm h}01^{\rm m}02\overset{\rm s}{.}1$	$+02^\circ52'43''$	22–25	0.562	0.588	0.575	0.021	0.385
35	Abell 1914	0.171	$14^{\rm h}26^{\rm m}02\overset{\rm s}{.}8$	$+37^\circ49'28''$	22–25	0.685	0.708	0.697	0.017	0.525
36	Abell 1942	0.224	$14^{\rm h}38^{\rm m}21\overset{\rm s}{.}9$	$+03^\circ40'13''$	22–25	0.607	0.633	0.620	0.021	0.434
37	Abell 2104	0.153	$15^{\rm h}40^{\rm m}07\overset{\rm s}{.}9$	$-03^\circ18'16''$	22–25	0.707	0.727	0.717	0.014	0.552
38	Abell 2111	0.229	$15^{\rm h}39^{\rm m}40\overset{\rm s}{.}5$	$+34^\circ25'27''$	22–25	0.599	0.625	0.612	0.021	0.425
39	Abell 2163	0.203	$16^{\rm h}15^{\rm m}49\overset{\rm s}{.}0$	$-06^\circ08'41''$	22–25	0.619	0.639	0.629	0.016	0.445
40	Abell 2204	0.152	$16^{\rm h}32^{\rm m}47\overset{\rm s}{.}0$	$+05^\circ34'33''$	22–25	0.708	0.728	0.718	0.014	0.554
41	Abell 2259	0.164	$17^{\rm h}20^{\rm m}09\overset{\rm s}{.}7$	$+27^\circ40'08''$	22–25	0.690	0.711	0.700	0.015	0.530
42	Abell 2261	0.224	$17^{\rm h}22^{\rm m}27\overset{\rm s}{.}2$	$+32^\circ07'58''$	22–25	0.606	0.632	0.619	0.020	0.433
43	Abell 2537	0.295	$23^{\rm h}08^{\rm m}22\overset{\rm s}{.}2$	$-02^\circ11'32''$	22–25	0.511	0.537	0.524	0.023	0.332
44	MS0440.5+0204	0.190	$04^{\rm h}43^{\rm m}09\overset{\rm s}{.}9$	$+02^\circ10'19''$	22–25	0.646	0.667	0.656	0.017	0.476
45	MS0451.6−0305	0.550	$04^{\rm h}54^{\rm m}10\overset{\rm s}{.}8$	$-03^\circ00'51''$	22–25	0.283	0.307	0.295	0.024	0.140
46	MS1008.1−1224	0.301	$10^{\rm h}10^{\rm m}32\overset{\rm s}{.}3$	$-12^\circ39'53''$	22–25	0.504	0.531	0.517	0.024	0.326
47	RXJ1347.5−1145	0.451	$13^{\rm h}47^{\rm m}30\overset{\rm s}{.}1$	$-11^\circ45'09''$	22–25	0.346	0.371	0.358	0.024	0.187
48	RXJ1524.6+0957	0.516	$15^{\rm h}24^{\rm m}38\overset{\rm s}{.}3$	$+09^\circ57'43''$	22–25	0.297	0.321	0.309	0.024	0.150
49	MACS J0717.5+3745	0.548	$07^{\rm h}17^{\rm m}30\overset{\rm s}{.}4$	$+37^\circ45'38''$	22–25	0.269	0.291	0.280	0.022	0.131
50	MACS J0913.7+4056	0.442	$09^{\rm h}13^{\rm m}45\overset{\rm s}{.}5$	$+40^\circ56'29''$	22–25	0.366	0.393	0.380	0.026	0.203
51	CIZA J1938+54	0.260	$19^{\rm h}38^{\rm m}18\overset{\rm s}{.}1$	$+54^\circ09'40''$	22–25	0.550	0.574	0.562	0.021	0.371
52	3C295	0.460	$14^{\rm h}11^{\rm m}20\overset{\rm s}{.}6$	$+52^\circ12'10''$	22–25	0.343	0.368	0.356	0.025	0.185

Column 2: cluster name; column 3: cluster redshift; column 4,5: right ascension and declination (J2000.0) of the adopted cluster centre. In all but four cases (Abell 520, Abell 851, Abell 1758 and Abell 1914) we take this to be the position of the brightest cluster galaxy (BCG). Column 6: magnitude range used for the source galaxies. For clusters 1–20 this is the R_C filter and r' for the remaining clusters. Column 7: the average value of $\beta = D_{\rm ls}/D_{\rm s}$ based on the photo-z analysis presented here; column 8: the values for β measured as described in Ilbert et al. (2013); column 9: the value for β we use to estimate masses, which is the average of the two measurements; column 10: estimate for the systematic uncertainty in β as described in the text; column 11: average value for $\langle\beta^2\rangle$.

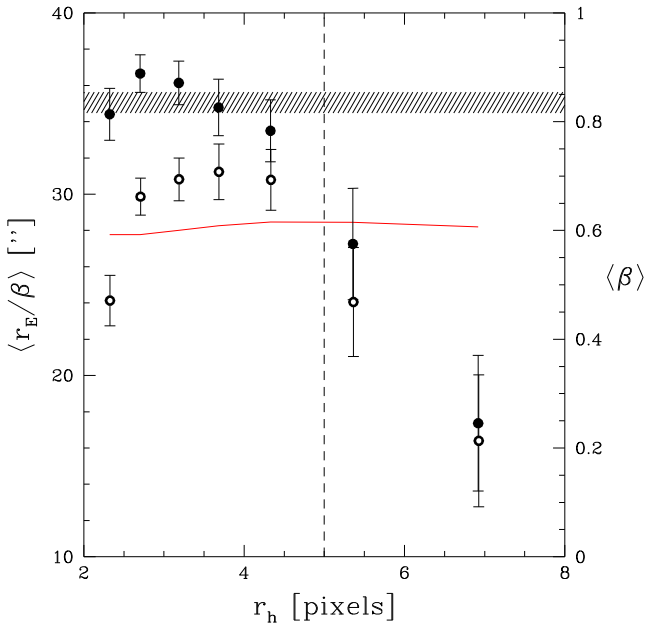

Figure 14. The ensemble averaged lensing signal $\langle r_E/\beta \rangle$ as a function of observed source half-light radius r_h. The value for the Einstein radius r_E was obtained by fitting an SIS model to the lensing signal at radii 0.75-4 h_{70}^{-1} Mpc. The solid points, which should not depend on source size, indicate the results when we apply our correction for the multiplicative bias, whereas the open points are the uncorrected values. The lensing signal is biased low for galaxies with $r_h > 5$, similar to what we observed in our simulated data. The red line indicates the estimate of $\langle \beta \rangle$ as a function of source size (the values are indicated on the rightmost vertical axis), which does not vary significantly with object size, suggesting that the adopted redshift distribution is not very sensitive to the size cuts applied.

H12 used galaxies fainter than $R_C = 25$, for which the redshift distribution is not well-constrained. We now limit the analysis to galaxies brighter than $R_C = 25$.

4.1 Parametric mass models

One approach to infer masses is to fit parametrized models to the lensing signal. The most commonly used profile is the NFW fitting function proposed by Navarro, Frenk & White (1997), which is a good description of the average density profiles of haloes in numerical simulations of structure formation in cold dark matter dominated universes. It also describes the stacked lensing signal for ensembles of clusters well (e.g. Okabe et al. 2013; Umetsu et al. 2014). The NFW profile is characterized by two parameters. We use the mass of the halo and the concentration c, although we do not fit for these parameters simultaneously: we use the fact that simulations show that the mass and concentration are correlated. However, as the concentration depends on the formation redshift of the halo, this relation depends on cosmology. H12 used the results from Duffy et al. (2008), which are based on the cosmological parameters from the five-year *Wilkinson Microwave Anisotropy Probe* observations (*WMAP*5; Komatsu et al. 2009). These have since been superseded by the measurements of the *Planck* satellite (Planck Collaboration XVI 2014b). Dutton & Macciò (2014) present fitting functions for the mass–concentration relation for this cosmology, which we use when we estimate cluster masses (cf. Table 2).

Becker & Kravtsov (2011) have shown that fitting an NFW model to the observed lensing signal can lead to mass estimates that are biased low when measurements at large radii are included. For this reason we restrict the fit to 0.5-2 h_{70}^{-1} Mpc from the cluster, where biases should be negligible. The resulting masses,[8] for different overdensities Δ are presented in columns 9–11 in Table 2. The statistical uncertainties on the measurements are estimated as described in Hoekstra (2007) and H12. The uncertainties in the mass estimates include the contribution from distant large-scale structure (Hoekstra 2001, 2003; Hoekstra et al. 2011). For reference with other studies we also present the velocity dispersion corresponding to the best-fitting singular isothermal sphere (SIS).

For reference we note that if we had used the mass–concentration relation from Duffy et al. (2008), which yields concentrations that are ∼20 per cent lower compared to the values used here, our masses would change as follows: M_{2500} decreases on average by 7 per cent, while M_{500} and M_{200} increase by 5 and 9 per cent, respectively. The relative change in mass does not depend significantly on the cluster redshift. However, for a direct comparison with the existing literature, we present mass measurements using the Duffy et al. (2008) mass–concentration relation for the *WMAP*5 cosmology in Appendix F.

4.1.1 *Comparison to other weak lensing studies*

Several studies have determined weak lensing masses for large samples of clusters using observations with the Subaru telescope. The most relevant for the comparison with CCCP is the WtG project, described in von der Linden et al. (2014a), which targeted 51 massive clusters. For a subset of the clusters WtG determined photometric redshifts for the sources (see Kelly et al. 2014, for details). However, these are predominantly the high-redshift systems where the overlap with CCCP is limited. For this reason we compare with the 'colour-cut' masses, which are presented in Applegate et al. (2014).

A closer inspection of the sample studied by von der Linden et al. (2014a) and Applegate et al. (2014) shows that they associated MS0906.5+1110 with the cluster Abell 750 which is located only 3 arcmin away in projection. However, as discussed by Rines et al. (2013) the latter is a different cluster, which is clearly separated in redshift. The location of A750 provided in von der Linden et al. (2014a) is in fact that of MS0906.5+1110, and we therefore include this cluster in the comparison. Abell 1758 is a merging cluster and therefore H12 considered the Eastern and Western component separately (also see Ragozzine et al. 2012). However, other studies consider this a single cluster and we therefore decided to provide results for the location listed by von der Linden et al. (2014a), who refer to this cluster as Abell 1758N. As a result we have 18 clusters in common with WtG.

To compare the results for these clusters, we follow Applegate et al. (2014) and fit an NFW model with a concentration $c_{200} = 4$ to the tangential distortion within $0.75-3\,h_{70}^{-1}$ Mpc and compute the mass within a sphere of radius $1.5\,h_{70}^{-1}$ Mpc. The results are presented in Fig. 15. We find that the WtG masses are somewhat larger: the dashed line indicates the best-fitting linear relation $M_{\rm WtG} = (1.082 \pm 0.038) M_{\rm CCCP}$. Repeating the comparison using the results from H12 yields $M_{\rm WtG} = (1.263 \pm 0.048) M_{\rm H12}$. Hence the analysis presented here reduces the discrepancy considerably. We note that differences in the fitting procedure can lead to additional uncertainty, and it is therefore not clear whether the

[8] M_Δ is the mass enclosed within a radius where the mean density of the halo is Δ times the critical density at the redshift of the cluster; the virial mass is defined relative to the background density. See Hoekstra (2007) for more details about our choice of definition.

Table 2. Weak lensing mass estimates for the CCCP sample.

(1) Name	(2) σ	(3) $M_{0.5}^{\rm proj}$	(4) $M_{1.0}^{\rm proj}$	(5) r_{2500}	(6) $M_{2500}^{\rm ap}$	(7) r_{500}	(8) $M_{500}^{\rm ap}$	(9) $M_{\rm vir}^{\rm NFW}$	(10) $M_{2500}^{\rm NFW}$	(11) $M_{500}^{\rm NFW}$
Abell 68	1117^{+67}_{-71}	4.5 ± 0.5	8.3 ± 1.3	552	$3.1^{+0.4}_{-0.4}$	1380	$9.7^{+1.9}_{-2.0}$	$12.9^{+2.7}_{-2.7}$	$3.0^{+0.6}_{-0.6}$	$7.5^{+1.6}_{-1.6}$
Abell 209	970^{+78}_{-84}	3.2 ± 0.5	7.3 ± 1.2	490	$2.1^{+0.4}_{-0.4}$	1230	$6.5^{+1.4}_{-1.3}$	$8.4^{+2.2}_{-2.4}$	$2.0^{+0.5}_{-0.6}$	$4.9^{+1.3}_{-1.4}$
Abell 267	1006^{+85}_{-92}	3.9 ± 0.5	7.2 ± 1.4	525	$2.6^{+0.4}_{-0.4}$	1203	$6.3^{+1.8}_{-1.6}$	$7.9^{+2.4}_{-2.3}$	$1.9^{+0.6}_{-0.6}$	$4.7^{+1.4}_{-1.3}$
Abell 370	1489^{+75}_{-79}	6.9 ± 0.7	15.3 ± 1.6	638	$5.5^{+0.8}_{-0.8}$	1637	$18.5^{+2.3}_{-2.3}$	$30.4^{+5.5}_{-5.3}$	$6.5^{+1.2}_{-1.1}$	$17.5^{+3.2}_{-3.1}$
Abell 383	821^{+117}_{-135}	2.5 ± 0.6	6.1 ± 1.3	430	$1.4^{+0.6}_{-0.5}$	1217	$6.2^{+2.2}_{-2.2}$	$5.8^{+2.6}_{-2.4}$	$1.4^{+0.6}_{-0.6}$	$3.4^{+1.5}_{-1.4}$
Abell 963	1106^{+76}_{-82}	3.4 ± 0.5	6.6 ± 1.4	506	$2.3^{+0.5}_{-0.4}$	1185	$5.8^{+1.5}_{-1.5}$	$12.3^{+3.0}_{-3.0}$	$2.9^{+0.7}_{-0.7}$	$7.1^{+1.7}_{-1.7}$
Abell 1689	1429^{+59}_{-62}	7.0 ± 0.5	13.2 ± 1.4	702	$5.9^{+0.7}_{-0.7}$	1571	$13.3^{+2.4}_{-2.2}$	$30.9^{+5.0}_{-4.8}$	$6.6^{+1.1}_{-1.0}$	$17.3^{+2.8}_{-2.7}$
Abell 1763	1229^{+70}_{-75}	4.9 ± 0.6	10.0 ± 1.4	604	$3.9^{+0.6}_{-0.6}$	1511	$12.3^{+3.2}_{-2.9}$	$16.9^{+3.5}_{-3.5}$	$3.8^{+0.8}_{-0.8}$	$9.7^{+2.0}_{-2.0}$
Abell 2218	1181^{+77}_{-82}	5.0 ± 0.6	8.7 ± 1.4	630	$4.2^{+0.7}_{-0.7}$	1379	$8.9^{+2.2}_{-2.1}$	$16.4^{+3.8}_{-3.6}$	$3.7^{+0.9}_{-0.8}$	$9.4^{+2.2}_{-2.1}$
Abell 2219	1041^{+75}_{-80}	4.1 ± 0.5	9.5 ± 1.4	552	$3.0^{+0.5}_{-0.5}$	1408	$10.0^{+2.0}_{-1.8}$	$11.3^{+2.5}_{-2.5}$	$2.7^{+0.6}_{-0.6}$	$6.6^{+1.5}_{-1.5}$
Abell 2390	1331^{+61}_{-64}	4.9 ± 0.5	9.9 ± 1.3	602	$3.9^{+0.5}_{-0.5}$	1351	$8.8^{+1.5}_{-1.5}$	$23.1^{+3.8}_{-3.6}$	$5.1^{+0.8}_{-0.8}$	$13.2^{+2.1}_{-2.1}$
MS 0015.9+1609	1456^{+117}_{-127}	6.8 ± 0.6	17.9 ± 1.9	601	$5.6^{+0.7}_{-0.8}$	1617	$21.8^{+3.2}_{-3.2}$	$28.9^{+7.0}_{-6.8}$	$6.2^{+1.5}_{-1.5}$	$16.9^{+4.1}_{-4.0}$
MS 0906.5+1110	1077^{+70}_{-74}	3.8 ± 0.5	8.7 ± 1.2	549	$2.8^{+0.5}_{-0.5}$	1423	$9.7^{+1.4}_{-1.6}$	$12.6^{+2.9}_{-2.8}$	$2.9^{+0.7}_{-0.6}$	$7.3^{+1.7}_{-1.6}$
MS 1224.7+2007	860^{+118}_{-136}	1.8 ± 0.6	2.4 ± 1.8	345	$0.8^{+0.4}_{-0.7}$	782	$1.9^{+1.0}_{-0.9}$	$4.9^{+2.3}_{-2.1}$	$1.2^{+0.6}_{-0.5}$	$3.0^{+1.4}_{-1.3}$
MS 1231.3+1542	590^{+115}_{-141}	1.0 ± 0.5	0.4 ± 1.3	344	$0.7^{+0.2}_{-0.2}$	565	$0.6^{+0.5}_{-0.4}$	$1.9^{+1.3}_{-1.1}$	$0.5^{+0.3}_{-0.3}$	$1.1^{+0.8}_{-0.7}$
MS 1358.4+6245	1167^{+74}_{-79}	4.3 ± 0.5	8.7 ± 1.5	529	$3.0^{+0.5}_{-0.5}$	1291	$8.6^{+2.0}_{-2.0}$	$13.4^{+3.2}_{-3.1}$	$3.1^{+0.7}_{-0.7}$	$7.8^{+1.9}_{-1.8}$
MS 1455.0+2232	1131^{+63}_{-66}	3.7 ± 0.4	7.2 ± 1.2	510	$2.5^{+0.4}_{-0.4}$	1158	$5.7^{+1.2}_{-1.1}$	$13.2^{+2.4}_{-2.6}$	$3.1^{+0.6}_{-0.6}$	$7.7^{+1.4}_{-1.5}$
MS 1512.4+3647	733^{+111}_{-130}	1.5 ± 0.6	4.5 ± 1.4	282	$0.5^{+0.3}_{-0.3}$	853	$2.6^{+1.5}_{-1.6}$	$3.9^{+1.7}_{-1.7}$	$1.0^{+0.4}_{-0.4}$	$2.4^{+1.1}_{-1.1}$
MS 1621.5+2640	1300^{+83}_{-89}	4.9 ± 0.6	11.4 ± 1.7	543	$3.6^{+0.8}_{-0.8}$	1286	$9.5^{+2.0}_{-1.9}$	$19.1^{+4.2}_{-4.1}$	$4.3^{+0.9}_{-0.9}$	$11.2^{+2.5}_{-2.4}$
CL 0024.0+1652	1311^{+94}_{-101}	5.6 ± 0.6	11.4 ± 1.7	571	$4.0^{+0.6}_{-0.6}$	1333	$10.2^{+2.4}_{-2.2}$	$24.4^{+5.7}_{-5.4}$	$5.3^{+1.3}_{-1.2}$	$14.1^{+3.3}_{-3.1}$
Abell 115N	833^{+89}_{-99}	1.4 ± 0.4	5.3 ± 1.1	283	$0.4^{+0.3}_{-0.4}$	1098	$4.6^{+1.0}_{-1.1}$	$5.9^{+2.0}_{-2.0}$	$1.5^{+0.5}_{-0.5}$	$3.5^{+1.2}_{-1.2}$
Abell 115S	859^{+82}_{-91}	2.6 ± 0.4	5.9 ± 1.2	416	$1.2^{+0.4}_{-0.5}$	1127	$5.0^{+1.3}_{-1.2}$	$7.0^{+2.0}_{-2.0}$	$1.7^{+0.5}_{-0.5}$	$4.1^{+1.2}_{-1.2}$
Abell 222	916^{+85}_{-93}	2.9 ± 0.5	6.9 ± 1.4	450	$1.6^{+0.6}_{-0.8}$	1174	$5.7^{+1.5}_{-1.3}$	$6.4^{+2.1}_{-2.1}$	$1.6^{+0.5}_{-0.5}$	$3.8^{+1.3}_{-1.3}$
Abell 223N	989^{+80}_{-86}	3.0 ± 0.5	7.5 ± 1.3	463	$1.7^{+0.5}_{-0.6}$	1236	$6.6^{+1.3}_{-1.3}$	$8.9^{+2.5}_{-2.5}$	$2.1^{+0.6}_{-0.6}$	$5.2^{+1.5}_{-1.5}$
Abell 223S	923^{+90}_{-99}	3.0 ± 0.5	8.3 ± 1.1	466	$1.8^{+0.5}_{-0.5}$	1370	$9.0^{+1.5}_{-1.5}$	$7.8^{+2.6}_{-2.4}$	$1.9^{+0.6}_{-0.6}$	$4.6^{+1.5}_{-1.4}$
Abell 520	1144^{+64}_{-67}	3.6 ± 0.4	7.3 ± 1.1	526	$2.5^{+0.5}_{-0.4}$	1208	$6.1^{+1.2}_{-1.1}$	$15.3^{+3.0}_{-3.0}$	$3.5^{+0.7}_{-0.7}$	$8.8^{+1.7}_{-1.7}$
Abell 521	944^{+94}_{-103}	3.1 ± 0.5	9.0 ± 1.4	448	$1.7^{+0.7}_{-0.8}$	1335	$8.8^{+2.0}_{-1.9}$	$10.7^{+3.0}_{-3.0}$	$2.5^{+0.7}_{-0.7}$	$6.3^{+1.7}_{-1.7}$
Abell 586	804^{+107}_{-122}	2.5 ± 0.6	6.1 ± 1.6	441	$1.4^{+0.5}_{-0.4}$	1221	$6.1^{+2.6}_{-2.6}$	$4.6^{+2.1}_{-2.0}$	$1.2^{+0.5}_{-0.5}$	$2.8^{+1.3}_{-1.2}$
Abell 611	995^{+94}_{-103}	3.7 ± 0.5	9.0 ± 1.4	502	$2.4^{+0.5}_{-0.5}$	1236	$7.2^{+1.5}_{-1.4}$	$9.4^{+2.8}_{-2.8}$	$2.2^{+0.7}_{-0.7}$	$5.5^{+1.6}_{-1.6}$
Abell 697	1146^{+74}_{-79}	4.6 ± 0.5	10.5 ± 1.4	565	$3.4^{+0.6}_{-0.6}$	1431	$11.2^{+1.5}_{-1.7}$	$14.1^{+3.1}_{-3.1}$	$3.3^{+0.7}_{-0.7}$	$8.2^{+1.8}_{-1.8}$
Abell 851	1328^{+91}_{-98}	5.4 ± 0.6	12.2 ± 1.6	553	$3.7^{+0.5}_{-0.5}$	1362	$11.1^{+2.2}_{-2.1}$	$21.4^{+5.3}_{-4.9}$	$4.7^{+1.2}_{-1.1}$	$12.5^{+3.1}_{-2.8}$
Abell 959	1257^{+70}_{-74}	5.0 ± 0.5	10.8 ± 1.4	596	$4.0^{+0.6}_{-0.6}$	1343	$9.2^{+1.6}_{-1.6}$	$19.7^{+3.8}_{-3.6}$	$4.4^{+0.8}_{-0.8}$	$11.4^{+2.2}_{-2.1}$
Abell 1234	969^{+77}_{-84}	2.5 ± 0.5	4.4 ± 1.3	447	$1.5^{+0.3}_{-0.3}$	989	$3.2^{+1.2}_{-1.0}$	$7.6^{+2.3}_{-2.1}$	$1.9^{+0.6}_{-0.5}$	$4.5^{+1.3}_{-1.2}$
Abell 1246	921^{+78}_{-85}	2.7 ± 0.4	5.6 ± 1.1	440	$1.5^{+0.3}_{-0.4}$	1089	$4.4^{+1.0}_{-1.0}$	$8.7^{+2.4}_{-2.2}$	$2.1^{+0.6}_{-0.5}$	$5.1^{+1.4}_{-1.3}$
Abell 1758	1278^{+60}_{-62}	5.5 ± 0.5	12.1 ± 1.4	651	$5.2^{+0.7}_{-0.7}$	1507	$12.9^{+1.9}_{-1.9}$	$19.4^{+3.2}_{-3.2}$	$4.3^{+0.7}_{-0.7}$	$11.2^{+1.9}_{-1.9}$
Abell 1835	1295^{+65}_{-68}	5.3 ± 0.5	10.7 ± 1.3	618	$4.3^{+0.5}_{-0.5}$	1398	$10.0^{+1.6}_{-1.6}$	$19.9^{+3.7}_{-3.5}$	$4.4^{+0.8}_{-0.8}$	$11.4^{+2.1}_{-2.0}$
Abell 1914	1098^{+57}_{-60}	3.7 ± 0.5	7.8 ± 1.2	531	$2.5^{+0.4}_{-0.4}$	1293	$7.3^{+1.3}_{-1.3}$	$13.5^{+2.4}_{-2.4}$	$3.1^{+0.6}_{-0.6}$	$7.8^{+1.4}_{-1.4}$
Abell 1942	1080^{+70}_{-74}	3.8 ± 0.6	7.5 ± 1.3	531	$2.7^{+0.6}_{-0.5}$	1212	$6.4^{+1.4}_{-1.3}$	$13.5^{+2.8}_{-2.6}$	$3.1^{+0.6}_{-0.6}$	$7.8^{+1.6}_{-1.5}$
Abell 2104	1135^{+71}_{-76}	4.2 ± 0.5	10.2 ± 1.2	596	$3.5^{+0.6}_{-0.6}$	1426	$9.6^{+1.7}_{-1.4}$	$15.7^{+3.5}_{-3.3}$	$3.6^{+0.8}_{-0.8}$	$9.0^{+2.0}_{-1.9}$
Abell 2111	996^{+77}_{-83}	3.9 ± 0.5	6.6 ± 1.5	528	$2.6^{+0.5}_{-0.5}$	1170	$5.7^{+1.9}_{-1.8}$	$9.4^{+2.4}_{-2.2}$	$2.3^{+0.6}_{-0.5}$	$5.5^{+1.4}_{-1.3}$
Abell 2163	1188^{+74}_{-79}	4.4 ± 0.4	9.4 ± 1.2	574	$3.3^{+0.4}_{-0.4}$	1466	$11.0^{+2.0}_{-2.0}$	$17.4^{+3.8}_{-3.6}$	$3.9^{+0.9}_{-0.8}$	$10.0^{+2.2}_{-2.1}$
Abell 2204	1229^{+56}_{-58}	4.8 ± 0.5	10.8 ± 1.2	631	$4.2^{+0.5}_{-0.6}$	1491	$11.0^{+1.6}_{-1.5}$	$19.9^{+3.2}_{-3.1}$	$4.4^{+0.7}_{-0.7}$	$11.3^{+1.8}_{-1.7}$
Abell 2259	932^{+89}_{-98}	2.4 ± 0.6	5.6 ± 1.2	427	$1.3^{+0.5}_{-0.4}$	1113	$4.6^{+1.2}_{-1.1}$	$7.9^{+2.6}_{-2.4}$	$1.9^{+0.6}_{-0.6}$	$4.6^{+1.5}_{-1.4}$
Abell 2261	1307^{+65}_{-68}	6.0 ± 0.5	14.1 ± 1.4	682	$5.7^{+0.6}_{-0.6}$	1663	$16.4^{+1.9}_{-1.9}$	$24.4^{+4.1}_{-3.9}$	$5.3^{+0.9}_{-0.9}$	$13.9^{+2.3}_{-2.2}$
Abell 2537	1285^{+71}_{-75}	5.4 ± 0.6	10.0 ± 1.4	599	$4.1^{+0.6}_{-0.6}$	1312	$8.7^{+1.6}_{-1.5}$	$20.9^{+4.0}_{-3.8}$	$4.6^{+0.9}_{-0.9}$	$12.1^{+2.3}_{-2.2}$
MS 0440.5+0204	780^{+112}_{-130}	2.9 ± 0.6	2.8 ± 1.3	468	$1.8^{+0.6}_{-0.5}$	896	$2.5^{+0.7}_{-0.7}$	$3.5^{+1.8}_{-1.8}$	$0.9^{+0.5}_{-0.5}$	$2.1^{+1.1}_{-1.1}$

Table 2 – *continued.*

(1) Name	(2) σ	(3) $M_{0.5}^{\rm proj}$	(4) $M_{1.0}^{\rm proj}$	(5) r_{2500}	(6) $M_{2500}^{\rm ap}$	(7) r_{500}	(8) $M_{500}^{\rm ap}$	(9) $M_{\rm vir}^{\rm NFW}$	(10) $M_{2500}^{\rm NFW}$	(11) $M_{500}^{\rm NFW}$
MS 0451.6−0305	1302^{+129}_{-142}	4.4 ± 0.6	8.4 ± 2.1	466	$2.6^{+0.6}_{-0.5}$	1082	$6.5^{+1.7}_{-1.9}$	$17.4^{+5.7}_{-5.3}$	$3.9^{+1.3}_{-1.2}$	$10.3^{+3.4}_{-3.1}$
MS 1008.1−1224	1218^{+74}_{-79}	4.1 ± 0.4	8.2 ± 1.4	520	$2.7^{+0.4}_{-0.4}$	1176	$6.3^{+1.2}_{-1.2}$	$16.3^{+3.5}_{-3.3}$	$3.7^{+0.8}_{-0.8}$	$9.5^{+2.0}_{-1.9}$
RX J1347.5−1145	1358^{+95}_{-102}	5.2 ± 0.7	10.1 ± 2.2	547	$3.8^{+0.9}_{-0.9}$	1323	$10.7^{+3.5}_{-3.7}$	$19.9^{+5.2}_{-5.0}$	$4.4^{+1.2}_{-1.1}$	$11.7^{+3.0}_{-3.0}$
RX J1524.6+0957	839^{+199}_{-258}	2.1 ± 0.7	6.5 ± 2.1	249	$0.4^{+0.3}_{-12.2}$	977	$4.6^{+2.2}_{-2.0}$	$6.2^{+4.2}_{-3.7}$	$1.5^{+1.0}_{-0.9}$	$3.8^{+2.5}_{-2.2}$
MACS J0717.5+3745	1617^{+119}_{-128}	6.4 ± 0.8	19.3 ± 2.3	612	$5.9^{+1.2}_{-1.4}$	1489	$17.1^{+3.2}_{-3.1}$	$38.4^{+8.9}_{-8.6}$	$8.0^{+1.8}_{-1.8}$	$22.3^{+5.2}_{-5.0}$
MACS J0913.7+4056	919^{+143}_{-168}	3.1 ± 0.8	5.3 ± 1.8	397	$1.4^{+0.9}_{-0.7}$	945	$3.9^{+1.3}_{-1.2}$	$6.5^{+3.4}_{-2.9}$	$1.6^{+0.8}_{-0.7}$	$3.9^{+2.0}_{-1.8}$
CIZA J1938+54	1186^{+84}_{-90}	5.3 ± 0.6	11.3 ± 1.5	601	$4.0^{+0.6}_{-0.6}$	1573	$14.4^{+2.4}_{-2.4}$	$17.2^{+4.2}_{-4.0}$	$3.9^{+0.9}_{-0.9}$	$9.9^{+2.4}_{-2.3}$
3C295	1076^{+113}_{-125}	4.6 ± 0.7	8.1 ± 1.9	501	$2.9^{+0.7}_{-0.6}$	1101	$6.2^{+1.6}_{-1.7}$	$12.0^{+3.9}_{-3.8}$	$2.8^{+0.9}_{-0.9}$	$7.1^{+2.3}_{-2.2}$

Column 1: cluster name; column 2: line-of-sight velocity dispersion (in units of km s^{-1}) of the best-fitting SIS model; columns 3 and 4: projected mass within an aperture of radius 0.5 and 1 h_{70}^{-1} Mpc, respectively; columns 5 and 7: r_Δ (in units of h_{70}^{-1}kpc) determined using aperture masses; columns 6 and 8: deprojected aperture masses within r_Δ; columns 9–11: masses from best-fitting NFW model. All masses are listed in units of ($10^{14}\,h_{70}^{-1}\,{\rm M}_\odot$).

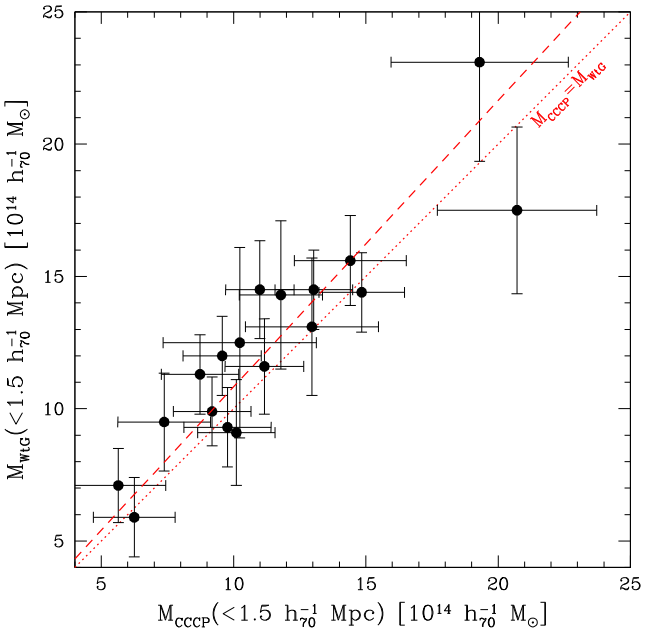

Figure 15. Comparison with the mass estimates from Applegate et al. (2014). The CCCP masses are computed from the best-fitting NFW model to the lensing measurements within $0.75-3\,h_{70}^{-1}$ Mpc, adopting a concentration $c_{200} = 4$. This matches the procedure described in Applegate et al. (2014), although our source redshift distribution is somewhat different, as explained in the text. The dotted line indicates the line of equality, whereas the dashed line is the best fit, which has a slope 1.08.

difference is significant. Furthermore, the 'colour-cut' masses from Applegate et al. (2014) are derived using the photometric redshift catalogue from Ilbert et al. (2009), which are based on the original COSMOS-30 data. Using this redshift distribution we find $M_{\rm WtG} = (1.063 \pm 0.038) M_{\rm CCCP}$. Interestingly, when we compute deprojected aperture masses within a radius of $1\,h_{70}^{-1}$ Mpc (in this case adopting $c_{200} = 4$; see Section 4.2 for details), the agreement with the corresponding masses from Applegate et al. (2014) is excellent: $M_{\rm WtG} = (1.018 \pm 0.036) M_{\rm CCCP}^{\rm ap}$.

The NFW model is fit to relatively small radii, where the contamination by cluster members is large (although the inner $750\,h_{70}^{-1}$ kpc are not used): if we omit the correction for the contamination of cluster members our masses decrease, as expected, and $M_{\rm WtG} \sim 1.28 \times M_{\rm CCCP}$. Although the correction is substantial, Fig. 12 suggests that the bias after correction should be <2 per cent.

We investigated this further by restricting the fit to small ($r_{\rm in} = 0.75-1.5\,h_{70}^{-1}$ Mpc) and large ($r_{\rm out} = 1.5-3\,h_{70}^{-1}$ Mpc) radii. If the contamination correction is adequate, the resulting average masses should agree, whereas a ratio $M_{\rm out}/M_{\rm in} > 1$ would imply residual contamination by cluster members. For the 18 clusters in common with WtG we find $M_{\rm out}/M_{\rm in} = 1.05 \pm 0.05$, suggesting that the correction has worked well (the ratio is 1.16 ± 0.05 if we do not correct for contamination). For the full CCCP sample we find $M_{\rm out}/M_{\rm in} = 1.00 \pm 0.03$.

Umetsu et al. (2014) present results for 20 clusters studied as part of the Cluster Lensing and Supernova survey with Hubble (CLASH). Of these 17 clusters overlap with WtG, but only 6 overlap with CCCP. For the six clusters we have in common, the CLASH masses are 12 ± 5 per cent higher than the CCCP results. However, we note that Umetsu et al. (2014) use a different fit range, while leaving the concentration a free parameter. Although they find a best-fitting concentration of ∼4 when they stack the clusters in their sample, cluster-to-cluster variation complicates a more direct comparison. As discussed below, we analysed the CLASH data using our pipeline and find better agreement with our masses derived from CFHT observations.

Another large study that does overlap considerably with CCCP is the LoCuSS. Results for 30 clusters, of which 13 overlap with CCCP are presented in Okabe et al. (2010). Okabe et al. (2013) suggest that a revised analysis leads to higher masses, but only present results for an ensemble stacked lensing signal and do not provide updated masses for the individual clusters.

4.1.2 *Direct comparison with Subaru data*

The processed Subaru imaging data used by CLASH have been publicly released[9] for 9 of the 20 clusters. We retrieved the data for the four clusters that overlap with CCCP. To extend the comparison sample, Keiichi Umetsu kindly provided us with the data for Abell 209 and Abell 611. Observation of Abell 1758 were provided by James Jee. We analysed these data using our CCCP weak lensing

[9] http://archive.stsci.edu/prepds/clash/

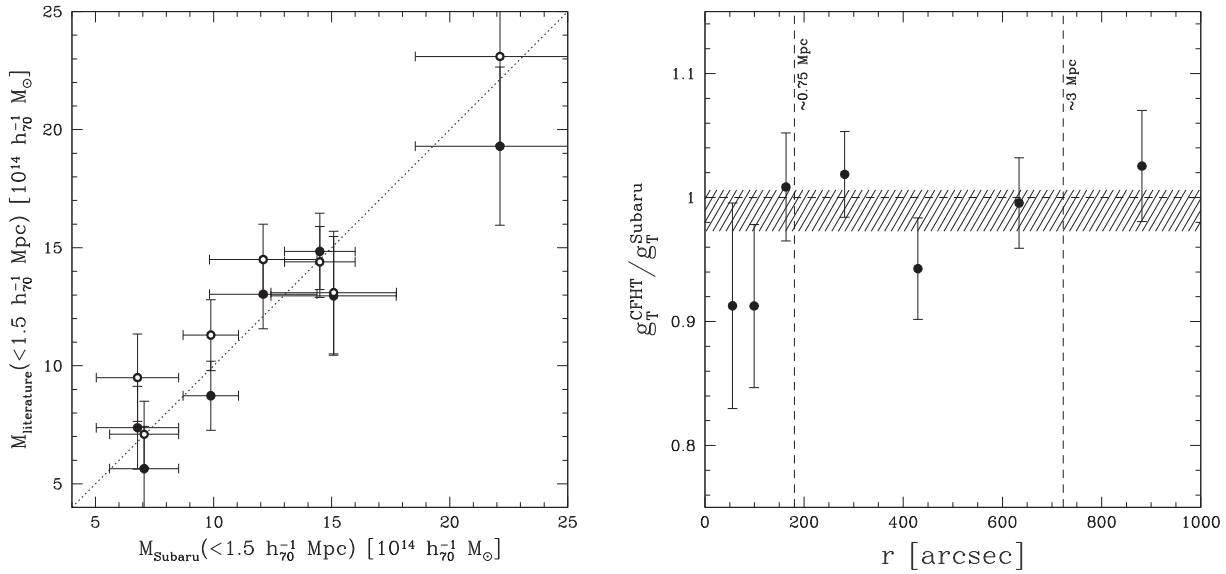

Figure 16. Left-hand panel: black points indicate the mass based on our analysis of CFHT data as a function of the mass obtained from the Subaru data. As opposed to the results presented in the right-hand panel, the object catalogues were not matched. We find that the Subaru-based masses are 4 ± 6 per cent higher than the CFHT-based results we report in this paper. The open points indicate the comparison of our Subaru-based masses with those from Applegate et al. (2014), which are 7 ± 6 per cent higher. The dotted line indicates the line of equality. Right-hand panel: ratio of the ensemble averaged tangential distortion as a function of radius measured from CFHT and Subaru data. The ratio is computed by combining the measurements for the sample of seven clusters, where the catalogues are matched, such that they contain essentially the same objects. The hatched region indicates the 68 per cent confidence region for the average ratio $\langle g_{\rm T}^{\rm CFHT}/g_{\rm T}^{\rm Subaru}\rangle = 0.99 \pm 0.02$. For reference we also indicate the fitting range used in the comparison with Applegate et al. (2014).

pipeline. We made no modifications and thus assume that our empirical correction for noise bias also applies to the typically deeper Subaru data (note that we also apply the size cut of $r_{\rm h} < 5$ pixels in this case). The results presented in Appendix D suggest that our approach, which is a function of SNR and galaxy size, is sufficiently flexible.

The data that were provided are stacks of dithered exposures. As a consequence multiple chips can contribute to a given location, which can lead to a more complex PSF. Our observing strategy with Megacam allowed us to avoid this, but we note that the same problem occurs for our CFH12k data. However, we did not measure noticeable differences in the scaling relations based on CFH12k or Megacam data. As we did for the CFH12k data, we use the weight images to split the data into regions that more or less correspond to the chips of the camera, and analyse the resulting images using the pipeline described earlier. This analysis is done completely independently from the analysis based on the CFHT data. Hence we redo the object detection and masking, identify the stars which are used to model the PSF, etc.

The results are presented in Fig. 16. The left-hand panel shows a comparison of the weak lensing masses when the source catalogue is determined independently from the CFHT analysis. Following Applegate et al. (2014), these masses are based on the best-fitting NFW model to the lensing measurements within $0.75{-}3\,h_{70}^{-1}$ Mpc, adopting a concentration $c_{200} = 4$. For most clusters the Subaru data are deeper, resulting in a different effective source redshift distribution, which we account for. Consequently the correction for contamination by cluster members also differs somewhat. It is important to stress this correction works well on average (as shown in Fig. 12), but the statistical uncertainty is larger when comparing a small sample of clusters.

The filled points in the left-hand panel of Fig. 16 compare the masses based on the CFHT data ($M_{\rm literature}$) to those determined from our analysis of the Subaru data ($M_{\rm Subaru}$). We find excellent agreement with a best-fitting $M_{\rm CFHT} = (0.97 \pm 0.06) M_{\rm Subaru}$. We can use this result as a measure of the systematic uncertainty when cluster masses are determined independently using different instruments, albeit with the same shape measurement pipeline. This comparison is made by fitting a parametric model to the lensing signal at relatively small radii, where our correction for contamination by cluster members is largest. As discussed in more detail in Section 4.3, we expect our aperture masses to be more reliable. We find that the masses from Applegate et al. (2014), indicated by the open points, are somewhat higher. We obtain $M_{\rm WtG} = (1.07 \pm 0.06) M_{\rm Subaru}$. A similar result is obtained if we compare with the six clusters in common with Umetsu et al. (2014), where we note that our Subaru-based masses are in excellent agreement (they are ∼2.4 per cent lower on average).

To examine the performance of the shape measurement algorithm further, we created a source catalogue where the objects were matched by position (note that we do apply the size cut before matching). Although blending in the inner regions may cause some misidentification, such a comparison should eliminate differences in the source redshift distribution and the contamination by cluster members. The shape measurements for individual galaxies are too noisy, and we therefore compare the tangential shear profiles from the two telescopes. To improve the SNR even further, we combine the signals from the seven clusters. This allowed us to measure the ratio of the lensing signal as a function of distance from the cluster centre. The results are presented in the right-hand panel of Fig. 16. We find an average ratio $\langle g_{\rm T}^{\rm CFHT}/g_{\rm T}^{\rm Subaru}\rangle = 0.99 \pm 0.02$, indicated by the hatched region. This result suggests that our pipeline is able to recover the shapes to within 1 ± 2 per cent for different data sets.

In addition we have matched our CFHT-based measurements to catalogues provided by the WtG team (Von der Linden, private communication). This direct comparison for the overlapping sample of clusters showed a remarkable agreement with $\langle g_{\rm T}^{\rm CCCP}/g_{\rm T}^{\rm WtG}\rangle = 0.991 \pm 0.018$. These direct comparisons of shear

catalogues obtained from observations using different telescopes and teams suggest that the pipelines are robust. The differences we observe are consistent with the statistical uncertainties associated with comparing such a small sample of clusters.

4.1.3 Comparison to Hectospec Cluster Survey

The infall regions of galaxy clusters provide an interesting alternative way to estimate cluster masses at relatively large radii. The Hectospec Cluster Survey targeted a sample of 58 clusters, 14 of which overlap with our study. This survey is described in Rines et al. (2013), who measured cluster-mass profiles using the caustic technique (e.g. Diaferio 1999). We compare the estimates for M_{200} from the best-fitting NFW models to those obtained by Rines et al. (2013). The results are presented in Fig. 17. We note that a comparison to the velocity dispersions yields similar results.

The lensing masses are higher than the dynamical masses; the average ratio is 1.22 ± 0.07. If we adopt the mass–concentration relation from Duffy et al. (2008) the agreement is worse with an average ratio of 1.33 ± 0.08. Although the comparison sample is small, we observe substantial scatter. The most significant outliers in Fig. 17 are MS0906.5+1110, Abell 1758 and Abell 2261. In section 3.6 of their paper, Rines et al. (2013) comment on individual clusters, including these outliers. The first is part of a pair of clusters, but the infall patterns can be separated. Rines et al. (2013) find the higher mass for the other component, Abell 750, but comment that MS0906.5+1110 has the higher X-ray luminosity. Okabe et al. (2010) also find a higher lensing mass for this cluster, suggesting that the dynamical mass is too low. Coe et al. (2012) present a detailed study of Abell 2261, suggesting that the dynamical mass estimate for Abell 2261 is low compared to the lensing and X-ray estimates. We note the large range in hydrostatic mass estimates for this cluster, including the measurement from Mahdavi et al. (2013); although lower than the lensing mass, the X-ray mass is still larger than the dynamical mass from Rines et al. (2013). Finally Abell 1758 is a merging system, at the high-redshift end of the sample studied by Rines et al. (2013). As a result the dynamical mass is not that well-constrained.

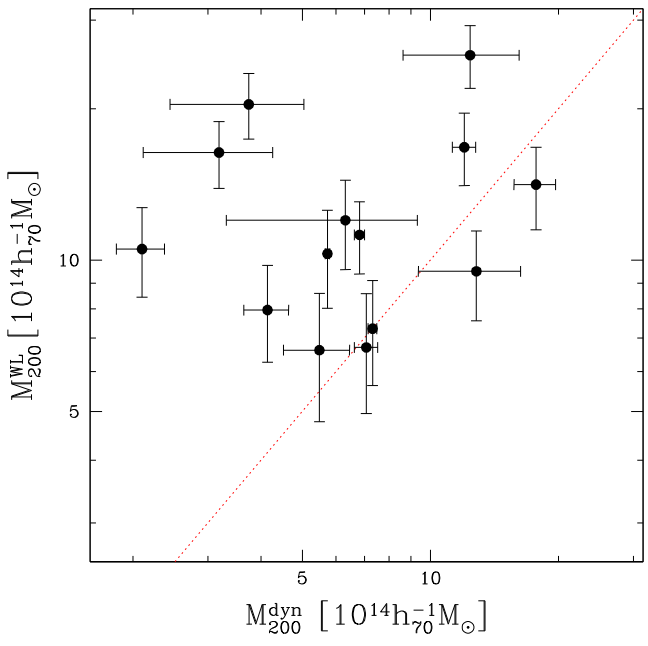

Figure 17. Comparison of the dynamical estimates for M_{200} from the Hectospec Cluster Survey from Rines et al. (2013) with the best-fitting NFW model to the lensing data. The dotted line indicates the line of equality.

4.2 Aperture masses

Although the NFW profile is a good description for an ensemble of clusters, as is suggested by stacked weak lensing studies (Okabe et al. 2013; Umetsu et al. 2014), it may not be a good model to fit to individual systems. In particular the presence of substructure may lead to incorrect masses (see e.g. Hoekstra et al. 2000, for a clear example). The CCCP sample contains several complex, or merging clusters, of which Abell 520 has been studied in particular detail (Mahdavi et al. 2007; Jee et al. 2012, 2014). In these cases a more direct estimate of the projected mass, without having to rely on a particular profile, would be preferable.

It is possible to estimate the projected mass within an aperture with minimal assumptions about the actual mass distribution. However, comparison to other proxies typically still relies on deprojected masses, which do depend on the assumed density profile. Various estimators are available, but we prefer to use the one proposed by Clowe et al. (1998):

$$\zeta_c(r_1) = 2\int_{r_1}^{r_2} \mathrm{d}\ln r \langle\gamma_t\rangle + \frac{2r_{\rm max}^2}{r_{\rm max}^2 - r_2^2}\int_{r_2}^{r_{\rm max}} \mathrm{d}\ln r \langle\gamma_t\rangle, \quad (11)$$

which can be expressed in terms of the mean dimensionless surface density interior to r_1 relative to the mean surface density in an annulus from r_2 to $r_{\rm max}$

$$\zeta_c(r_1) = \bar{\kappa}(r' < r_1) - \bar{\kappa}(r_2 < r' < r_{\rm max}). \quad (12)$$

Hence we can determine the mass up to constant, which is determined by the mean convergence in the annulus $r_2 < r' < r_{\rm max}$. Assumptions about the mass distribution enter in two ways. First of all, we do not measure the tangential shear $\gamma_{\rm T}$, but the reduced shear $g_{\rm T} = \gamma_{\rm T}/(1-\kappa)$. For this conversion we use the best-fitting NFW model. However, the estimate of $\zeta_c(r)$ depends on the lensing measurements at large radii and consequently this correction is small.

The more important dependence on the density profile is through the need to estimate the average convergence in the annulus. Although the contribution is relatively small if we consider large radii, it cannot be ignored for our analysis. H12 used an outer radius $r_{\rm max} = 1000$ arcsec for the CFH12k data. However, for these data the azimuthal coverage is incomplete. We therefore keep the inner radius of 600 arcsec, but reduce the outer radius to 800 arcsec. The annuli are unchanged for the Megacam data, i.e. we use $r_2 = 900$ arcsec and $r_{\rm max} = 1500$ arcsec.

As was done by H12, we use the best-fitting NFW model to estimate the mean convergence in the annulus. We quantify the sensitivity to the profile by varying the normalization of the $c(M)$ relation by 20 per cent. This should capture the variation in results for different cosmologies (e.g. Bhattacharya et al. 2013). The resulting change in projected mass within an aperture of $1\,h_{70}^{-1}$ Mpc depends on redshift, with a reduction (increase) of <1 per cent for clusters with $z > 0.3$ if the normalization is increased (decreased). The changes are somewhat larger at lower redshifts but at most only ∼2 per cent. Hence the systematic uncertainty in the estimate of the projected mass at this radius is remarkably small.

We do make one small change with respect to H12: we include a contribution from neighbouring haloes. Oguri & Hamana (2011) show that such a two-halo term dominates over the NFW profile on large scales. A convenient way to describe such contributions

is provided by the so-called *halo model* (see e.g. Cooray & Sheth 2002, for a review). The implementation we use here is described in Cacciato et al. (2013). It was used in Cacciato, van Uitert & Hoekstra (2014) to model the lensing signal around galaxies. In this context, the cluster lensing signal is simply the lensing signal around the brightest cluster galaxy of a very massive halo. The contribution from the clustering of haloes, the two-halo term, is in its most basic implementation proportional to the linear matter power spectrum (but see e.g. van den Bosch et al. 2013, for a more sophisticated implementation). The constant of proportionality is determined by the product of the bias of the halo of interest and an average of the halo bias over all halo masses in the range of interest weighted by the halo mass function (see e.g. Cacciato et al. 2009).

To compute the contribution to the average convergence in the annulus, we use the functions for the halo mass and bias provided by Tinker et al. (2010). We find that the correction is small, less than 1 per cent for the radii we are interested in. This also depends somewhat on the assumed density profile, but the main source of theoretical uncertainty is the halo bias function, especially at the high-mass end considered here. We gauge the systematic uncertainty in the correction by considering different fitting functions from the literature, namely those from Sheth & Tormen (1999), Sheth, Mo & Tormen (2001) and Seljak & Warren (2004). Although these may differ by up to a factor of a few at the highest masses of interest, the corresponding cluster bias varies by at most ∼20 per cent because of the exponential drop-off of the halo mass function. We find that the resulting systematic uncertainty in the correction itself is ∼10 per cent, which can be safely ignored.

The resulting projected masses within fixed apertures of 0.5 and 1 h_{70}^{-1} Mpc are listed in Table 2. The uncertainties include the contributions from cosmic noise and shape noise. Although the variation in the source redshift distribution is in principle an additional source of noise (Hoekstra et al. 2011), the impact is smaller for the larger angular scales used to compute the aperture masses. Furthermore, Hoekstra et al. (2011) find that the combination of cosmic noise and the variation in the redshift distribution leads to a slight reduction in the uncertainty, compared to the situation where only cosmic noise is considered.

In Fig. 18 we compare the projected mass to the results from H12 as a function of cluster redshift. We find no significant trend with redshift and obtain a weighted average ratio $\langle M/M_{\rm H12}\rangle = 1.093 \pm 0.016$. The increase in the amplitude of the lensing signal because of the correction for the effects of noise in the images is partly offset by the increase in the mean source redshift and the change in the mass–concentration relation.

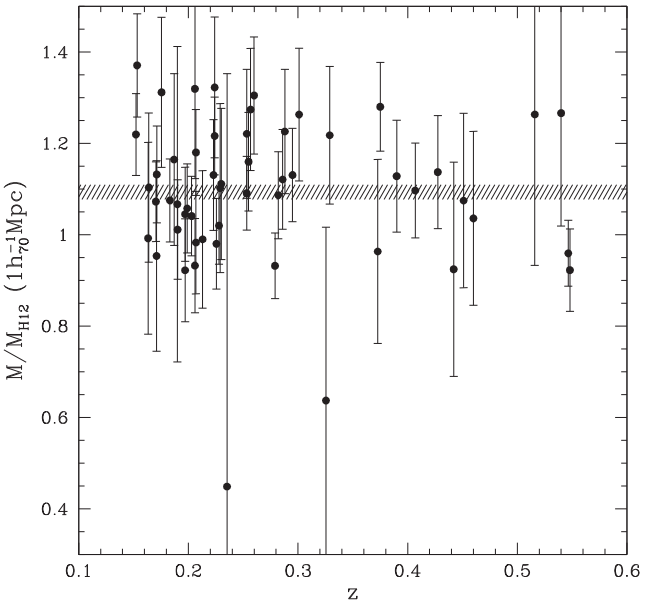

Figure 18. Ratio of the projected aperture mass within an aperture of radius 1 h_{70}^{-1} Mpc and the mass obtained by H12. The hatched region indicates the weighted average ratio $\langle M/M_{\rm H12}\rangle = 1.093 \pm 0.016$. Note that the error bars indicate only the uncertainty in the updated mass estimates.

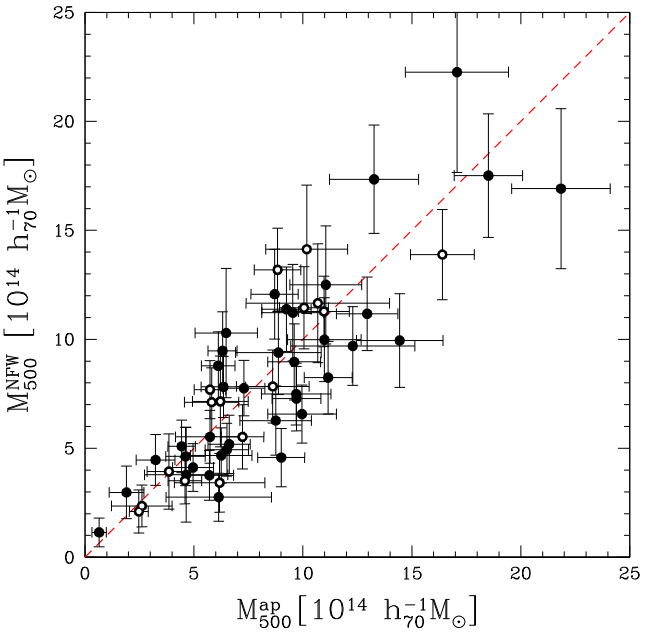

Figure 19. Comparison of $M_{500}^{\rm NFW}$ from the best-fitting NFW model and the value for $M_{500}^{\rm ap}$ using the deprojected aperture mass. The former is based on measurements of the lensing signal at radii 0.5-2 h_{70}^{-1} Mpc, whereas the latter uses data from radii larger than r_{500} which is typically larger than 1 h_{70}^{-1} Mpc. As a consequence the measurements are almost independent. The red dashed line indicates the line of equality. The open points indicate clusters for which Mahdavi et al. (2013) measured a central entropy $K_0 < 70\,{\rm keV\,cm^2}$.

4.2.1 *Deprojected masses*

Although the projected masses can be determined robustly within an aperture of fixed radius, comparison with other observations requires the deprojection of the mass estimates. To do so we follow the procedure described in Hoekstra (2007), which was also used in H12: at each radius we find the NFW model that yields the same projected mass. We take the mass of this model, which depends on the adopted mass–concentration relation, to be the estimate for the deprojected mass. The results for $M_{2500}^{\rm ap}$ and $M_{500}^{\rm ap}$, using the relation between concentration and mass from Dutton & Macciò (2014), are listed in columns 6 and 8 of Table 2 (for reference we also list the corresponding radius r_Δ). We present results for the $c(M)$ relation from Duffy et al. (2008) in Appendix F.

In Fig. 19, we compare $M_{500}^{\rm ap}$ to the estimate obtained from the best-fitting NFW model to the lensing signal at radii 0.5-2 h_{70}^{-1} Mpc. The two measurements are nearly independent, because the aperture mass is based on the lensing signal at radii larger than r_{500} which is larger than 1 h_{70}^{-1} Mpc for most of the clusters. From a linear fit to the

masses we find that $M_{500}^{\rm NFW}/M_{500}^{\rm ap} = 0.97 \pm 0.03$, i.e. the estimates from the NFW model are in good agreement on average.

It is interesting to compare the results for clusters with a low and high central entropy K_0, the value of the deprojected entropy profile at a radius of $20\,h_{70}^{-1}$ kpc, as measured by Mahdavi et al. (2013). The open points in Fig. 19 indicate the clusters with $K_0 < 70\,{\rm keV\,cm^2}$, for which we find a ratio 1.04 ± 0.06; for the remaining clusters $M_{500}^{\rm NFW}$ is on average 0.95 ± 0.04 times smaller than the aperture mass estimate. In neither case do we observe a significant difference between the two mass estimates.

As discussed in H12, the deprojection depends on the mass–concentration relation. The results listed in Table 2 are based on the mass–concentration relation from Dutton & Macciò (2014). If we instead consider the concentrations from Duffy et al. (2008), which are ∼20 per cent lower, we find that the resulting $M_{2500}^{\rm ap}$ is 8 per cent lower, and $M_{500}^{\rm ap}$ is 3 per cent lower. Hence, the value for $M_{500}^{\rm ap}$ is fairly robust against changes in the concentration.

If we compare to the deprojected masses from H12 we find that $M_{500}^{\rm ap}$ has increased by 19 per cent on average, whereas the increase in M_{2500} is 28 per cent. The larger change in M_{2500} is due to the higher concentrations from Dutton & Macciò (2014), which affect the deprojection. As discussed in H12 (see fig. 4 in their paper), lowering the concentration increases the ratio $M_{500}^{\rm NFW}/M_{500}^{\rm ap}$.

4.3 Error budget

Fig. 20 shows the distribution of the relative change in the projected mass within an aperture of radius $1\,h_{70}^{-1}$ Mpc when one of the corrections is not included. The red histogram shows that the average mass would be on average reduced by a factor 0.84 if we do not include the correction for the multiplicative bias in the shape measurement. Ignoring the correction for contamination by cluster members reduces the mass by a factor 0.90. As indicated by the black histogram, this correction varies more from cluster to cluster. This is expected, as the correction depends on the cluster redshift, the spatial distribution of cluster galaxies and the richness of the cluster. If we use the source redshift distribution used by H12 the masses would increase by 4 per cent, as indicated by the blue histogram.

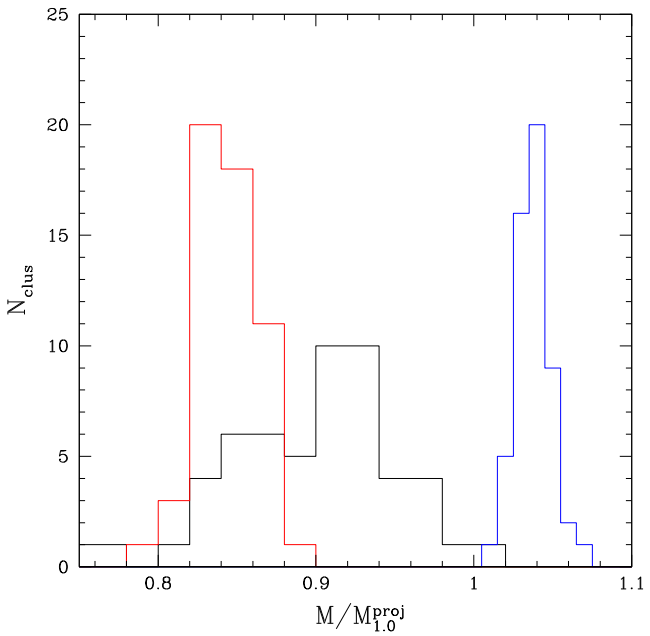

Figure 20. Ratio of the projected aperture mass within an aperture of radius $1\,h_{70}^{-1}$ Mpc when one of the corrections is not included, and the final mass estimate. The red histogram shows the distribution if we ignore the correction for the multiplicative bias in the shape measurement. The black histogram shows the decrease in mass if we ignore the contamination by cluster members. The blue histogram shows the change in mass if we use the source redshift distribution used by H12.

Although each of these corrections is substantial and cannot be ignored, they are well-determined. The associated remaining systematic errors are much smaller than the statistical uncertainties for individual clusters. However, the precision that is afforded by the full CCCP sample is much higher, and we therefore summarize the systematic error budget in this section. These are relevant for the discussion in Section 5 where we examine the scaling relation between the lensing mass and the SZ measurements from *Planck*.

We assume a 2 per cent systematic uncertainty in the calibration of the multiplicative bias. Although Fig. 9 suggests that the empirical correction results in biases <1 per cent if we adopt $\epsilon_0 = 0.25$, the variation in seeing and the dependence of the bias on morphology are expected to lead to additional error. The latter will need to be investigated in more detail, but the observed variation in bias as a function of Sérsic index (see Fig. 8) suggests that the contribution from morphology is ∼1 per cent.

As shown by the black points in Fig. 12 the mean residual contamination by cluster members is at most ∼2 per cent, except for the innermost regions. Although the contamination exceeds this value at some radii for clusters with $z > 0.25$ (red hatched region), the contamination is still smaller than 2 per cent when averaged over the radii of interest. For the projected masses, the uncertainty in the extrapolated density profile changes the masses by less than 1 per cent.

The largest contribution comes from the uncertainty in the source redshift distribution, despite the wavelength coverage and depth of the COSMOS data: for clusters at $z = 0.2$ the systematic uncertainty in $\langle\beta\rangle$ is ∼3 per cent, increasing to ∼8 per cent for the highest redshift clusters in our sample ($z = 0.55$). These estimates include the difference in $\langle\beta\rangle$ between our photometric redshift distribution and that of Ilbert et al. (2013), and the field-to-field variation in the four CFHTLS Deep fields. These systematic errors are independent from one another and thus can be added in quadrature, resulting in a total systematic uncertainty of 4.2 per cent at $z = 0.2$ and 8.5 per cent at $z = 0.55$ for the projected masses in a fixed aperture.

The deprojected masses are more sensitive to the assumed mass–concentration relation. For instance, $M_{500}^{\rm ap}$ decreases by ∼3 per cent if we lower the concentration by 20 per cent (i.e. by switching to the Duffy et al. 2008 values). This shift in concentration is caused by the change in the cosmological parameters determined by *WMAP*5, used by Duffy et al. (2008), and the more precise *Planck* values used by Dutton & Macciò (2014). However, these estimates are based on simulations that only include dark matter and thus ignore the additional effects of baryon physics. The impact of this has been studied by Duffy et al. (2010) using hydrodynamic simulations. Duffy et al. (2010) found that the change in concentration is <10 per cent for cluster-mass haloes. We therefore adopt a similar uncertainty in the concentrations, which implies a systematic contribution of ∼2 per cent for $M_{500}^{\rm ap}$.

Importantly the systematic errors for the deprojected masses are increased because a change in the lensing signal also affects the radius corresponding to a particular overdensity. This increases the uncertainty compared to a fixed aperture because the enclosed mass increases with radius. As a result we estimate a total systematic uncertainty in $M_{500}^{\rm ap}$ of 6 per cent for clusters at $z = 0.2$, which

increases to 12 per cent at $z = 0.55$. Given the observational cost of calibrating scaling relations, it makes sense to avoid introducing such unnecessary sources of uncertainty: numerical simulations should instead be used to make predictions for the observed lensing measurements.

In Section 4.1.2, we determined masses for Abell 1758 and the six clusters in common with Umetsu et al. (2014) using available Subaru observations. We found that these masses were 4 ± 6 per cent higher than our estimates based on the CFHT data. Given the differences between the data from the two telescopes (e.g. depth, masked areas), the results agree very well. Moreover, when we match the object catalogues, we find that we recover the average tangential distortion within 1 ± 2 per cent, suggesting our shape measurement pipeline is robust.

5 *Planck* SZE SCALING RELATION

Planck Collaboration XX (2014c) present constraints on cosmological parameters using the cluster number counts as a function of redshift for a sample of 189 clusters of galaxies detected through the SZE by *Planck*. The estimates for the masses from *Planck* are based on the scaling relation between the X-ray hydrostatic mass and Y_X, the product of the X-ray temperature and the gas mass. This relation is calibrated using measurements from Arnaud et al. (2010), who studied a sample of 20 nearby relaxed clusters. Consequently a measurement of Y_X can be converted into a hydrostatic mass estimate $M_{500}^{Y_X}$. These results are then used to establish the relation between the SZ signal Y_{500} and $M_{500}^{Y_X}$ (see appendix A.2.2 of Planck Collaboration XX 2014c, for details). We denote this hydrodynamic mass estimate as $M_{500}^{\rm Planck}$.

Numerical simulations (e.g. Rasia et al. 2006; Nagai et al. 2007; Lau et al. 2009) suggest that such mass estimates are biased low. Similarly, Mahdavi et al. (2013) studied the scaling relations between X-ray observations and the weak lensing masses from H12 and found that hydrostatic masses underestimate the weak lensing masses by 10–15 per cent at r_{500}. Our updated masses do not change this conclusion, and in fact strengthen it. In their analysis Planck Collaboration XX (2014c) assume that the hydrostatic masses are biased low by a factor $(1 - b) = 0.8$ based on a comparison with numerical simulations. They find that their best-fitting parameters for σ_8 and $\Omega_{\rm m}$ are in tension with the measurements obtained from the analysis of the primary CMB by Planck Collaboration XVI (2014b). The results can be reconciled by considering a low value of $(1 - b) \sim 0.6$.

Recently, von der Linden et al. (2014b) estimated the bias using the lensing masses for the 38 clusters in common between *Planck* and WtG. They compared their estimates for M_{500} based on the NFW fits with $c_{200} = 4$ from Applegate et al. (2014) to the hydrostatic mass estimates from Planck Collaboration XXIX (2014a). They obtained an average ratio $(1 - b) = 0.69 \pm 0.07$, which alleviates the tension. As our comparison in Section 4.1.1 and Fig. 15 shows, the WtG masses are slightly higher than our estimates when we follow the same approach, but when we compare the masses from WtG to our deprojected aperture masses, which are more robust and therefore used here, we find that the agreement is excellent.

There are 38 clusters in common between CCCP and the catalogue provided by Planck Collaboration XXIX (2014a), although we omit Abell 115 from the comparison as we determine masses for the two separate components of this merging cluster. The left-hand panel in Fig. 21 shows the deprojected aperture mass $M_{500}^{\rm WL}$ as a function of the hydrostatic mass $M_{500}^{\rm Planck}$ from Planck Collaboration XXIX (2014a). Note that the observed value for Y_X was used to estimate the radius r_{500} used to determine $M_{500}^{\rm Planck}$, whereas $M_{500}^{\rm WL}$ is based on the value for r_{500} listed in Table 2. For the cosmological analysis, Planck Collaboration XX (2014c) restricted the sample to clusters above a SNR threshold of 7 in unmasked areas. In our case, the mask only impacts the merging cluster Abell 2163, which corresponds to the rightmost point in Fig. 21. There are 20 SNR>7 clusters in common with CCCP and these are indicated as filled points in Fig. 21, whereas the remaining clusters are indicated by

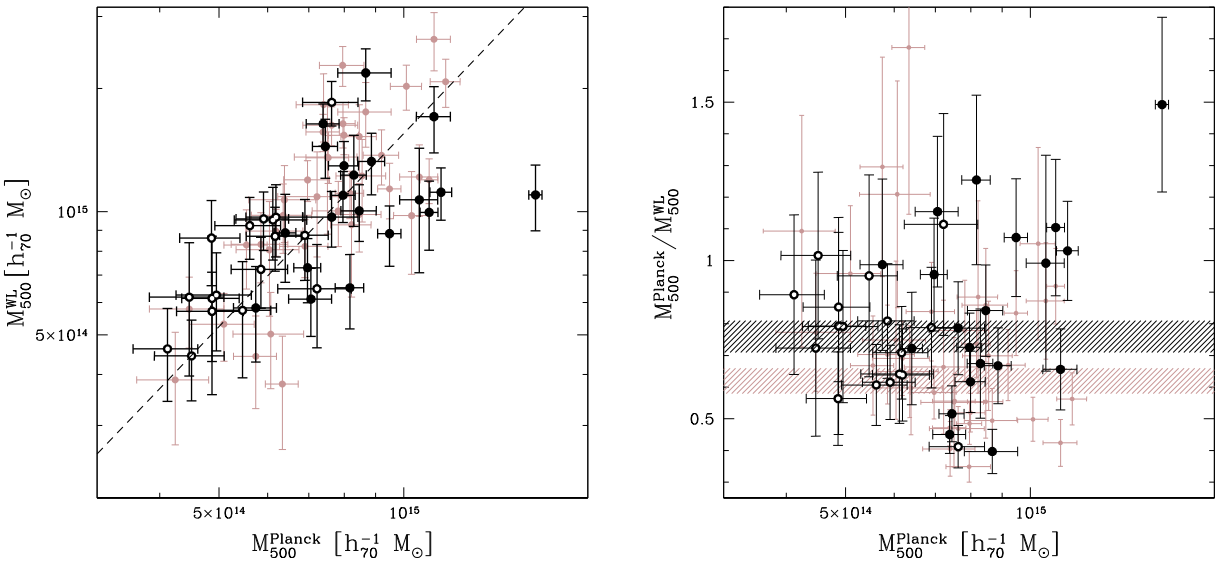

Figure 21. Left-hand panel: the deprojected aperture mass M_{500} from weak lensing as a function of the hydrostatic mass from Planck Collaboration XXIX (2014a). Note that $M_{500}^{\rm Planck}$ is measured using r_{500} from the estimate of Y_X, and $M_{500}^{\rm WL}$ is determined using the lensing derived value for r_{500}. The black points show our CCCP measurements, with the filled symbols indicating the clusters detected by *Planck* with a signal-to-noise ratio $SNR > 7$ and the open points the remainder of the sample. The dashed line shows the best-fitting power-law model. The WtG results are shown as rosy brown coloured points. Right-hand panel: ratio of the hydrostatic and the weak lensing mass as a function of mass. The dark hatched area indicates the average value of 0.76 ± 0.05 for the CCCP sample, whereas the rosy brown coloured hatched region is the average for the published WtG measurements, for which we find 0.62 ± 0.04.

the open points. We find that the SNR threshold is essentially a selection by mass. For reference, the measurements from von der Linden et al. (2014b) are indicated by the rosy brown coloured points.

The right-hand panel shows the ratio of the hydrostatic masses from *Planck* and our weak lensing estimates for all 37 clusters in common. The hatched region indicates our estimate for $(1 - b) = 0.76 \pm 0.05(\mathrm{stat}) \pm 0.06(\mathrm{syst})$, which was obtained from a linear fit to $M_{500}^{\mathrm{Planck}}$ as a function of M_{500}^{WL} that accounts for intrinsic scatter (Hogg, Bovy & Lang 2010). The systematic error is based on the estimates presented in Section 4.3. We measure an intrinsic scatter of (28 ± 6) per cent, most of which can be attributed to the triaxial nature of dark matter haloes (e.g. Corless & King 2007; Meneghetti et al. 2010). If we restrict the comparison to the clusters with SNR > 7 (black points) we obtain $(1 - b) = 0.78 \pm 0.07$, whereas $(1 - b) = 0.69 \pm 0.05$ for the remaining clusters. For reference, the rosy brown coloured points and hatched region indicate the results for WtG, used in von der Linden et al. (2014b). We refit these measurements, which yields $(1 - b) = 0.62 \pm 0.04$ and an intrinsic scatter of (26 ± 5) per cent. Our measurement of the bias is in agreement with the nominal value adopted by Planck Collaboration XX (2014c) and we conclude that a large bias in the hydrostatic mass estimate is unlikely to be the explanation of the tension of the cluster counts and the primary CMB.

von der Linden et al. (2014b) find modest evidence for a mass dependence of the bias, with $M_{\mathrm{Planck}} \propto M_{\mathrm{WtG}}^{0.68}$. It is therefore interesting to repeat this for our measurements. If we restrict the fit to the clusters with a SNR>7, the range is too small to obtain a useful constraint on the slope. We therefore fit a power law to the CCCP measurements of the 37 clusters that overlap with Planck Collaboration XXIX (2014a), which yields

$$\frac{M_{\mathrm{Planck}}}{10^{15}\, h_{70}^{-1}\, \mathrm{M}_\odot} = (0.76 \pm 0.04) \times \left(\frac{M_{\mathrm{CCCP}}}{10^{15}\, h_{70}^{-1}\, \mathrm{M}_\odot} \right)^{0.64 \pm 0.17},$$

and an intrinsic scatter of (21 ± 4) per cent. The slope is similar to that found by von der Linden et al. (2014b) and our results therefore support their conclusion that the bias in the hydrostatic masses used by *Planck* depends on the cluster mass, but our normalization is 9 per cent higher.

As noted above, Planck Collaboration XX (2014c) use X-ray data to relate the observed SZ-signal to cluster mass. It is, however, more convenient to directly constrain the scaling relation between the lensing mass and the observed SZ signal. H12 presented results for the early data release from Planck Collaboration VIII (2011), but here we expand the sample to the 37 clusters in common with CCCP and use the measurements for Y_{500} provided by Planck Collaboration XXIX (2014a).

Assuming a constant gas fraction and self-similarity, the SZ signal Y scales with mass as $M_{500}^{5/3} \propto D_{\mathrm{ang}}^2 E(z)^{-2/3} Y_{500}$ (e.g. McCarthy et al. 2003), where D_{ang} is the angular diameter distance to the cluster and $E(z) = H(z)/H_0$ is the normalized Hubble parameter. The results are presented in Fig. 22, where the open points indicate the clusters that *Planck* detected with an SNR<7. The dashed line indicates the best-fitting power-law relation to all clusters in common, for which we find

$$\frac{M_{500}^{\mathrm{WL}}}{10^{15}\, h_{70}^{-1}\, \mathrm{M}_\odot} = (1.01 \pm 0.06) \times \left(\frac{10^4\, D_{\mathrm{ang}}^2\, Y_{500}}{E(z)^{2/3}\, \mathrm{Mpc}^2} \right)^{0.53 \pm 0.13},$$

and an intrinsic scatter of (27 ± 6) per cent. For this comparison we used measurements of the SZ signal and the lensing mass in apertures that were determined independently. Although this is what one needs for the cosmological interpretation of the *Planck* cluster sample, the use of different apertures introduces additional noise as well as an offset because the lensing aperture is larger on average.

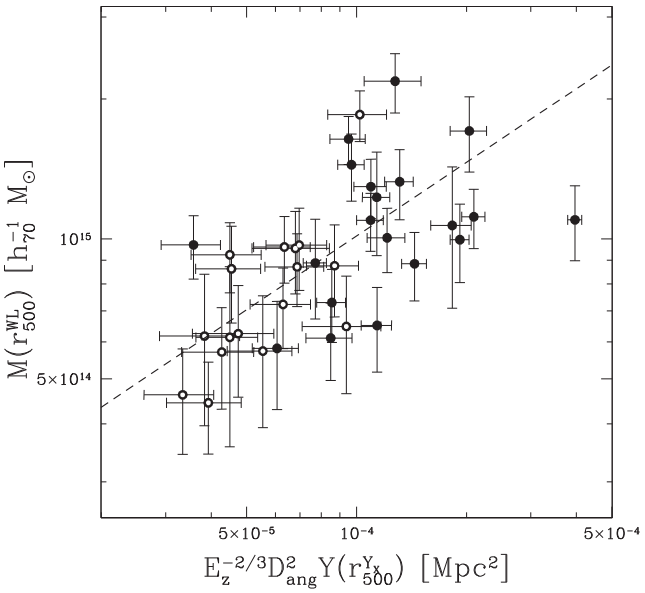

Figure 22. Plot of M_{500}, the aperture mass estimate within an aperture r_{500}^{WL}, as a function of the SZ signal measured within an aperture $r_{500}^{Y_{\mathrm{X}}}$ from Planck Collaboration XXIX (2014a). The dashed line indicates the best-fitting power law, which has a slope of 0.53 ± 0.13. We measure an intrinsic scatter of (27 ± 6) per cent. The open circles indicate clusters which *Planck* detected with an SNR < 7.

In the case of hydrodynamic simulations of clusters the comparison can be made at a common radius, as r_{500} is known. It is therefore useful to consider also the scaling relation for the SZ signal within r_{500}^{WL}. The SZ measurements within r_{500}^{WL} were estimated (Arnaud & Pratt, private communication) using the two-dimensional marginal probability distribution between the SZ signal and size available from the 2013 Planck SZ catalogue (Planck Collaboration XXIX 2014a). They correlate very well with Y_{500}. The best-fitting power-law scaling relation is given by

$$\frac{M_{500}^{\mathrm{WL}}}{10^{15}\, h_{70}^{-1}\, \mathrm{M}_\odot} = (0.98 \pm 0.05) \times \left(\frac{10^4\, D_{\mathrm{ang}}^2\, Y(r_{500}^{\mathrm{WL}})}{E(z)^{2/3}\, \mathrm{Mpc}^2} \right)^{0.64 \pm 0.12},$$

and the intrinsic scatter is reduced to (22 ± 7) per cent. Note that for both scaling relations, the slopes are consistent with the value of 0.6 expected for a self-similar model.

As a caveat, we note that the combination of relatively low significance of the SZ detections and large intrinsic scatter leads to Malmquist bias. As a result, the average SZ signal of the observed sample is biased high compared to the average of the parent population the clusters are drawn from. Consequently, the best-fitting parameters for the scaling relations are expected to be slightly biased. If the CCCP sample were a well-defined sample one could account for this, which was done by Planck Collaboration XX (2014c). Although Mahdavi et al. (2013) show that the CCCP sample is representative of more carefully selected samples of X-ray luminous clusters, the selection may impact the scaling relation. We therefore do not attempt to correct our results, but note that, based on the findings of Planck Collaboration XX (2014c) and Mahdavi et al. (2013), we expect the bias to be small.

6 CONCLUSIONS

Accurate cluster masses are necessary to interpret the cluster counts from wide-area surveys. In particular the scaling relations and their scatter need to be determined observationally. Weak lensing masses are particularly well-suited as they provide a direct estimate of the projected mass and do not depend on the dynamical state of the cluster. In this paper we revisited the analysis of a sample of 50 massive clusters by H12, with a particular focus on improving the corrections for various sources of systematic error in the cluster masses.

We use extensive image simulations to quantify the bias in our shape measurement algorithm. The bias is a strong function of SNR and size, but depends relatively weakly on surface brightness profile and ellipticity distribution. We demonstrate that the inferred bias depends on the input parameters used to create the simulated data. For instance we find that the bias is underestimated if faint galaxies are lacking from the simulations: the bias converges if the simulation includes galaxies that are at least 1.5 mag fainter than the limiting magnitude of the sample of sources used for the lensing analysis. The large number of simulated galaxies enables us to determine an empirical correction, which is found to be accurate and robust to the main uncertainties. We estimate that the systematic uncertainty due to the shape measurements is at most 2 per cent.

The dominant source of systematic error is the source redshift distribution, which is needed to convert the lensing signal into an estimate of the mass. We use the latest state-of-the-art photometric redshift catalogues that are based on measurements in 29 bands in the COSMOS field (Capak et al. 2007; Scoville et al. 2007), including new deep NIR observations from UltraVISTA (McCracken et al. 2012). Compared to our previous analysis, this leads to a small increase in mean source redshift, or a modest (∼4 per cent) reduction in the cluster mass. Despite the unprecedented quality of the data, different analyses show variations in the source redshift distributions that result in systematic uncertainties that are substantial compared to the statistical uncertainties for the full sample of clusters.

We find that the projected aperture masses within apertures of fixed radius provide the most robust measurements. The dependence on the assumed density profile is minimal and the systematic error is dominated by the uncertainty in the source redshift distribution. We estimate a total systematic error of 4.2 per cent for clusters at $z = 0.2$, which increases to 8.5 per cent at $z = 0.55$. To compare with measurements at other wavelengths, we deproject the masses. This leads to an increased sensitivity to the assumed density profile, although this is modest in the case of M_{500}: we estimate a systematic uncertainty of 2 per cent. Compared to the masses within a fixed aperture, the additional uncertainty in determining r_{500} increases both the statistical and systematic errors. Comparison of the deprojected aperture masses within r_{500} and the corresponding mass for the best-fitting NFW profile shows good agreement, even though the two estimates are nearly independent from each other.

We compare masses for the 18 clusters in common with Applegate et al. (2014). To do so we fit an NFW model with $c = 4$ to our lensing signal within $0.75{-}3\,h_{70}^{-1}$ Mpc. We find that the resulting CCCP masses are on average 8 ± 4 per cent lower, which reduces to 6 per cent if we use the same source redshift distribution as Applegate et al. (2014). Given the limitations of the comparison, it is not clear whether this difference is a sign of residual systematics or merely statistical in nature.

We did examine the robustness of our analysis. To this end we determined masses for Abell 1758 and the six clusters in common with Umetsu et al. (2014) using available Subaru observations. We found good agreement with our mass estimates based on the CFHT data, suggesting our pipeline yields robust results when data from different telescopes are analysed. Interestingly, a direct comparison of our CFHT measurements to catalogues provided by the WtG team (Von der Linden, private communication) also showed remarkable agreement with $\langle g_{\rm T}^{\rm CCCP}/g_{\rm T}^{\rm WtG}\rangle = 0.991 \pm 0.018$.

The overlap of 37 clusters with the sample of clusters for which *Planck* detected the SZ signal (Planck Collaboration XXIX 2014a) enables us to calibrate the bias in the hydrostatic masses used by Planck Collaboration XX (2014c) to infer cosmological parameters. The resulting estimates for σ_8 and $\Omega_{\rm m}$ are in tension with the measurements from the primary CMB Planck Collaboration XVI (2014b). Our measurements for the overlapping clusters yield $1 - b = 0.76 \pm 0.05({\rm stat}) \pm 0.06({\rm syst})$, in good agreement with the nominal value used by Planck Collaboration XX (2014c). Our results do not support a large bias in the hydrostatic masses, which could alleviate the tension. We also directly calibrate the scaling relation between the SZ signal Y_{500} and the weak lensing mass. When we compare the lensing mass to the SZ signal measured in the same aperture, we find a best-fitting slope of 0.64 ± 0.12 which is in good agreement with the expectation of a self-similar model (e.g. McCarthy et al. 2003).

The constraints from the current CCCP sample are already limited by systematic uncertainties. The most dominant of these is our limited knowledge of the source redshift distribution. Although this can be alleviated by measuring photometric redshifts for the sources in the cluster fields (e.g. Applegate et al. 2014; Umetsu et al. 2014), biases may remain due to limited wavelength coverage. Improving this situation is critical to calibrate cluster scaling relations to the level of accuracy afforded by the next generation of cluster surveys.

ACKNOWLEDGEMENTS

We thank Monique Arnaud, Gabriel Pratt, Keiichi Umetsu, James Jee, Edo van Uitert, Anja von der Linden, Doug Applegate, Ken Rines and Douglas Scott for useful discussions. We are also grateful to Keiichi Umetsu and James Jee for providing some of their Subaru images, and to Anja von der Linden for sharing the WtG lensing catalogues. HH and MC were supported by the Netherlands organization for Scientific Research (NWO) Vidi grant 639.042.814; HH, RH and MV acknowledge support from the European Research Council FP7 grant number 279396. AB would like to acknowledge research funding from NSERC Canada through its Discovery Grant program. AM acknowledges support from NASA ADAP grant NNX12AE45G. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency.

REFERENCES

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Applegate D. E. et al., 2014, MNRAS, 439, 48
Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92
Bahcall N. A., Fan X., 1998, ApJ, 504, 1
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25
Bernstein G. M., 2010, MNRAS, 406, 2793
Bernstein G. M., Armstrong R., 2014, MNRAS, 438, 1880
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bhattacharya S., Habib S., Heitmann K., Vikhlinin A., 2013, ApJ, 766, 32

Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Bridle S. et al., 2010, MNRAS, 405, 2044
Cacciato M., van den Bosch F. C., More S., Li R., Mo H. J., Yang X., 2009, MNRAS, 394, 929
Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, MNRAS, 430, 767
Cacciato M., van Uitert E., Hoekstra H., 2014, MNRAS, 437, 377
Capak P. et al., 2007, ApJS, 172, 99
Clowe D., Luppino G. A., Kaiser N., Henry J. P., Gioia I. M., 1998, ApJ, 497, L61
Coe D. et al., 2012, ApJ, 757, 22
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Corless V. L., King L. J., 2007, MNRAS, 380, 149
Diaferio A., 1999, MNRAS, 309, 610
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., Battye R. A., Booth C. M., 2010, MNRAS, 405, 2161
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Fernández-Soto A., Lanzetta K. M., Yahil A., 1999, ApJ, 513, 34
Gruen D. et al., 2014, MNRAS, 442, 1507
Hasselfield M. et al., 2013, J. Cosmol. Astropart. Phys., 7, 8
Henry J. P., 2000, ApJ, 534, 565
Henry J. P., Evrard A. E., Hoekstra H., Babul A., Mahdavi A., 2009, ApJ, 691, 1307
Heymans C. et al., 2006, MNRAS, 368, 1323
Hildebrandt H. et al., 2010, A&A, 523, A31
Hoekstra H., 2001, A&A, 370, 743
Hoekstra H., 2003, MNRAS, 339, 1155
Hoekstra H., 2004, MNRAS, 347, 1337
Hoekstra H., 2007, MNRAS, 379, 317
Hoekstra H., Franx M., Kuijken K., Squires G., 1998, ApJ, 504, 636
Hoekstra H., Franx M., Kuijken K., 2000, ApJ, 532, 88
Hoekstra H., Hartlap J., Hilbert S., van Uitert E., 2011, MNRAS, 412, 2095
Hoekstra H., Mahdavi A., Babul A., Bildfell C., 2012, MNRAS, 427, 1298 (H12)
Hoekstra H., Bartelmann M., Dahle H., Israel H., Limousin M., Meneghetti M., 2013, Space Sci. Rev., 177, 75
Hogg D. W., Pahre M. A., McCarthy J. K., Cohen J. G., Blandford R., Smail I., Soifer B. T., 1997, MNRAS, 288, 404
Hogg D. W., Bovy J., Lang D., 2010, preprint (arXiv:1008.4686)
Ilbert O. et al., 2006, A&A, 457, 841
Ilbert O. et al., 2009, ApJ, 690, 1236
Ilbert O. et al., 2013, A&A, 556, A55
Jee M. J., Mahdavi A., Hoekstra H., Babul A., Dalcanton J. J., Carroll P., Capak P., 2012, ApJ, 747, 96
Jee M. J., Hoekstra H., Mahdavi A., Babul A., 2014, ApJ, 783, 78
Kacprzak T., Zuntz J., Rowe B., Bridle S., Refregier A., Amara A., Voigt L., Hirsch M., 2012, MNRAS, 427, 2711
Kacprzak T., Bridle S., Rowe B., Voigt L., Zuntz J., Hirsch M., MacCrann N., 2014, MNRAS, 441, 2528
Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460
Kelly P. L. et al., 2014, MNRAS, 439, 28
Kitching T. D. et al., 2012, MNRAS, 423, 3163
Komatsu E. et al., 2009, ApJS, 180, 330
Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129
Lilly S. J. et al., 2009, ApJS, 184, 218
Luppino G. A., Kaiser N., 1997, ApJ, 475, 20
McCarthy I. G., Babul A., Holder G. P., Balogh M. L., 2003, ApJ, 591, 515
McCracken H. J. et al., 2012, A&A, 544, A156
Mahdavi A., Hoekstra H., Babul A., Balam D. D., Capak P. L., 2007, ApJ, 668, 806
Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, MNRAS, 384, 1567
Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116
Mandelbaum R. et al., 2014a, preprint (arXiv:1412.1825)
Mandelbaum R. et al., 2014b, ApJS, 212, 5
Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, MNRAS, 406, 1759
Mantz A. B. et al., 2015, MNRAS, 446, 2205
Martin D. C. et al., 2005, ApJ, 619, L1
Massey R., Refregier A., Conselice C. J., David J., Bacon J., 2004, MNRAS, 348, 214
Massey R. et al., 2007, MNRAS, 376, 13
Massey R. et al., 2013, MNRAS, 429, 661
Melchior P., Viola M., 2012, MNRAS, 424, 2757
Mellier Y., 1999, ARA&A, 37, 127
Meneghetti M., Rasia E., Merten J., Bellagamba F., Ettori S., Mazzotta P., Dolag K., Marri S., 2010, A&A, 514, A93
Miller L. et al., 2013, MNRAS, 429, 2858
Muzzin A. et al., 2013, ApJS, 206, 8
Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Oguri M., Hamana T., 2011, MNRAS, 414, 1851
Okabe N., Takada M., Umetsu K., Futamase T., Smith G. P., 2010, PASJ, 62, 811
Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, ApJ, 769, L35
Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266
Planck Collaboration VIII, 2011, A&A, 536, A8
Planck Collaboration XVI, 2014b, A&A, 571, A16
Planck Collaboration XX, 2014c, A&A, 571, A20
Planck Collaboration XXIX, 2014a, A&A, 571, A29
Ragozzine B., Clowe D., Markevitch M., Gonzalez A. H., Bradač M., 2012, ApJ, 744, 94
Rasia E. et al., 2006, MNRAS, 369, 2013
Reichardt C. L. et al., 2013, ApJ, 763, 127
Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716
Rines K., Geller M. J., Diaferio A., Kurtz M. J., 2013, ApJ, 767, 15
Rix H.-W. et al., 2004, ApJS, 152, 163
Robin A. C., Reylé C., Derrière S., Picaud S., 2003, A&A, 409, 523
Rowe B. et al., 2014, preprint (arXiv:1407.7676)
Sanders D. B. et al., 2007, ApJS, 172, 86
Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Scoville N. et al., 2007, ApJS, 172, 1
Seljak U., Warren M. S., 2004, MNRAS, 355, 129
Semboloni E. et al., 2013, MNRAS, 432, 2385
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
Sifón C., Hoekstra H., Cacciato M., Viola M., Köhlinger F., van der Burg R., Sand D., Graham M. L., 2015, A&A, 575, 48
Simet M., Mandelbaum R., 2014, preprint (arXiv:1406.4908)
Sunyaev R. A., Zeldovich Y. B., 1972, Comments Astrophys. Space Phys., 4, 173
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Umetsu K. et al., 2014, ApJ, 795, 163
van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725
van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14
Vikhlinin A. et al., 2009, ApJ, 692, 1060
Viola M., Kitching T. D., Joachimi B., 2014, MNRAS, 439, 1909
von der Linden A. et al., 2014a, MNRAS, 439, 2
von der Linden A. et al., 2014b, MNRAS, 443, 1973

APPENDIX A: EFFECTS OF PSF ANISOTROPY

To quantify the multiplicative bias in the shear measurement resulting from noise in the images and limitations in the correction for the size of the PSF and the weight function, we created simulated images with a circular PSF. In real data, however, the PSF is generally anisotropic, albeit to varying degree. PSF anisotropy leads to additive biases by introducing coherent alignments in the observed shapes. In the case of cosmic shear this is a dominant source of

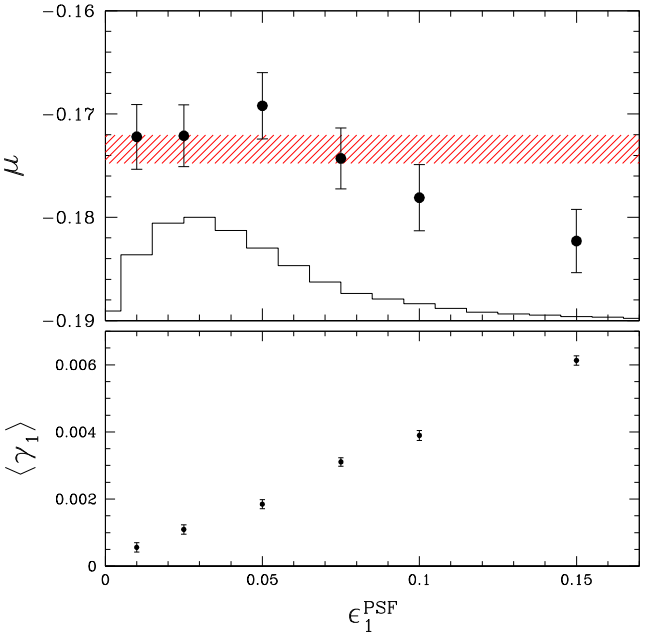

Figure A1. Top panel: multiplicative bias as a function of the ellipticity of the PSF for galaxies with $22 < m_r < 25$. The red hatched region indicates the results obtained for a circular PSF. The histogram indicates the distribution of PSF ellipticities observed in the CFHT data used in this paper. Bottom panel: the average shear γ_1 measured after PSF correction as a function of PSF ellipticity. The average should vanish if the correction is perfect. These results suggest that the residual is ∼4 per cent of the original PSF ellipticity.

systematic, but in our case the signal is averaged azimuthally around the clusters, which also averages out most of the PSF anisotropy.[10]

None the less it is important to examine whether the results obtained using circular PSFs can be applied to our CFHT data. We therefore created a set of simulations in which the PSF is elliptical (along the x-axis) using our fiducial ellipticity distribution ($\epsilon_0 = 0.25$). These images were analysed as usual, and the resulting multiplicative bias as a function of PSF ellipticity is presented in the top panel of Fig. A1. Even for rather elongated PSFs ($\epsilon > 0.1$), the increase in bias is at the per cent level at most. For comparison, the histogram shows the observed distribution of PSF ellipticity in the CCCP data. These results suggest that we do not need to account for the PSF anisotropy explicitly in order to quantify the multiplicative bias in our data.

We do not expect the correction to work perfectly, either due to limitation of the correction method itself, or due to the fact that the noise in the images biases the polarizabilities. The bottom panel shows how well the correction for PSF anisotropy performs. We observe a linear trend of the recovered average shear as a function of PSF ellipticity. For the galaxies with $22 < m_r < 25$, this results in a residual bias of ∼4 per cent of the original PSF ellipticity. Although the bias is smaller for bright galaxies it does not vanish, suggesting that a large part of the observed bias is due to a fundamental limitation of the KSB method.[11]

[10] We note that this is not always the case, as some images show strong radial patterns, which can bias the azimuthally averaged lensing signal if the cluster is located at the centre of the field of view.

[11] The KSB algorithm assumes that the PSF is described as the convolution of a circular kernel with a compact anisotropic one.

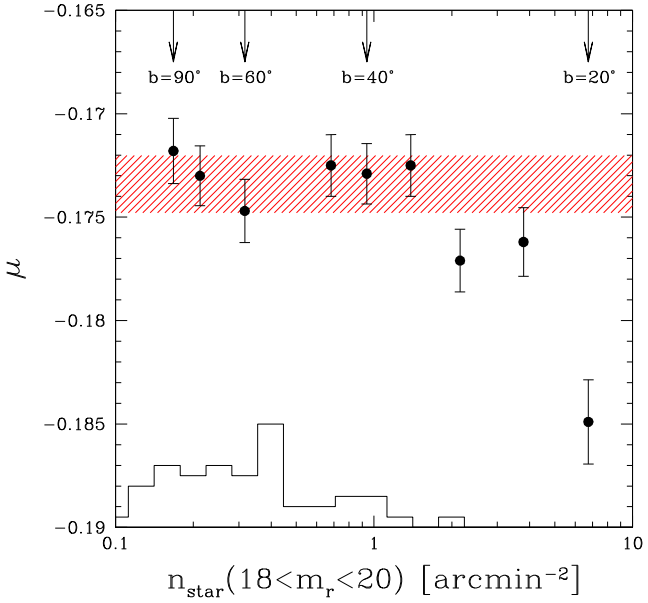

Figure B1. Multiplicative bias for galaxies with $22 < m_r < 25$ as a function of the density of stars with $18 < m_r < 20$, based on the Besançon model of stellar population synthesis of the Galaxy (Robin et al. 2003). The red hatched region indicates the results for the simulations without stars. The arrows indicate the corresponding Galactic latitudes (for the longitude of Abell 1835, $l = 340°$). The histogram shows the distribution of star densities of the CCCP data.

APPENDIX B: CONTAMINATION BY STARS

The simulated data that were used to study the multiplicative bias contained a small number of bright stars, which were only included to keep track of the PSF properties. In real data, however, stars may blend with galaxies or other stars. As a consequence they might be misidentified and be included in the galaxy catalogue, contributing to multiplicative bias (as they are not sheared).

In this section, we study how the multiplicative bias depends on star density by including stars in the images with a realistic number density and magnitude distribution. We use the Besançon model of stellar population synthesis of the Galaxy for the stars (Robin et al. 2003). As a reference we consider again Abell 1835, with $(l, b) = (340, 60)$, where we change the Galactic latitude b to change the star number density. We place the stars at random locations in the images, with $\epsilon_0 = 0.25$, and proceed as before with the shape analysis.

The results for galaxies with $22 < m_r < 25$ are presented in Fig. B1, which shows the multiplicative bias as a function of the density of stars with $18 < m_r < 20$. For densities $n_{\rm star} < 1.5$ arcmin^{-2} the bias is consistent with the results without stars (indicated by the red hatched region). For the adopted longitude of $l = 340°$ this corresponds to $b > 35°$. For lower Galactic latitudes, the higher number density of stars can lead to appreciable levels of bias. The histogram in Fig. B1 shows the distribution of star density in the CCCP data, which suggests that we can safely ignore the contribution from stars.

APPENDIX C: DETERMINING THE EMPIRICAL CORRECTION

In this appendix, we describe the empirical correction used in the analysis of the CCCP data. We assume that the bias is a function of the SNR (e.g. Kacprzak et al. 2012; Melchior & Viola 2012) and depends on how well a galaxy is resolved (e.g. Massey et al.

2013; Miller et al. 2013). That these parameters are important is also suggested by Figs 5 and 6, which show a larger bias for fainter galaxies and larger PSFs.

Our implementation of KSB provides an estimate for σ_e, the uncertainty in the polarization (see Hoekstra et al. 2000, for details). The reciprocal of this quantity is a useful proxy for the SNR ν, hence we use $\nu = 1/\sigma_e$. To quantify how well a galaxy is resolved we use

$$\mathcal{R}^2 = \frac{r^2_{\rm h,*}}{r^2_{\rm h,gal} - r^2_{\rm h,*}}, \tag{C1}$$

where $r_{\rm h,*}$ denotes the half-light radius of the PSF and $r_{\rm h,\,gal}$ that of the observed galaxy. The denominator corresponds to the unconvolved size of a source if galaxies were Gaussians. Importantly, these are quantities that can be measured for individual sources.

We consider values of ϵ_0 between 0 and 0.5 with steps of 0.05 and create 169 pairs of images, each 10 000 by 10 000 pixels, with constant shears of $-0.06 \leq \gamma_i \leq 0.06$ applied. As a result, for each ϵ_0 we analyse $\sim 10^7$ galaxies with $20 < m_r < 25$ (note that the input catalogue does contain fainter galaxies). For a given ϵ_0 we bin the measurements in fine bins of ν and $r_{\rm h}$ and determine the bias.

We considered various fitting functions, with the aim to find an adequate correction that is also robust against the uncertainty in the true value of ϵ_0. We consider a conservative range of $\epsilon_0 \in [0.15, 0.3]$, where we note that a stronger prior on the input ellipticity distribution would allow the uncertainties to be reduced, and the fitting functions to be optimized.

Although further optimization is possible by introducing additional parameters, we opted for a correction with four free parameters: one to describe the size dependence of the bias for a given ν, and three to capture the dependence on ν. The correction takes the form

$$1 + \mu(\nu, \mathcal{R}) = \frac{b(\nu)}{1 + \alpha(\epsilon_0)\mathcal{R}}. \tag{C2}$$

The first step is to determine $\alpha(\epsilon_0)$, which is done by examining μ as a function of size for narrow bins in ν. The left-hand panel of Fig. C1 shows the measurements as a function of half-light radius for galaxies with $20 < \nu < 30$ for three ellipticity distributions. The number of objects as a function of observed half-light radii is shown in the bottom panel; most sources are small (the value for the PSF is $r_{\rm h,*} = 2.056$ pixels).

Table C1. Parameters for the empirical correction for multiplicative bias.

ϵ_0	α	b_0	b_1	b_2
0	0.056 ± 0.0015	0.952 ± 0.0143	0.84 ± 0.15	-5.50 ± 0.32
0.05	0.067 ± 0.0017	0.957 ± 0.0145	0.86 ± 0.15	-5.47 ± 0.33
0.10	0.107 ± 0.0022	0.958 ± 0.0164	1.20 ± 0.17	-6.15 ± 0.37
0.15	0.164 ± 0.0029	0.959 ± 0.0233	1.61 ± 0.24	-6.81 ± 0.53
0.20	0.247 ± 0.0038	0.970 ± 0.0235	2.01 ± 0.24	-7.34 ± 0.53
0.25	0.348 ± 0.0051	0.988 ± 0.0217	2.29 ± 0.22	-7.40 ± 0.49
0.30	0.473 ± 0.0067	0.993 ± 0.0220	2.73 ± 0.22	-7.72 ± 0.50
0.35	0.616 ± 0.0087	1.017 ± 0.0298	3.04 ± 0.30	-7.70 ± 0.68
0.40	0.729 ± 0.0104	0.997 ± 0.0260	3.57 ± 0.26	-8.32 ± 0.59
0.45	0.864 ± 0.0120	1.018 ± 0.0289	3.85 ± 0.29	-8.43 ± 0.65
0.50	0.921 ± 0.0121	1.002 ± 0.0265	3.95 ± 0.27	-8.18 ± 0.60

For each ellipticity distribution we determine the best-fitting value for α, under the assumption it only depends on ϵ_0. The results are presented in the right-hand panel of Fig. C1 and listed in Table C1. The right-hand panel in Fig. C1 shows that $\alpha(\epsilon_0)$ increases smoothly with increasing ϵ_0. The lines in the left-hand panel indicate the predicted bias, with the amplitude a free parameter (which is used to determine $b(\nu)$, see below). Our parametrization for the size dependence does fairly well for the bulk of the sources, but it does not capture the results for $r_{\rm h} > 4$ pixels. Better agreement is obtained if we include an additional term $\propto \mathcal{R}^2$, but we found that this did not improve the robustness of the correction. Similarly, we found that we could have included a dependence on $1/\nu$, again with limited effect. Note there is covariance between some of the parameters. For instance, for small values for ν we do expect $b(\nu)$ and $\alpha(\epsilon_0)$ to be correlated because those galaxies have on average larger values for $\mathcal{R}$.

Closer inspection of the objects with large observed sizes revealed that most of these are faint objects for which the input sizes were

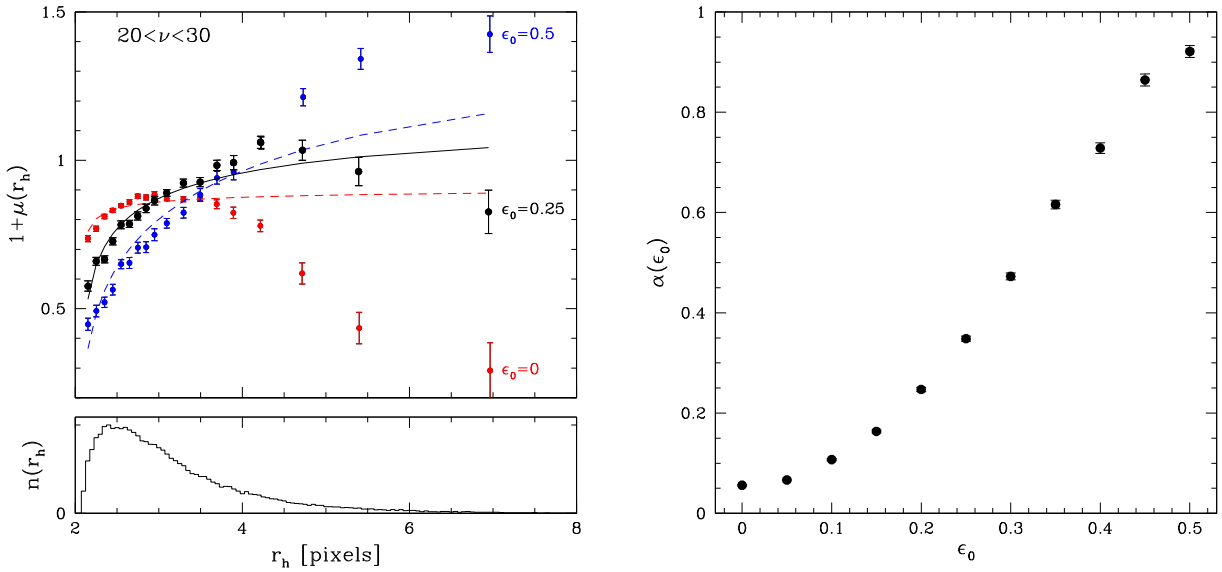

Figure C1. Left-hand panel: bias $1 + \mu$ as a function of half-light radius for galaxies with $20 < \nu < 30$. The black points indicate the results for our adopted value of $\epsilon_0 = 0.25$; the red (blue) points are for the extreme values of $\epsilon_0 = 0$ ($\epsilon_0 = 0.5$). The lines indicate the fits using our best-fitting value for $\alpha(\epsilon_0)$. The bottom panel shows the distribution of half-light radii for these sources. Right-hand panel: resulting best-fitting value for $\alpha(\epsilon_0)$ for our simulated CCCP data. The values are listed in Table C1.

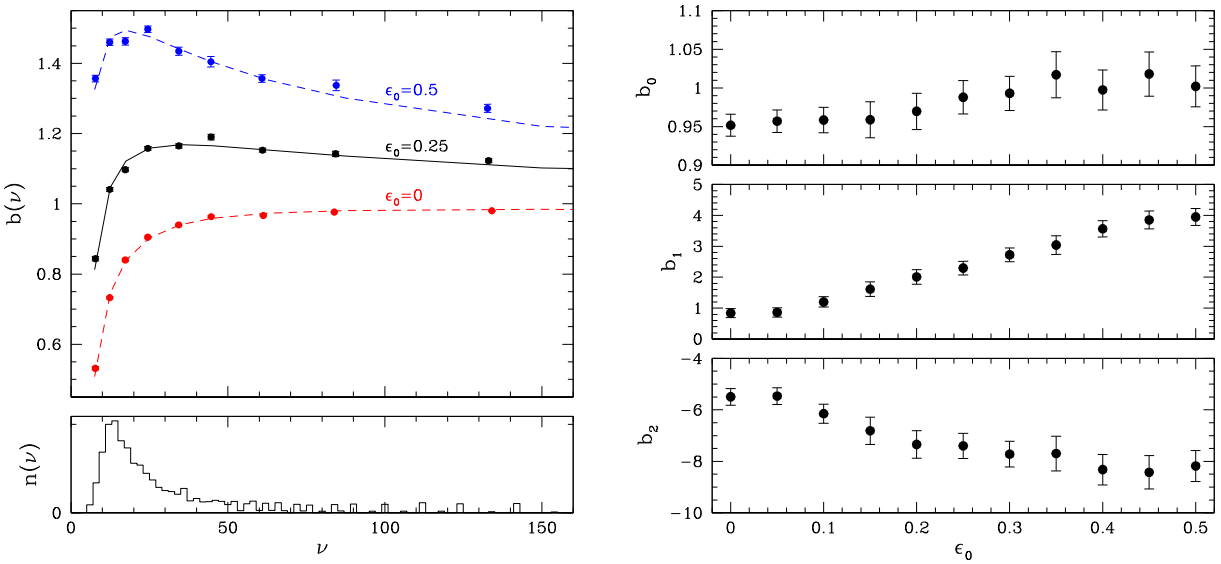

Figure C2. Left-hand panel: $b(\nu)$ for the three ellipticity distributions, as well as the best-fitting models. The bottom panel shows the distribution of SNRs for galaxies with $20 < m_r < 25$. Right-hand panel: the model parameters b_i as a function of ϵ_0, which show a smooth dependence. The values are listed in Table C1. Note that b_0 is fairly close to unity.

much smaller or blends with other galaxies. As the latter might be particularly relevant for the study of galaxy clusters, we decided to only use galaxies with $r_h < 5$ pixels for the actual lensing analysis. The ensemble averaged lensing signal as a function of source size presented in Fig. 14 indicates that the results for large sources are indeed biased low. Note that we do not apply this cut when fitting for $\alpha(\epsilon_0)$, because we found that the correction performed a bit better when we considered the full range in sizes.

The next step is to quantify how μ depends on the SNR. The results for $\epsilon_0 = 0$, 0.25 and 0.5 are presented in the left-hand panel of Fig. C2. For low ϵ_0 the bias increases monotonically, asymptoting to a value ∼0.95 for large ν, whereas for $\epsilon_0 = 0.5$, $b(\nu)$ increases first before declining. To capture the variation in $b(\nu)$ we adopt

$$b(\nu) = b_0(\epsilon_0) + \frac{b_1(\epsilon_0)}{\sqrt{\nu}} + \frac{b_2((\epsilon_0)}{\nu}. \quad \text{(C3)}$$

We note that fixing $b_0 \sim 0.95$ also gave reasonable results. Furthermore b_1 and b_2 are highly (anti-) correlated, and it thus might be possible to reduce the number of free parameters in principle. Interestingly the value for b_0, which corresponds to the bias for bright, large sources, is fairly close to unity. We list the best-fitting parameters in Table C1 and the right-hand panel in Fig. C2 shows that the parameters vary smoothly with ϵ_0.

In Section 2.4, we discuss how well the correction performs as a function of ellipticity distribution and seeing. We also examined $\mu_{\rm cor}$ as a function of apparent magnitude and found that residuals are <1 per cent for galaxies fainter than $m_r = 22$, whereas $\mu_{\rm cor} \sim 0.02$ for galaxies with $20 < m_r < 22$. This overcorrection is probably caused by the fact that our parametrization tends to underestimate the bias for large objects.

The bias depends on Sérsic index (see Fig. 8), and therefore the performance of the empirical correction may differ. Fig. C3 shows the residual bias as a function of ellipticity distribution for four different Sérsic indices. The black points with error bars indicate the results for the ensemble of galaxies, with half-light radii and Sérsic indices drawn from the observed distribution of values. The lines indicate the results for the simulations where the Sérsic indices were fixed to the indicated values. The range in bias is similar to what

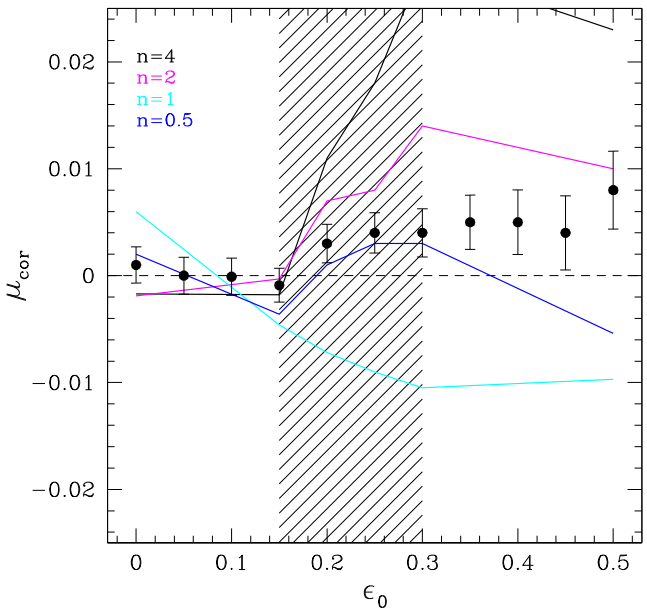

Figure C3. Bottom panel: residual multiplicative bias for galaxies with different values for the Sérsic index as a function of ϵ_0. The simulated galaxies have the same distribution in half-light radii as the regular simulations (indicated by the black points with error bars) and $22 < m_r < 25$. As in the real data we only include galaxies with an observed size $r_h < 5$ pixels. The full range spans about 0.03, but the uncertainty for the ensemble of sources is much less as the distribution of Sérsic profiles has been obtained from *HST* observations.

was observed in Fig. 8, with a positive bias in the expected range of ϵ_0 for $n = 4$ and a negative bias for $n = 1$. Given the relatively weak dependence of Sérsic index and the fact that the distribution of Sérsic values is well-determined from *HST* observations (Rix et al. 2004), we conclude that our correction for the ensemble is robust. Hence, we ignore the dependence of the bias on Sérsic index.

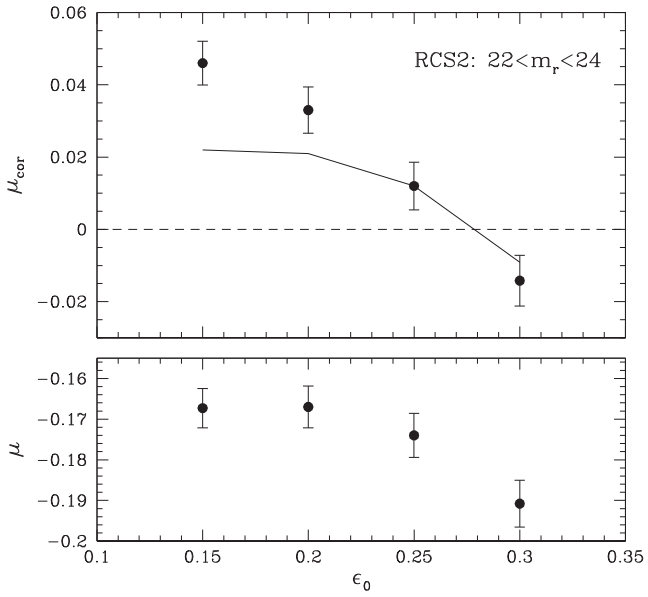

Figure D1. Bottom panel: multiplicative bias for galaxies with $20 < m_r < 24$ as a function of ϵ_0 for a simulation of RCS2 data, which are shallower than the CCCP data. Top panel: residual bias after we apply our correction, using parameters optimized for CCCP data, to the simulated RCS2 data. The solid line indicates $\mu_{\rm cor}$ if we use $\epsilon_0 = 0.25$ to correct the simulations for other ellipticity distributions, suggesting our approach is adequate for these data as well.

APPENDIX D: APPLICATION TO SIMULATED RCS2 DATA

Our KSB implementation has also been used by van Uitert et al. (2011) to measure the lensing signal around galaxies using data from the RCS2. It is therefore interesting to examine the impact of our findings on those results. Compared to our CCCP data (with a total integration time of 3600 s per position), the RCS2 data are much shallower, consisting of a single 480 s exposure in the r' band. We create a separate set of simulations where the noise level matches that of the RCS2 data. The resulting bias for galaxies with $22 < m_r < 24$ (the range used by van Uitert et al. 2011 for the RCS2 analysis) as a function of ϵ_0 is presented in the lower panel of Fig. D1. The actual bias is smaller because van Uitert et al. (2011) used the STEP1 implementation, also used in the CCCP analysis of H12. We also note that van Uitert et al. (2011) used SExtractor to detect objects, and used the resulting half-light radius for the weight function in the subsequent shape analysis.

Here we are interested whether our correction scheme can be used to data with different noise properties. We therefore apply the correction to the simulated RCS2 results and present the residual bias in the top panel of Fig. D1. The black line indicates $\mu_{\rm cor}$ if we use our reference value of $\epsilon_0 = 0.25$ to correct the measurements for the various ellipticity distributions. The results suggest that the parameters that were optimized to correct CCCP can be used for the shallower RCS2 data as well.

APPENDIX E: OBSCURATION BY CLUSTER MEMBERS

As described in Section 3.1, we account for the dilution of the lensing signal due to cluster galaxies in the source galaxy sample using the excess galaxy counts as a function of cluster-centric radius. Fig. 11 shows that the number density of bright cluster galaxies is substantial at small radii, which may affect the detection of sources. As shown in Simet & Mandelbaum (2014), this is an important source of bias for the measurement of the magnification signal, but also is relevant here as it leads to biases in the correction for contamination by cluster galaxies.

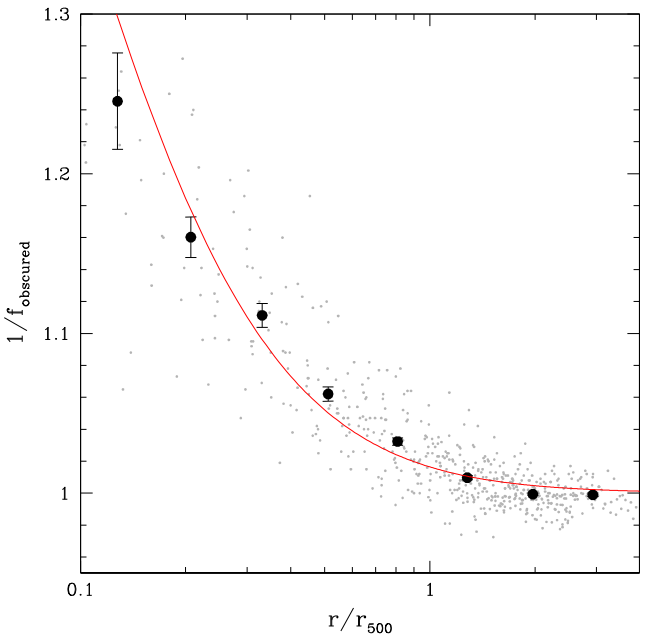

Figure E1. Plot of the correction of the counts of galaxies with $22 < m_r < 25$ for the obscuration by cluster galaxies as a function of radius in units of r_{500}. The black points indicate the average from a representative subset of clusters, whereas the light grey points show the individual measurements. The red line is the best-fitting model as described in the text.

To quantify the impact of the obscuration by cluster galaxies on the source galaxy counts, we use our image simulations: we add the cluster observations to the simulated image and perform the object detection and analysis on the images with and without the cluster added.[12] We identify the objects detected in both catalogues and measure their number density as a function of cluster-centric radius. As expected, at small radii we observe a decrease in the recovered number density.

Fig. E1 shows the corresponding correction for the source counts as a function of radius in units of r_{500}. The black points correspond to the average from a representative sample of clusters that we used in this study. The individual measurements are indicated by the light-grey points. For radii larger than r_{500} the observed excess counts are biased low by a few per cent. Hence, our aperture mass estimates for M_{500} are unaffected. On the smallest scales considered for the NFW fits, the observed counts are biased low by ∼10 per cent. Even in this case the impact is small, as this is a 10 per cent correction to a correction that itself is 30 per cent; the resulting change in the best-fitting mass is an ∼1–2 per cent increase. We find that the correction can be described by

$$\frac{1}{f_{\rm obscured}} = 1 + \frac{0.021}{0.14 + (r/r_{500})^2}, \qquad \text{(E1)}$$

[12] We ignore magnification which increases the fluxes of sources and thus reduces the effects of obscuration somewhat. Note that the number density of sources is not affected significantly by magnification because the power-law slope of the number counts $d\log N_{\rm gal}/dM \sim 0.38$–$0.4$, as discussed in Section 3.

which is the red line shown in Fig. E1. We use this model to correct the measurements of the contamination by cluster members.

APPENDIX F: MASS ESTIMATES USING DUFFY ET AL. (2008) $C(M)$-RELATION

Previous cluster weak lensing studies, including H12, presented mass measurements using the relation between mass and concentration from Duffy et al. (2008), which is based on numerical simulations assuming a *WMAP5* cosmology (Komatsu et al. 2009). The first results from *Planck* presented by Planck Collaboration XVI (2014b) suggest higher values for both the normalization of the matter power spectrum σ_8 and the mean density Ω_m. For this reason we adopted the relation from Dutton & Macciò (2014), which yields a concentration that is ∼20 per cent higher than Duffy et al. (2008) for a given mass.

This affects the masses inferred from parametric NFW fits and the deprojected aperture masses, but not necessarily in the same sense. For instance, if we switch to the Duffy et al. (2008) relation, the estimate for M_{2500} decreases by ∼7 per cent, no matter whether we consider the aperture mass or the best-fitting NFW model. On the other hand, the estimate for M_{500} increases by 5 per cent for the NFW model fit, whereas the deprojected aperture mass decreases by ∼3 per cent (also see section 4.3 in H12). Hence, to allow for a straightforward comparison with previous mass measurements, we present in Table F1 the results if we use the mass–concentration relation from Duffy et al. (2008).

Table F1. Weak lensing mass estimates for the CCCP sample using the mass–concentration from Duffy et al. (2008).

(1) Name	(2) $M_{0.5}^{\rm proj}$	(3) $M_{1.0}^{\rm proj}$	(4) r_{2500}	(5) $M_{2500}^{\rm ap}$	(6) r_{500}	(7) $M_{500}^{\rm ap}$	(8) $M_{\rm vir}^{\rm NFW}$	(9) $M_{2500}^{\rm NFW}$	(10) $M_{500}^{\rm NFW}$
Abell 68	5.0 ± 0.4	8.7 ± 1.2	560	$3.2^{+0.3}_{-0.3}$	1391	$9.9^{+1.5}_{-1.6}$	$14.2^{+3.1}_{-3.1}$	$2.8^{+0.6}_{-0.6}$	$7.8^{+1.7}_{-1.7}$
Abell 209	3.7 ± 0.5	7.8 ± 1.5	508	$2.3^{+0.4}_{-0.4}$	1249	$6.8^{+1.6}_{-1.5}$	$9.3^{+2.7}_{-2.5}$	$1.9^{+0.5}_{-0.5}$	$5.1^{+1.5}_{-1.4}$
Abell 267	4.6 ± 0.5	7.8 ± 1.5	546	$2.9^{+0.4}_{-0.4}$	1222	$6.6^{+1.6}_{-1.6}$	$8.6^{+2.8}_{-2.6}$	$1.8^{+0.6}_{-0.5}$	$4.8^{+1.5}_{-1.4}$
Abell 370	8.1 ± 0.6	16.9 ± 1.7	661	$6.1^{+0.6}_{-0.6}$	1659	$19.3^{+2.5}_{-2.5}$	$34.6^{+6.6}_{-6.3}$	$6.2^{+1.2}_{-1.1}$	$18.7^{+3.6}_{-3.4}$
Abell 383	2.8 ± 0.6	6.4 ± 1.4	445	$1.5^{+0.5}_{-0.6}$	1232	$6.4^{+2.2}_{-2.4}$	$6.2^{+2.9}_{-2.6}$	$1.3^{+0.6}_{-0.6}$	$3.5^{+1.6}_{-1.5}$
Abell 963	4.5 ± 0.5	7.9 ± 1.4	555	$3.0^{+0.5}_{-0.5}$	1274	$7.2^{+1.7}_{-1.7}$	$13.6^{+3.5}_{-3.4}$	$2.7^{+0.7}_{-0.7}$	$7.5^{+1.9}_{-1.8}$
Abell 1689	8.6 ± 0.5	15.0 ± 1.4	734	$6.7^{+0.6}_{-0.6}$	1616	$14.4^{+2.4}_{-2.2}$	$35.0^{+6.0}_{-5.7}$	$6.4^{+1.1}_{-1.0}$	$18.5^{+3.2}_{-3.0}$
Abell 1763	5.4 ± 0.5	10.5 ± 1.4	615	$4.1^{+0.6}_{-0.5}$	1529	$12.7^{+3.3}_{-2.9}$	$19.0^{+4.1}_{-3.9}$	$3.6^{+0.8}_{-0.8}$	$10.3^{+2.2}_{-2.1}$
Abell 2218	5.7 ± 0.5	9.3 ± 1.4	646	$4.6^{+0.7}_{-0.6}$	1403	$9.4^{+2.2}_{-2.2}$	$18.3^{+4.4}_{-4.2}$	$3.5^{+0.8}_{-0.8}$	$9.9^{+2.4}_{-2.3}$
Abell 2219	4.5 ± 0.6	10.1 ± 1.4	562	$3.2^{+0.6}_{-0.6}$	1418	$10.2^{+1.9}_{-1.7}$	$12.5^{+2.9}_{-2.9}$	$2.5^{+0.6}_{-0.6}$	$6.9^{+1.6}_{-1.6}$
Abell 2390	6.1 ± 0.5	11.6 ± 1.3	648	$4.9^{+0.6}_{-0.6}$	1407	$10.0^{+1.7}_{-1.5}$	$26.4^{+4.5}_{-4.3}$	$4.9^{+0.8}_{-0.8}$	$14.2^{+2.4}_{-2.3}$
MS 0015.9+1609	8.5 ± 0.6	20.8 ± 2.0	644	$6.9^{+0.7}_{-0.7}$	1654	$23.4^{+3.1}_{-3.0}$	$33.4^{+8.5}_{-8.0}$	$5.9^{+1.5}_{-1.4}$	$18.4^{+4.7}_{-4.4}$
MS 0906.5+1110	4.5 ± 0.5	9.6 ± 1.3	574	$3.2^{+0.5}_{-0.5}$	1457	$10.4^{+1.7}_{-1.8}$	$14.2^{+3.2}_{-3.2}$	$2.8^{+0.6}_{-0.6}$	$7.7^{+1.8}_{-1.8}$
MS 1224.7+2007	2.0 ± 0.6	2.6 ± 1.6	359	$0.9^{+0.4}_{-0.4}$	794	$2.0^{+0.8}_{-0.7}$	$5.2^{+2.6}_{-2.3}$	$1.1^{+0.6}_{-0.5}$	$3.0^{+1.5}_{-1.3}$
MS 1231.3+1542	1.2 ± 0.4	0.5 ± 1.2	359	$0.8^{+0.2}_{-0.2}$	584	$0.7^{+0.4}_{-0.4}$	$1.9^{+1.3}_{-1.2}$	$0.5^{+0.3}_{-0.3}$	$1.1^{+0.8}_{-0.7}$
MS 1358.4+6245	4.8 ± 0.7	9.4 ± 1.7	542	$3.2^{+0.7}_{-0.7}$	1316	$9.1^{+2.2}_{-2.2}$	$15.0^{+3.8}_{-3.6}$	$2.9^{+0.7}_{-0.7}$	$8.3^{+2.1}_{-2.0}$
MS 1455.0+2232	4.3 ± 0.4	8.1 ± 1.3	538	$2.9^{+0.4}_{-0.4}$	1191	$6.2^{+1.3}_{-1.2}$	$14.6^{+2.8}_{-2.8}$	$2.9^{+0.6}_{-0.6}$	$8.0^{+1.6}_{-1.6}$
MS 1512.4+3647	1.5 ± 0.6	4.5 ± 1.6	274	$0.4^{+0.3}_{-0.3}$	829	$2.4^{+1.8}_{-1.4}$	$4.1^{+2.0}_{-1.7}$	$0.9^{+0.4}_{-0.4}$	$2.4^{+1.2}_{-1.0}$
MS 1621.5+2640	5.4 ± 0.6	12.3 ± 1.6	565	$4.0^{+0.7}_{-0.7}$	1301	$9.9^{+2.0}_{-1.8}$	$21.8^{+5.0}_{-4.8}$	$4.0^{+0.9}_{-0.9}$	$12.0^{+2.7}_{-2.6}$
CL 0024.0+1652	6.7 ± 0.6	12.5 ± 1.8	597	$4.6^{+0.6}_{-0.6}$	1357	$10.7^{+1.9}_{-1.9}$	$27.7^{+6.8}_{-6.4}$	$5.1^{+1.2}_{-1.2}$	$15.1^{+3.7}_{-3.5}$
Abell 115N	1.4 ± 0.5	5.3 ± 1.3	283	$0.4^{+0.3}_{-0.4}$	1087	$4.4^{+1.3}_{-1.9}$	$6.6^{+2.2}_{-2.2}$	$1.4^{+0.5}_{-0.5}$	$3.7^{+1.2}_{-1.2}$
Abell 115S	2.6 ± 0.5	6.0 ± 1.3	399	$1.1^{+0.5}_{-0.5}$	1116	$4.8^{+1.3}_{-1.2}$	$7.9^{+2.4}_{-2.4}$	$1.6^{+0.5}_{-0.5}$	$4.4^{+1.3}_{-1.3}$
Abell 222	2.9 ± 0.4	6.9 ± 1.2	433	$1.4^{+0.5}_{-0.7}$	1165	$5.6^{+1.2}_{-1.1}$	$7.1^{+2.4}_{-2.3}$	$1.5^{+0.5}_{-0.5}$	$4.0^{+1.4}_{-1.3}$
Abell 223N	3.0 ± 0.5	7.6 ± 1.2	448	$1.6^{+0.5}_{-0.7}$	1226	$6.5^{+1.3}_{-1.2}$	$10.0^{+3.0}_{-2.8}$	$2.0^{+0.6}_{-0.6}$	$5.5^{+1.7}_{-1.6}$
Abell 223S	3.0 ± 0.4	8.4 ± 1.2	451	$1.6^{+0.5}_{-0.6}$	1355	$8.7^{+1.6}_{-1.5}$	$8.8^{+2.9}_{-2.7}$	$1.8^{+0.6}_{-0.6}$	$4.9^{+1.6}_{-1.5}$
Abell 520	3.6 ± 0.5	7.4 ± 1.4	513	$2.3^{+0.5}_{-0.5}$	1201	$6.0^{+1.5}_{-1.3}$	$16.9^{+3.5}_{-3.4}$	$3.3^{+0.7}_{-0.7}$	$9.2^{+1.9}_{-1.8}$
Abell 521	3.1 ± 0.5	9.0 ± 1.4	401	$1.2^{+1.0}_{-0.5}$	1321	$8.5^{+1.7}_{-1.6}$	$12.2^{+3.5}_{-3.4}$	$2.4^{+0.7}_{-0.7}$	$6.7^{+2.0}_{-1.9}$
Abell 586	2.5 ± 0.6	6.1 ± 1.4	430	$1.3^{+0.5}_{-0.4}$	1203	$5.9^{+2.4}_{-2.2}$	$5.0^{+2.4}_{-2.2}$	$1.1^{+0.5}_{-0.5}$	$2.8^{+1.3}_{-1.2}$
Abell 611	3.7 ± 0.5	9.0 ± 1.3	489	$2.2^{+0.6}_{-0.5}$	1226	$7.1^{+1.2}_{-1.2}$	$10.1^{+3.4}_{-3.1}$	$2.0^{+0.7}_{-0.6}$	$5.6^{+1.9}_{-1.7}$
Abell 697	4.6 ± 0.5	10.5 ± 1.4	551	$3.2^{+0.5}_{-0.5}$	1417	$10.8^{+1.7}_{-2.0}$	$15.7^{+3.8}_{-3.6}$	$3.1^{+0.7}_{-0.7}$	$8.6^{+2.1}_{-2.0}$
Abell 851	5.4 ± 0.5	12.2 ± 1.4	540	$3.4^{+0.4}_{-0.4}$	1348	$10.7^{+1.8}_{-1.7}$	$24.1^{+6.2}_{-5.7}$	$4.4^{+1.1}_{-1.1}$	$13.2^{+3.4}_{-3.2}$
Abell 959	5.0 ± 0.5	10.9 ± 1.4	580	$3.7^{+0.6}_{-0.6}$	1333	$9.0^{+1.5}_{-1.5}$	$22.3^{+4.4}_{-4.4}$	$4.2^{+0.8}_{-0.8}$	$12.1^{+2.4}_{-2.4}$
Abell 1234	2.5 ± 0.5	4.4 ± 1.4	439	$1.4^{+0.3}_{-0.3}$	982	$3.2^{+1.3}_{-1.1}$	$8.2^{+2.6}_{-2.4}$	$1.7^{+0.5}_{-0.5}$	$4.5^{+1.4}_{-1.3}$
Abell 1246	2.7 ± 0.5	5.7 ± 1.1	426	$1.3^{+0.4}_{-0.5}$	1082	$4.4^{+0.9}_{-0.9}$	$9.7^{+2.7}_{-2.5}$	$2.0^{+0.5}_{-0.5}$	$5.4^{+1.5}_{-1.4}$

Table F1 – *continued.*

(1) Name	(2) $M_{0.5}^{\rm proj}$	(3) $M_{1.0}^{\rm proj}$	(4) r_{2500}	(5) $M_{2500}^{\rm ap}$	(6) r_{500}	(7) $M_{500}^{\rm ap}$	(8) $M_{\rm vir}^{\rm NFW}$	(9) $M_{2500}^{\rm NFW}$	(10) $M_{500}^{\rm NFW}$
Abell 1758	5.5 ± 0.5	12.2 ± 1.4	633	$4.8^{+0.6}_{-0.6}$	1491	$12.6^{+1.9}_{-1.9}$	$21.6^{+3.6}_{-3.6}$	$4.1^{+0.7}_{-0.7}$	$11.7^{+2.0}_{-2.0}$
Abell 1835	5.3 ± 0.4	10.7 ± 1.2	603	$4.0^{+0.4}_{-0.4}$	1387	$9.8^{+1.5}_{-1.5}$	$22.4^{+4.2}_{-4.2}$	$4.2^{+0.8}_{-0.8}$	$12.1^{+2.3}_{-2.3}$
Abell 1914	3.7 ± 0.5	7.9 ± 1.2	522	$2.4^{+0.4}_{-0.4}$	1287	$7.2^{+1.4}_{-1.3}$	$15.1^{+2.8}_{-2.7}$	$3.0^{+0.5}_{-0.5}$	$8.2^{+1.5}_{-1.5}$
Abell 1942	3.8 ± 0.5	7.6 ± 1.2	519	$2.5^{+0.4}_{-0.4}$	1204	$6.2^{+1.3}_{-1.2}$	$15.1^{+3.2}_{-3.1}$	$3.0^{+0.6}_{-0.6}$	$8.2^{+1.8}_{-1.7}$
Abell 2104	4.2 ± 0.5	10.3 ± 1.2	583	$3.3^{+0.6}_{-0.6}$	1422	$9.5^{+1.9}_{-1.6}$	$17.8^{+4.0}_{-3.8}$	$3.4^{+0.8}_{-0.7}$	$9.6^{+2.2}_{-2.1}$
Abell 2111	3.9 ± 0.5	6.6 ± 1.3	517	$2.5^{+0.4}_{-0.4}$	1156	$5.5^{+1.6}_{-1.5}$	$10.1^{+2.5}_{-2.5}$	$2.1^{+0.5}_{-0.5}$	$5.6^{+1.4}_{-1.4}$
Abell 2163	4.4 ± 0.5	9.5 ± 1.4	562	$3.1^{+0.6}_{-0.5}$	1456	$10.8^{+2.2}_{-2.1}$	$19.6^{+4.4}_{-4.2}$	$3.7^{+0.8}_{-0.8}$	$10.6^{+2.4}_{-2.3}$
Abell 2204	4.8 ± 0.5	11.0 ± 1.0	619	$3.9^{+0.5}_{-0.5}$	1490	$10.9^{+1.4}_{-1.3}$	$22.4^{+3.8}_{-3.6}$	$4.2^{+0.7}_{-0.7}$	$12.0^{+2.0}_{-1.9}$
Abell 2259	2.4 ± 0.5	5.6 ± 1.3	417	$1.2^{+0.4}_{-0.4}$	1106	$4.5^{+1.2}_{-1.3}$	$8.8^{+3.1}_{-2.7}$	$1.8^{+0.6}_{-0.6}$	$4.8^{+1.7}_{-1.5}$
Abell 2261	6.0 ± 0.4	14.2 ± 1.3	666	$5.3^{+0.5}_{-0.5}$	1654	$16.1^{+1.6}_{-1.6}$	$28.1^{+4.9}_{-4.7}$	$5.2^{+0.9}_{-0.9}$	$15.1^{+2.6}_{-2.5}$
Abell 2537	5.4 ± 0.6	10.1 ± 1.4	585	$3.9^{+0.6}_{-0.6}$	1302	$8.5^{+1.6}_{-1.5}$	$23.6^{+4.6}_{-4.6}$	$4.4^{+0.9}_{-0.9}$	$12.9^{+2.5}_{-2.5}$
MS 0440.5+0204	2.9 ± 0.5	2.8 ± 1.3	457	$1.6^{+0.4}_{-0.4}$	890	$2.4^{+0.7}_{-0.7}$	$3.5^{+2.0}_{-1.8}$	$0.8^{+0.4}_{-0.4}$	$2.0^{+1.1}_{-1.0}$
MS 0451.6−0305	4.4 ± 0.7	8.4 ± 2.1	453	$2.4^{+0.6}_{-0.6}$	1071	$6.3^{+1.7}_{-1.9}$	$19.4^{+6.7}_{-6.2}$	$3.6^{+1.2}_{-1.1}$	$10.8^{+3.7}_{-3.5}$
MS 1008.1−1224	4.1 ± 0.4	8.2 ± 1.4	507	$2.5^{+0.4}_{-0.3}$	1168	$6.2^{+1.3}_{-1.2}$	$18.2^{+4.0}_{-3.8}$	$3.5^{+0.8}_{-0.7}$	$10.0^{+2.2}_{-2.1}$
RX J1347.5−1145	5.2 ± 0.8	10.1 ± 1.9	530	$3.4^{+0.9}_{-0.8}$	1301	$10.1^{+3.2}_{-3.0}$	$22.0^{+6.0}_{-5.7}$	$4.1^{+1.1}_{-1.0}$	$12.2^{+3.3}_{-3.1}$
RX J1524.6+0957	2.1 ± 0.9	6.5 ± 2.1	245	$0.4^{+0.4}_{-1003.0}$	961	$4.4^{+2.2}_{-1.9}$	$7.1^{+4.8}_{-4.2}$	$1.4^{+1.0}_{-0.9}$	$4.1^{+2.8}_{-2.4}$
MACS J0717.5+3745	6.4 ± 0.9	19.3 ± 2.3	586	$5.2^{+1.1}_{-3.1}$	1470	$16.4^{+3.4}_{-3.0}$	$44.8^{+10.8}_{-10.3}$	$7.6^{+1.8}_{-1.8}$	$24.4^{+5.9}_{-5.6}$
MACS J0913.7+4056	3.1 ± 0.5	5.3 ± 1.5	380	$1.2^{+0.5}_{-0.5}$	935	$3.7^{+1.0}_{-0.9}$	$6.9^{+3.7}_{-3.2}$	$1.4^{+0.8}_{-0.7}$	$4.0^{+2.1}_{-1.8}$
CIZA J1938+54	5.3 ± 0.5	11.3 ± 1.3	588	$3.8^{+0.4}_{-0.4}$	1557	$14.0^{+2.2}_{-2.2}$	$19.4^{+5.0}_{-4.7}$	$3.7^{+1.0}_{-0.9}$	$10.6^{+2.7}_{-2.6}$
3C295	4.6 ± 0.6	8.1 ± 1.9	488	$2.7^{+0.5}_{-0.5}$	1090	$6.0^{+1.5}_{-1.6}$	$13.4^{+4.6}_{-4.2}$	$2.6^{+0.9}_{-0.8}$	$7.5^{+2.6}_{-2.4}$

Column 1: cluster name; columns 2 and 3: projected mass within an aperture of 0.5 and 1 h_{70}^{-1} Mpc, respectively; columns 4 and 6: r_Δ (in units of h_{70}^{-1} kpc) determined using aperture masses; columns 5 and 7: deprojected aperture masses within r_Δ; columns 8–10: masses from best-fitting NFW model. All masses are listed in units of ($10^{14}\, h_{70}^{-1}\, {\rm M}_\odot$).

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