

# Improved optical mass tracer for galaxy clusters calibrated using weak lensing measurements

R. Reyes,<sup>1\*</sup> R. Mandelbaum,<sup>2\*†</sup> C. Hirata,<sup>3</sup> N. Bahcall<sup>1</sup> and U. Seljak<sup>4,5</sup>

<sup>1</sup>*Princeton University Observatory, Peyton Hall, Ivy Lane, Princeton, NJ 08544, USA*

<sup>2</sup>*Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

<sup>3</sup>*Mail Code 130-33, Caltech, Pasadena, CA 91125, USA*

<sup>4</sup>*Institute for Theoretical Physics, University of Zurich, Zurich, 8057 Switzerland*

<sup>5</sup>*Department of Physics, University of California, Berkeley, CA 94720, USA*

Accepted 2008 August 11. Received 2008 August 11; in original form 2008 February 17

## ABSTRACT

We develop an improved mass tracer for clusters of galaxies from optically observed parameters, and calibrate the mass relation using weak gravitational lensing measurements. We employ a sample of  $\sim 13\,000$  optically selected clusters from the Sloan Digital Sky Survey (SDSS) maxBCG catalogue, with photometric redshifts in the range 0.1–0.3. The optical tracers we consider are cluster richness, cluster luminosity, luminosity of the brightest cluster galaxy (BCG) and combinations of these parameters. We measure the weak lensing signal around stacked clusters as a function of the various tracers, and use it to determine the tracer with the least amount of scatter. We further use the weak lensing data to calibrate the mass normalization. We find that the best mass estimator for massive clusters is a combination of cluster richness,  $N_{200}$ , and the luminosity of the BCG,  $L_{\text{BCG}}$ :  $M_{200\bar{p}} = (1.27 \pm 0.08)(N_{200}/20)^{1.20 \pm 0.09} [L_{\text{BCG}}/\bar{L}_{\text{BCG}}(N_{200})]^{0.71 \pm 0.14} \times 10^{14} h^{-1} M_{\odot}$ , where  $\bar{L}_{\text{BCG}}(N_{200})$  is the observed mean BCG luminosity at a given richness. This improved mass tracer will enable the use of galaxy clusters as a more powerful tool for constraining cosmological parameters.

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

## 1 INTRODUCTION

Clusters of galaxies trace the matter density distribution in the Universe, and they have long been used successfully as powerful cosmological probes. Relating the observed cluster abundance to the dark matter halo abundance predicted by cosmological simulations provides powerful constraints on a range of cosmological parameters, including the amplitude of matter fluctuations, neutrino mass and dark energy density (Bahcall & Cen 1992; Haiman, Mohr & Holder 2001; Weller & Battye 2003; Wang et al. 2005; Albrecht et al. 2006; Mandelbaum & Seljak 2007). The strength of these constraints arises from the exponential cut-off in the cluster mass function for the most massive clusters, which depends strongly on both the amplitude of matter fluctuations and the matter density.

Currently, the use of clusters as precise cosmological probes is limited by the lack of reliable mass estimates for a large sample of clusters. While hydrodynamic simulations can provide estimates for the relation between X-ray observable parameters and cluster

mass (e.g. Kravtsov, Vikhlinin & Nagai 2006; Nagai, Kravtsov & Vikhlinin 2007), it is not clear that all the relevant physics determining these relations exist in the simulations. Estimating the virial mass of individual clusters using X-ray measurements (e.g. Schmidt & Allen 2007) requires the assumption of hydrostatic equilibrium, which introduces potential systematics for non-relaxed clusters, and neglects the effects of non-thermal pressure support, such as that from turbulence, cosmic rays and magnetic fields. There is a hint of a  $\sim 20$  per cent conflict between theoretical predictions and observations for the normalizations of these mass relations (Arnaud, Pointecouteau & Pratt 2007; Nagai, Kravtsov & Vikhlinin 2007). This discrepancy between hydrostatic masses and total mass also appears to be supported by observational results (Mahdavi et al. 2008). Thus, a careful treatment is necessary before they can be used for precision cosmology.

A way to estimate cluster masses that is insensitive to the dynamical state of the system is through weak gravitational lensing measurements. These directly probe the total (dark plus luminous) matter distribution. Estimates of the mass of individual clusters using weak lensing are currently limited to  $\sim 30$  per cent uncertainties by the signal-to-noise ratio of the lensing measurements, for clusters with  $M_{500} \sim \text{few} \times 10^{14} h^{-1} M_{\odot}$  (e.g. Hoekstra 2007; Pedersen & Dahle 2007). They are also subject to systematics such

\*E-mail: rreyes@astro.princeton.edu (RR); rmandelb@ias.edu (RM)

†Hubble Fellow.

as the shear and source redshift calibration, and limitations due to projection effects of matter near the cluster or along the line-of-sight (Metzler, White & Loken 2001; Hoekstra 2003). These probes can be augmented by strong gravitational lensing measurements (Bradač et al. 2005; Cacciato et al. 2006) and velocity dispersion measurements (Becker et al. 2007) to aid in the cluster mass determination (e.g. using the methods of Mahdavi et al. 2007 and Sereno 2007).

Here, we calibrate the mass relations for a range of optical parameters using measurements of the stacked weak lensing signal around a large set of clusters. This approach is complementary to those methods that provide mass estimates for individual clusters, which cannot currently be fully applied to large data sets. For example, velocity dispersion measurements are limited by the practical difficulty of obtaining spectroscopic observations for a large number of clusters. Our method for mass calibration can be readily applied to data sets from upcoming large-scale surveys, such as Dark Energy Survey (DES),<sup>1</sup> Pan-STARRS<sup>2</sup> and Large Synoptic Survey Telescope (LSST).<sup>3</sup>

We employ the largest available sample of  $\sim 13\,000$  galaxy clusters (maxBCG cluster catalogue; Koester et al. 2007a,b) selected from the Sloan Digital Sky Survey (SDSS; York et al. 2000). Stacking the weak lensing signals around many clusters increases the signal-to-noise ratio that we can achieve. The availability of accurate photometric redshifts for all objects in the sample also improves our mass measurements. Independent weak lensing analyses of clusters in this catalogue have been performed (Johnston et al. 2007; Sheldon et al. 2007a,b). Closest to this work is Johnston et al. (2007), where scaling relations of cluster mass with optical richness and cluster luminosity were obtained using a different method for estimating the cluster mass.

In this work, we consider optical tracers available in large cluster surveys, such as cluster richness, cluster luminosity and luminosity of the brightest cluster galaxy (BCG), and assess how well these parameters trace the cluster mass. In addition, we consider combinations of these parameters and assess whether they provide better mass determinations. Finding the most faithful tracer of cluster mass among the available options will allow us to fully harness the power of clusters in constraining cosmological parameters.

The paper is organized as follows. In Section 2, we describe the cluster catalogue and the weak lensing measurements. In Section 3, we describe how we use stacked weak lensing measurements to estimate cluster masses, and discuss our approach for assessing mass tracers in Section 3.5. Section 4 deals with various tests of systematics. We present our results in Section 5 and conclude in Section 6.

## 2 DATA

In this section, we describe the SDSS data (Section 2.1), the lens cluster sample from the maxBCG cluster catalogue (Section 2.2) and the source galaxy catalogue used in the weak lensing analysis (Section 2.3).

### 2.1 SDSS data

The maxBCG cluster catalogue and the lensing source catalogue come from the SDSS, a survey to image roughly  $\pi$  steradians of

the sky, and follow up approximately one million of the detected objects spectroscopically (Eisenstein et al. 2001; Richards et al. 2002; Strauss et al. 2002). The imaging is carried out by drift scanning the sky in photometric conditions (Hogg et al. 2001; Ivezić et al. 2004) in five bands (*ugriz*) (Fukugita et al. 1996; Smith et al. 2002) using a specially designed wide-field camera (Gunn et al. 1998). These imaging data are used to create the source catalogue that we use in this paper. In addition, objects are targeted for spectroscopy using these data (Blanton et al. 2003a) and are observed with a double 320-fibre spectrograph on the same telescope (Gunn et al. 2006). All of these data are processed by automated pipelines that detect and measure photometric properties of sources, and astrometrically calibrate the data (Lupton et al. 2001; Pier et al. 2003; Tucker et al. 2006). The SDSS is nearly complete, and has had seven major data releases (Stoughton et al. 2002; Abazajian et al. 2003, 2004, 2005; Finkbeiner et al. 2004; Adelman-McCarthy et al. 2006, 2007, 2008).

### 2.2 Cluster lens sample

Our lens sample consists of 12 612 clusters from the public maxBCG catalogue, with richness in red galaxies of  $N_{200} \geq 10$  (where the galaxy count includes galaxies brighter than  $0.4L_*$  and located within a scaled radius of  $r_{200}$ , defined in equation (1)). The clusters have photometric redshifts in the range of  $z = 0.1\text{--}0.3$ , selected over a  $0.5 (h^{-1} \text{Gpc})^3$  volume covering  $7500 \text{ deg}^2$  of sky. Our sample excludes  $\sim 9$  per cent of the solid angle covered by the survey where lensing shape measurements of source galaxies are currently not available. The maxBCG catalogue is presented and discussed in detail by Koester et al. (2007a,b). In this section, we briefly describe the cluster finder algorithm, and define the cluster properties used in this work.

The maxBCG cluster finder exploits the existence of the E/S0 red ridgeline of cluster galaxies in the colour–magnitude diagram, and of a BCG found near the centre of most clusters. For each galaxy, it obtains a photometric redshift estimate by maximizing the likelihood that (i) it is located in an overdensity of E/S0 ridgeline galaxies of similar colours, and (ii) it has colours and magnitudes of a typical BCG at that redshift. It also determines  $N_{1 \text{ Mpc}}$ , the number of E/S0 ridgeline galaxies located within a projected distance of  $1 h^{-1} \text{ Mpc}$  of the galaxy, which are dimmer than the galaxy and brighter than  $0.4L_*$ , where  $L^* = 2.08 \times 10^{10} h^{-2} L_{\odot}$  in the *i* band at  $z = 0.1$ , with a dependence on redshift determined from a Pegase-2 stellar population/galaxy formation model, similar to that of Eisenstein et al. (2001). It then chooses the galaxy with the highest likelihood and  $N_{1 \text{ Mpc}}$  as a bona fide BCG.

To identify cluster members, the cluster size is estimated to be  $r_{200}$ , the radius within which the galaxy number density of the cluster is  $200\Omega_m^{-1}$  times the mean density of galaxies in the present Universe. The scaled radius  $r_{200}$  is estimated from the empirical relation from Hansen et al. (2005):

$$r_{200} = 0.156 N_{1 \text{ Mpc}}^{0.6} h^{-1} \text{ Mpc}. \quad (1)$$

The cluster finder identifies galaxies within a scaled radius  $r_{200}$  of the BCG, removes them from the list of potential cluster centres, and continues down the list of galaxies with lower likelihood and lower  $N_{1 \text{ Mpc}}$  until all candidates are exhausted. For more details, see Koester et al. (2007a,b).

Koester et al. (2007a,b) performed tests of purity and completeness of the maxBCG catalogue using mock catalogues from *N*-body simulations. They found that the sample is more than 90 per cent pure for clusters with  $N_{200} \geq 10$ ; and 90–95 per cent pure for

<sup>1</sup> <https://www.darkenergysurvey.org/>

<sup>2</sup> <http://pan-starrs.ifa.hawaii.edu/public/>

<sup>3</sup> <http://www.lsst.org/>

clusters with  $N_{200} \geq 20$ . The sample is >90 per cent complete for masses  $M_{200} \gtrsim 2 \times 10^{14} h^{-1} M_{\odot}$ , and >95 per cent complete for masses  $M_{200} \gtrsim 3 \times 10^{14} h^{-1} M_{\odot}$ , where  $M_{200}$  is the mass within  $r_{200}$ . These results are of course subject to the assumption that the mock catalogues are a faithful representation of the clusters.

In this work, we use three optical properties of clusters that are reported in the maxBCG catalogue.

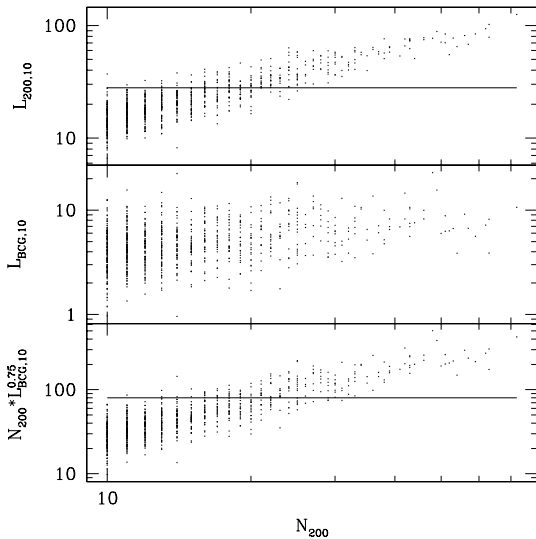
(i)  $N_{200}$  (cluster richness): the number of E/S0 ridgeline member galaxies fainter than the BCG, brighter than  $0.4L^*$ , and located within a projected distance  $r_{200}$  (given by equation 1) from the BCG.

(ii)  $L_{200}$  (cluster luminosity): the summed  $r$ -band luminosities of the BCG and the ridgeline member galaxies included in  $N_{200}$ ,  $k$ -corrected to  $z = 0.25$ . We usually express this luminosity in units of  $10^{10} h^{-2} L_{\odot}$  and denote it by  $L_{200,10}$ .

(iii)  $L_{\text{BCG}}$  (BCG luminosity): the  $r$ -band luminosity of the BCG,  $k$ -corrected to  $z = 0.25$ . We usually express this luminosity in units of  $10^{10} h^{-2} L_{\odot}$  and denote it by  $L_{\text{BCG},10}$ .

These luminosities are based on SDSS ‘cmodel’ magnitudes, which are constructed from a weighted combination of de Vaucouleurs and exponential magnitudes. The weights are determined by fitting the galaxy surface brightness profile with a linear combination of the best-fitting de Vaucouleurs and exponential profiles.  $K$ -corrections are calculated from the luminous red galaxy (LRG) template in v4.1.4 of KCORRECT (Blanton et al. 2003b), using photometric redshifts and without applying a correction for evolution. Galactic extinction correction is applied using the extinction maps of Schlegel, Finkbeiner & Davis (1998). We note that these luminosities may be underestimated (at the 10 per cent level) due to systematic errors in sky subtraction, which is most severe in galaxies of large extent (Adelman-McCarthy et al. 2008).

Fig. 1 shows the correlation of the cluster richness in red galaxies  $N_{200}$  with other optical parameters for the richness-selected cluster sample ( $N_{200} \geq 10$ ). There is a strong correlation between  $N_{200}$  and  $L_{200}$  (with a rank correlation coefficient of 0.68). The sample is



**Figure 1.** Correlation of  $N_{200}$  (cluster richness in red galaxies) with  $L_{200}$  (cluster luminosity in red galaxies),  $L_{\text{BCG}}$  (BCG luminosity) and the combination  $N_{200}L_{\text{BCG}}^{0.75}$  (with luminosities in units of  $10^{10} h^{-2} L_{\odot}$ ). The cluster sample, which is selected by richness ( $N_{200} \geq 10$ ), is complete above  $L_{200,10} = 30$  and above  $N_{200}L_{\text{BCG},10}^{0.75} = 80$  (solid horizontal lines), but is not complete in  $L_{\text{BCG}}$  at any value.

complete for cluster luminosities  $L_{200,10} \geq 30$  (uppermost panel). On the other hand, while the minimum value of  $L_{\text{BCG}}$  correlates with  $N_{200}$ , the maximum value of  $L_{\text{BCG}}$  does not. The two parameters are weakly correlated, with rank correlation coefficient is 0.30. The scatter in  $L_{\text{BCG}}$  at fixed richness has a Gaussian distribution with width  $\gtrsim 0.17$  dex (Hansen et al. 2007). The sample is not complete in  $L_{\text{BCG}}$  even at the brightest end (middle panel). However, the sample is complete at  $N_{200}L_{\text{BCG},10}^{0.75} \geq 80$  (lowermost panel). The  $1\sigma$  statistical error in the luminosities is roughly 0.06 dex (dominated by photometric redshift error), and is much smaller than the observed scatter.

### 2.3 Source catalogue

The source galaxy sample used for the weak lensing measurements is the same as that originally described in Mandelbaum et al. (2005a, hereafter M05). This source sample includes over 30 million galaxies from the SDSS imaging data with  $r$ -band model magnitude brighter than 21.8, with shape measurements obtained using the REGLENS pipeline, including point spread function (PSF) correction done via re-Gaussianization (Hirata & Seljak 2003) and with cuts designed to avoid various shear calibration biases. A full description of this pipeline can be found in M05.

The REGLENS pipeline obtains galaxy images in the  $r$  and  $i$  filters from the SDSS ‘atlas images’ (Stoughton et al. 2002). The basic principle of shear measurement using these images is to fit a Gaussian profile with elliptical isophotes to the image, and define the components of the ellipticity

$$(e_+, e_x) = \frac{1 - (b/a)^2}{1 + (b/a)^2} (\cos 2\phi, \sin 2\phi), \quad (2)$$

where  $b/a$  is the axis ratio and  $\phi$  is the position angle of the major axis. The ellipticity is then an estimator for the shear,

$$(\gamma_+, \gamma_x) = \frac{1}{2\mathcal{R}} \langle (e_+, e_x) \rangle, \quad (3)$$

where  $\mathcal{R} \approx 0.87$  is called the ‘shear responsivity’ and represents the response of the ellipticity (equation 2) to a small shear (Kaiser, Squires & Broadhurst 1995; Bernstein & Jarvis 2002). In practice, a number of corrections need to be applied to obtain the ellipticity. The most important of these is the correction for the smearing and circularization of the galactic images by the PSF; M05 uses the PSF maps obtained from stellar images by the PSP pipeline (Lupton et al. 2001), and corrects for these using the re-Gaussianization technique of Hirata & Seljak (2003), which includes corrections for non-Gaussianity of both the galaxy profile and the PSF. In order for these corrections to be successful, we require that the galaxy be well resolved compared to the PSF in both  $r$  and  $i$  bands (the only ones used for shape measurement). To do this we define the Gaussian resolution factor:

$$R_2 = 1 - \frac{T^{(P)}}{T^{(I)}}, \quad (4)$$

where the  $T$  values are the traces of the adaptive covariance matrices, and the superscripts indicate whether they are of the PSF or of the galaxy image. A large galaxy (compared to the PSF) would have  $R_2 \approx 1$ , while a star or other unresolved source would have  $R_2 \approx 0$ . We require that  $R_2$  exceed  $1/3$  in both  $r$  and  $i$  bands.

### 3 CLUSTER MASSES FROM STACKED WEAK LENSING MEASUREMENTS

In this section, we describe how we estimate cluster masses using stacked weak lensing measurements. We discuss theory

(Section 3.1), computation of the lensing signal (Section 3.2), modelling of the density profiles (Section 3.3), fits to the observed lensing signal to obtain cluster masses (Section 3.4) and interpretation of the best-fitting masses (Section 3.5).

### 3.1 Theory

Cluster–galaxy lensing provides a simple way to probe the connection between galaxies and matter via their cross-correlation function:

$$\xi_{\text{gm}}(\mathbf{r}) = \langle \delta_{\text{g}}(\mathbf{x})\delta_{\text{m}}(\mathbf{x} + \mathbf{r}) \rangle, \quad (5)$$

where  $\delta_{\text{g}}$  and  $\delta_{\text{m}}$  are overdensities of galaxies and matter, respectively. This cross-correlation can be related to the projected surface density

$$\Sigma(R) = \bar{\rho} \int \left[ 1 + \xi_{\text{gm}}(\sqrt{R^2 + \chi^2}) \right] d\chi \quad (6)$$

(where  $r^2 = R^2 + \chi^2$ ) which is then related to the observable quantity for lensing:

$$\Delta\Sigma(R) = \gamma_t(R)\Sigma_c = \bar{\Sigma}(\langle R \rangle) - \Sigma(R), \quad (7)$$

where  $\gamma_t$  is the tangential shear. The second relation is true only in the weak lensing limit, for a matter distribution that is axisymmetric along the line of sight. This symmetry is naturally achieved by our procedure of stacking many clusters and determining their average lensing signal. This observable quantity can be expressed as the product of the tangential shear  $\gamma_t$  and a geometric factor:

$$\Sigma_c = \frac{c^2}{4\pi G} \frac{D_s}{D_L D_{LS} (1 + z_L)^2}, \quad (8)$$

where  $D_L$  and  $D_s$  are angular diameter distances to the lens and source,  $D_{LS}$  is the angular diameter distance between the lens and source and the factor of  $(1 + z_L)^{-2}$  arises due to our use of comoving coordinates. For a given lens redshift,  $\Sigma_c^{-1}$  rises from zero at  $z_s = z_L$  to an asymptotic value at  $z_s \gg z_L$ ; that asymptotic value is an increasing function of lens redshift.

In practice, we truncate the integral in equation (6) at the virial radius of the cluster (defined in equation 12), motivated by attempts to model the lensing signal in simulations (M05). Truncation at two times the virial radius would change the cluster mass estimates at the 5 per cent level.

### 3.2 Signal computation

To compute the average lensing signal  $\Delta\Sigma(R)$ , lens–source pairs are first assigned weights according to the error on the shape measurement via

$$w_{\text{ls}} = \frac{\Sigma_c^{-2}}{\sigma_s^2 + \sigma_{\text{SN}}^2}, \quad (9)$$

where  $\sigma_{\text{SN}}^2$ , the intrinsic shape noise, was determined as a function of magnitude in M05, fig. 3. The factor of  $\Sigma_c^{-2}$  downweights pairs that are close in redshift, converting the shape noise in the denominator to a noise in  $\Delta\Sigma$ .

Once we have computed these weights, we compute the lensing signal in 62 logarithmic radial bins from 0.02 to  $9 h^{-1}$  Mpc as a summation over lens–source pairs via

$$\Delta\Sigma(R) = \frac{\sum_{\text{ls}} w_{\text{ls}} \gamma_t^{(\text{ls})} \Sigma_c}{2\mathcal{R} \sum_{\text{ls}} w_{\text{ls}}}, \quad (10)$$

where the factor of 2 arises due to our definition of ellipticity.

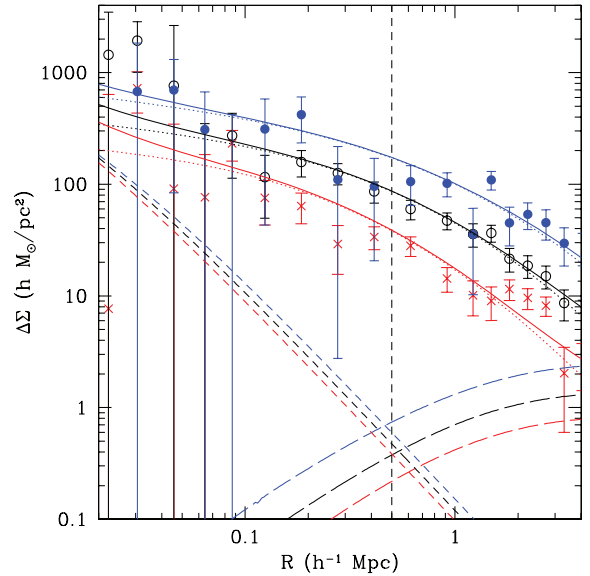
There are several additional procedures that must be done when computing the signal (for more detail, see M05). First, the signal computed around random points must be subtracted from the signal around real lenses to eliminate contributions from systematic shear. The measured signal around random points is consistent with zero over the range of radii we use. Subtraction of this signal introduces noise with rms of  $\sim 15$  per cent on scales from 0.5 to  $1 h^{-1}$  Mpc, and  $\sim 1$  per cent from 1 to  $9 h^{-1}$  Mpc.

Secondly, the signal must be boosted, i.e. multiplied by  $B(R) = n(R)/n_{\text{rand}}(R)$ , the ratio of the number density of sources relative to the number density around random points, in order to account for the dilution of the lensing signal due to sources that are physically associated with a lens (i.e. cluster galaxy members), and therefore not lensed. We find that  $B(R)$  decreases with increasing distance from the centre ranging from  $\sim 1.2$  to  $1.4$  at  $R = 0.5 h^{-1}$  Mpc (for low- to high-mass clusters), and dropping to unity for  $R \gtrsim 4 h^{-1}$  Mpc.

To determine errors on the lensing signal, we divide the survey area into 200 bootstrap subregions, and generate 2500 bootstrap-resampled data sets. Furthermore, to decrease noise in the covariance matrices due to the bootstrap, we rebin the signal into 22 radial bins (of which seven are in the range of radii we use for our fits).

### 3.3 Density profiles

We model the lensing signal as a sum of contributions from the cluster-mass cross-correlation from the cluster (one-halo term) and from large-scale structure (halo–halo term). At small scales, contributions from the stars in the central galaxy are also important, but we show that their contribution is negligible for the range of scales we use for our fits ( $0.5\text{--}4 h^{-1}$  Mpc). Fig. 2 shows the relative



**Figure 2.** Observed mean lensing signals around stacked clusters in three richness bins (data points; from bottom to top):  $N_{200} = 10\text{--}11$ ,  $26\text{--}40$  and  $71\text{--}190$ , with best-fitting masses  $M_{200\bar{\rho}} = 0.65 \pm 0.30$ ,  $2.48 \pm 0.57$  and  $8.72 \pm 1.40 \times 10^{14} h^{-1} M_{\odot}$ , respectively. Also shown are the best-fitting one-halo and halo–halo profiles (dotted and long-dashed curves, respectively), the estimated stellar component (short-dashed curves) and the sum of these three (solid curves). The range of scales used for the fits is  $R = 0.5\text{--}4.0 h^{-1}$  Mpc (rightward of the vertical dashed line). For this range of scales, the stellar contribution is negligible and the halo–halo contribution is subdominant to the one-halo term. However, the halo–halo contribution becomes significant for  $R > 1 h^{-1}$  Mpc. We model the lensing signal as a sum of the one-halo and halo–halo profiles.

contributions of these three components for representative cases. The halo–halo term is significant on scales  $R > 1 h^{-1}$  Mpc, but subdominant to the one-halo term on all scales used for the fits.

The cluster mass distribution is modelled as a Navarro–Frenk–White (hereafter NFW) profile of cold dark matter haloes (Navarro, Frenk & White 1996),

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}, \quad (11)$$

defined by two parameters, the concentration  $c = r_{\text{vir}}/r_s$  and the halo mass  $M_{200\bar{\rho}}$ . While many definitions are used in the literature, here we define the virial radius  $r_{\text{vir}}$  as the radius within which the average density is equal to 200 times the mean density of the Universe  $\bar{\rho}$ , so that

$$M_{200\bar{\rho}} = \frac{4\pi}{3} r_{\text{vir}}^3 (200\bar{\rho}), \quad (12)$$

where the subscript denotes that this mass definition uses  $200\bar{\rho}$  rather than the oft-used  $200\rho_{\text{crit}}$ . The two mass definitions differ by roughly 30 per cent for typical values of concentration.

We take the concentration to be a fixed function of mass:

$$c(M_{200\bar{\rho}}) = 5.0 \left( \frac{M_{200\bar{\rho}}}{10^{14} M_{\odot}} \right)^{-0.10}. \quad (13)$$

In other words, we assume that the mass distribution only depends on a single parameter, the cluster mass  $M_{200\bar{\rho}}$ . The exponent in equation (13) matches the results of  $N$ -body simulations (Neto et al. 2007) and the normalization is determined from the observed density profiles of clusters in the maxBCG catalogue (Mandelbaum, Seljak & Hirata 2008b). We find that increasing the normalization from 5.0 to 6.0 results in a decrease in the best-fitting mass of  $\lesssim 3$  per cent for most of the mass range we consider. In particular, this means that when we use a fixed mass–concentration relation, we tend to slightly overestimate the masses of clusters with high-luminosity BCGs relative to those that have low-luminosity ones, since the former tend to have earlier formation times, and therefore, higher concentrations. This effect would lead to a small positive trend in mass with BCG luminosity at fixed richness, but we find that the induced slope (0.025) is negligible compared to the observed slopes,  $\gamma$  in Table 2. To estimate this slope, we have used a result from the simulations of Croton, Gao & White (2007, fig. 4) that indicates that a difference of  $\sim 1$  mag in BCG luminosity corresponds to a roughly 20 per cent difference in halo concentration.

The halo–halo contribution to the lensing signal is modelled using the galaxy–matter cross-power spectrum as in e.g. Mandelbaum et al. (2005b). It is proportional to the bias  $b$ , the ratio of the galaxy–matter correlation function to the matter autocorrelation function. We express the bias as a function of mass or peak height  $\nu$  (Sheth & Tormen 1999):

$$b(\nu) = 1 + \frac{a\nu - 1}{\delta_c} + \frac{2p}{\delta_c[1 + (a\nu)^p]}, \quad (14)$$

where the peak height  $\nu = \delta_c^2/\sigma^2(M)$ ,  $\delta_c = 1.686$  is the linear overdensity at which a spherical perturbation collapses at redshift  $z$  and  $\sigma(M)$  is the rms fluctuation in spheres that contain an average mass  $M$  at an initial time, extrapolated using linear theory to  $z$ ; we use  $z = 0.23$ , the median redshift of the sample. For the purposes of computing bias, we use  $a = 0.73$  and  $p = 0.15$  in order to match the results of Seljak & Warren (2004). For example, at  $z = 0.23$ , clusters of mass  $6 \times 10^{13}$  and  $6 \times 10^{14} h^{-1} M_{\odot}$  have biases of 2.2 and 5.5, respectively.

For illustration purposes, we model the stellar component by a Hernquist density profile (Hernquist 1990), which is similar to the

NFW profile in equation (11) but with an exponent of 3 instead of 2, so that it falls off faster at large scales. We estimate stellar masses from the mean  $k + e$ -corrected  $r$ -band magnitudes of BCGs in each bin, assuming a mass-to-light ratio of  $\approx 3 M_{\odot}/L_{\odot}$  (Padmanabhan et al. 2004), following Mandelbaum et al. (2006). We estimate the Hernquist profile scale radius by the measured de Vaucouleurs half-light radius multiplied by a factor of  $(\sqrt{2} - 1) \approx 0.414$ . Fig. 2 shows that the stellar contribution to the lensing signal is negligible in the range of scales used for our fits. Thus, we do not include a stellar component in our model of the cluster density profile.

### 3.4 Fits to the lensing signal

We perform fits to the lensing signal at scales  $R = 0.5\text{--}4.0 h^{-1}$  Mpc, which is around the virial radii of clusters in our sample. This choice of fitting range allows us to obtain robust mass estimates (discussed in Section 4.2). The stellar contribution to the lensing signal is negligible at these scales (see Fig. 2). We therefore model the lensing signal as a sum of one-halo and halo–halo profiles.

For any  $M_{200\bar{\rho}}$ , we can calculate the one-halo and halo–halo profiles using equations (6), (7), (11), (13) and (14). Given the observed lensing signal  $\Delta\Sigma(R)$ , we determine the best-fitting lensing profile by minimizing  $\chi^2$ , using the smooth, analytic (diagonal) covariance matrix. We determine formal  $1\sigma$  errors on the best-fitting parameter  $M_{200\bar{\rho}}$  using the distribution of parameters obtained from many bootstrap-resampled data sets. This procedure incorporates correlations between the radial bins.

Fig. 2 shows representative examples of observed lensing signals and best-fitting profiles. The halo–halo term becomes important at scales  $R > 1.0 h^{-1}$  Mpc. Neglecting to include this component would yield  $\sim 7$  per cent larger mass estimates compared to fits that include it.

### 3.5 Interpretation of the best-fitting mass

The stacked weak lensing signal that we measure is the mean signal around a set of clusters with a range of redshifts and masses. Previous studies (Mandelbaum et al. 2005b) and the quality of our fits indicate that the mean signal can be modelled as a single NFW profile to a high degree of accuracy. Moreover, Mandelbaum et al. (2005b) showed that if the mass distribution is narrow (with a typical width of less than a factor of  $\sim 5$  in mass), this model is able to determine the mean mass of the set of clusters accurately. If there is significant scatter in the mass distribution, then the cluster mass estimate falls between the distribution mean and median.

Here, we consider two kinds of stacking processes: (a) over a set of clusters that lie within a narrow range of observable properties (e.g. richness or luminosity), and (b) over a set of clusters that satisfy a threshold in a given property. For case (a), we interpret the best-fitting mass  $M_{200\bar{\rho}}$  as an estimate of the mean mass of the clusters. We use this approach to calibrate the mean relation between cluster mass and a given cluster observable property.

For case (b), while  $M_{200\bar{\rho}}$  may not be a faithful estimate of the true mean mass because of the broad mass distribution, it nevertheless allows us to assess the relative amount of scatter in a given mass–observable relation  $M = M(O)$ . Assuming a monotonic mass–observable relation without scatter, rank ordering the clusters by an observable is the same as rank ordering them by mass. Thus, selecting the top  $N$  clusters by observable would select the  $N$  most massive clusters. Moreover, if there are two tracers with no scatter they would produce the same sample, even if the functional forms  $M(O)$  differ. The effect of scatter is to bring in clusters with lower

mass, which would lower the mean weak lensing signal around the stacked clusters and the corresponding best-fitting mass. Thus, a higher best-fitting mass obtained from a given observable threshold at fixed number density indicates a lower scatter in the corresponding mass–observable relation. This analysis has been worked out explicitly for the case of lognormal scatter in Mandelbaum & Seljak (2007).

Finally, we note that the mass that we measure from the weak lensing signal around stacked clusters may differ from other mass definitions, such as from spherical overdensity, because the presence of substructure and filaments introduce scatter between the two quantities. This scatter may be large if only a small number of clusters is stacked and one should quantify this with simulations, which is beyond the scope of this paper. Here we simply take lensing-defined mass as the mass definition.

#### 4 TESTS OF SYSTEMATICS

In this section, we discuss various tests of systematics associated with the cluster lens catalogue, including photometric redshift errors (Section 4.1) and offsets from the cluster centre (Section 4.2), and with the weak lensing source galaxy catalogue, including lensing calibration (Section 4.3) and contamination from intrinsic alignments (Section 4.4).

##### 4.1 Cluster photometric redshift errors

Koester et al. (2007a,b) assessed the accuracy of photometric redshifts (photo- $z$ ) measurements in the maxBCG catalogue by comparing them with measured spectroscopic redshifts (available for  $\sim 40$  per cent of the sample). They found that the photo- $z$  dispersion  $\sqrt{((z_{\text{photo}} - z_{\text{spec}})^2)} \approx 0.01$ , and is essentially independent of redshift for the range covered by the sample  $0.1 < z < 0.3$ . In this section, we investigate the effect of photometric redshift errors on our results.

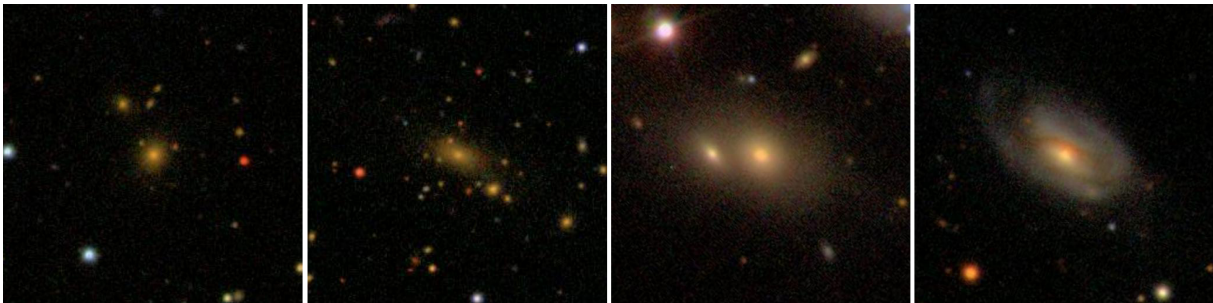
Cluster photo- $z$  errors affect both the measurement of cluster properties and the computation of the lensing signal. The reported luminosities in the maxBCG catalogue were converted from apparent magnitudes using distances from photometric redshifts, so an overestimate in the redshift would result in a corresponding overestimate in the reported luminosities. In addition,  $L_{200}$  and  $N_{200}$

would be affected because the change in both  $r_{200}$  and  $L_*$  would change which galaxies would be considered cluster members by the maxBCG cluster finder.

The lensing signal computation is affected in three ways: first, the lensing signal calibration depends on the lens-source geometry, and therefore on the assumed value for the cluster redshift; second, the conversion from angular distance to transverse separation depends on photometric redshift; third, the change in the observed property (luminosity or richness) would change the bin in which a given cluster belongs. Generically, we expect the first two errors to cancel out at some level for any given cluster: e.g. if the lens photo- $z$  is overestimated, then  $\Sigma_c$  and hence  $\Delta\Sigma$  are underestimated, but due to the error in the angular diameter distance we also overestimate the transverse separation  $R$ , which increases the signal at fixed transverse separation.

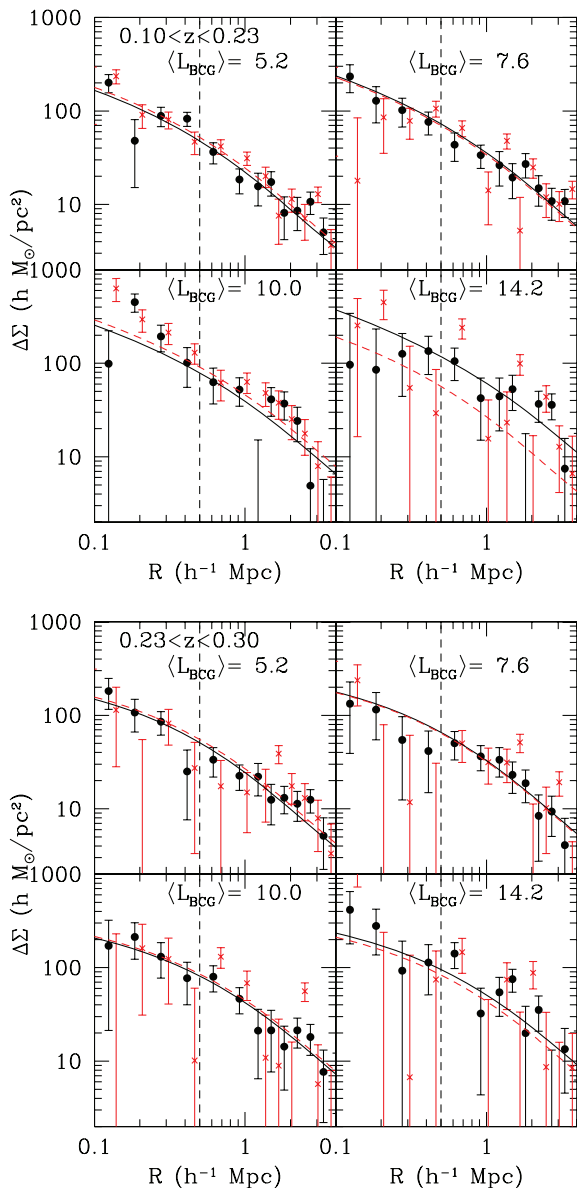
Out of 5423 BCGs (43 per cent of the sample) with measured spectroscopic redshifts, 131 galaxies (2.4 per cent) have severe photo- $z$  errors, corresponding to differences in distance moduli larger than 0.5 mag. The incidence of photo- $z$  errors is much higher for BCGs with the highest reported luminosities, as expected since these extremely luminous objects are rare and a few photo- $z$  failures on less luminous objects can lead to a large fractional contamination. Of the 49 objects with reported  $L_{\text{BCG}} > 16 \times 10^{10} h^{-2} L_{\odot}$ , 12 per cent (six objects) have severe photo- $z$  errors. We show some examples in Fig. 3.

To test for the effect of lens photo- $z$  errors on our weak lensing analysis, we divide the 5423 clusters (with measured spectroscopic redshifts) into two-redshift bins,  $0.10 < z < 0.23$  and  $0.23 < z < 0.30$ , and five bins in BCG luminosity. We calculate their lensing signal in two ways: (i) using photometric redshifts and the reported BCG luminosities, and (ii) using spectroscopic redshifts and BCG luminosities scaled to the measured spectroscopic redshifts. Fig. 4 compares the measured lensing signals for the two cases. Note that the binning assignment is different in the two cases because of the difference in assumed BCG luminosities. The lensing signals for the highest  $L_{\text{BCG}}$  bins tend to be noisier for case (ii) because these bins include very few objects once we correct for photo- $z$  errors. Within the error bars, we find no systematic difference between the two cases. Therefore, for our main analysis, we use the full cluster sample and the reported photometric redshifts and luminosities.



**Figure 3.** Examples of galaxies identified as BCGs in the maxBCG catalogue with reported  $r$  band luminosities ( $k$ -corrected to  $z = 0.25$ ) of  $L_{\text{BCG}} > 16 \times 10^{10} h^{-2} L_{\odot}$ ; these images are taken from the SDSS DR6 Skyserver. From left to right: (a) SDSS J085540.19–003257.2 ( $z = 0.271$ ); (b) SDSS J212939.95+000521.1 ( $z = 0.234$ ); (c) SDSS J085458.90+490832.3 ( $z = 0.052$ ) and (d) SDSS J102246.44+483813.6 ( $z = 0.050$ ). Objects (a) and (b) have accurate photometric redshifts. These fields show a dominant cD galaxy (the BCG) surrounded by other red galaxies, typical of clusters in the catalogue. Object (b)’s photo- $z$  was successful despite the presence of [O II], H $\alpha$  and [N II] emission lines that are unusual for a BCG. Objects (c) and (d) have severely overestimated photometric redshifts (0.127 and 0.138, respectively). For (c), the error in the photometric redshift is probably due to the difficulty in deblending the overlapping galaxies. Object (d) seems to be a face-on spiral galaxy with thick dust lanes, which was mistaken for a BCG. We estimate the contamination of the catalogue from such objects to be  $< 2.4$  per cent based on the incidence of very large errors in photometric redshift for those objects with spectra.





**Figure 4.** Test of systematics for the effect of cluster photo- $z$  errors. Clusters are divided into two ranges in redshift (upper and lower panels) and five bins in BCG luminosity (the four highest luminosity bins are shown above, with the mean  $L_{\text{BCG}}$  listed in units of  $10^{10} h^{-2} L_{\odot}$ ). The stacked weak lensing signal around clusters in each bin is calculated in two ways: (i) using photometric redshifts and the reported BCG luminosities (filled circles/black), and (ii) using spectroscopic redshifts and BCG luminosities scaled to the spectroscopic redshifts (crosses/red). The best-fitting one-halo + halo-halo profiles are shown in each case (solid and dashed curves, respectively). The data points have been slightly offset horizontally for clarity. The vertical dashed line marks the range of scales used in our fits  $R = 0.5\text{--}4.0 h^{-1} \text{Mpc}$ .

#### 4.2 Offsets from cluster centre

BCGs are generally expected to lie at or near the centres of clusters, where the potential well is the deepest, but this is not always observed. Using  $N$ -body mock galaxy catalogues, Johnston et al. (2007) found that only  $\sim 60\text{--}80$  per cent of the BCGs identified by the maxBCG cluster finder are located near the halo centre, and that the offsets of the rest of the BCGs can be modelled as a projected Gaussian distribution with a width of  $0.42 h^{-1} \text{Mpc}$ . These results

must however be seen in light of the fact that the haloes in the simulations do not correspond exactly to clusters in the data.

For our weak lensing measurements, we define the location of the BCG to be the centre of the cluster, but take steps to reduce the effect of offsets from the cluster centre on the mass estimates. Fits for the concentration from the lensing profiles of clusters in the maxBCG catalogue show that the effect of miscentring is important (leading to shallower derived concentrations and lower masses) when fits use transverse separations  $R < 0.5 h^{-1} \text{Mpc}$ , but not when the fits are restricted to  $R > 0.5 h^{-1} \text{Mpc}$  (Mandelbaum et al. in preparation). Fitting from 0.2 instead of  $0.5 h^{-1} \text{Mpc}$  tended to suppress the concentrations at the  $\sim 20$  per cent level. Therefore, we restrict the fitting range to  $R > 0.5 h^{-1} \text{Mpc}$  in this work.

#### 4.3 Lensing calibration

Lensing calibration systematics due to the source sample include source redshift uncertainties, shear calibration and stellar contamination. Since these effects do not vary with scale, they could only change the overall normalization in the derived mass–observable relation.

Comparison with spectroscopy from DEEP2 and zCOSMOS showed that to account for photometric redshift errors in the source redshifts, one has to multiply the signal by a calibration factor of  $0.97 \pm 0.02$  for the  $0.10 < z < 0.23$  sample, and  $0.98 \pm 0.04$  for the  $0.23 < z < 0.30$  sample (Mandelbaum et al. 2008a). Stellar contamination in the source catalogue, which would decrease the lensing signal, is tightly constrained to less than 1 per cent using COSMOS data (Mandelbaum et al. 2008a). Taking this into account, the calibration factors become  $0.98 \pm 0.02$  for the  $0.10 < z < 0.23$  sample and  $0.99 \pm 0.04$  for the  $0.23 < z < 0.30$  sample. Since these are within  $1\sigma$  of unity and are much smaller than the statistical error bars on the weak lensing signal, we choose not to apply these correction factors in this work. A conservative estimate of the total calibration uncertainty, including both these two effects and the shear calibration bias, is 8 per cent at the  $1\sigma$  level (M05). This can be taken into account by adding it in quadrature to the statistical error on the mass determinations.

#### 4.4 Intrinsic alignments

The important intrinsic alignment effect for cluster–galaxy lensing is the alignment between the intrinsic ellipticity of a galaxy and the direction to nearby cluster BCGs. This effect comes into play because we necessarily include some physically associated pairs (i.e. pairs of lenses and ‘sources’ that are really part of the same local structure); if these sources preferentially align tangentially or radially relative to the lens, they would provide an additive bias to the lensing signal.

The effect of intrinsic alignments on the lensing profile is more important at small transverse separations, since close physically associated pairs tend to be more aligned. Using the same source catalogue used here, Mandelbaum et al. (2006) found that intrinsic alignment contamination of the lensing signal for LRGs is only important at scales  $R < 0.1 h^{-1} \text{Mpc}$ , given our procedures for removing physically associated galaxies from the source sample. Since many cluster BCGs are also in this LRG sample, this result is relevant for the current work. Agustsson & Brainerd (2006) measured the mean tangential shear of spectroscopically determined satellites and found a tendency for satellites to align radially towards central

**Table 1.** Individual bins of clusters rank ordered according to  $N_{200}$  (cluster richness in red galaxies),  $L_{200}$  (cluster luminosity in red galaxies) and  $L_{\text{BCG}}$  (luminosity of the BCG). The number of clusters in each bin, their range of properties, mean  $N_{200}$ ,  $L_{200}$ ,  $L_{\text{BCG}}$  and the estimated mean cluster mass  $M_{200\bar{\rho}}$  are listed. The  $1\sigma$  errors on the mass estimates are derived from 2500 bootstrap-resampled data sets.

Number	Range	$\langle N_{200} \rangle$	$\langle L_{200} \rangle$ ( $10^{10} h^{-2} L_{\odot}$ )	$\langle L_{\text{BCG}} \rangle$ ( $10^{10} h^{-2} L_{\odot}$ )	$M_{200\bar{\rho}}$ ( $10^{14} h^{-1} M_{\odot}$ )
Bins in $N_{200}$					
4091	10–11	10.43	16.29	4.67	$0.65 \pm 0.30$
5164	12–17	13.88	21.67	5.27	$0.96 \pm 0.32$
2055	18–25	20.78	32.38	6.21	$1.43 \pm 0.42$
933	26–40	31.06	48.40	7.05	$2.48 \pm 0.57$
320	41–70	50.06	76.64	8.24	$3.96 \pm 0.77$
49	71–190	89.86	140.87	10.45	$8.72 \pm 1.40$
Bins in $L_{200}$					
4091	6.63–17.56	11.29	14.17	3.51	$0.56 \pm 0.29$
5164	17.56–28.51	13.89	22.22	5.57	$1.12 \pm 0.33$
2055	28.51–41.76	20.01	33.73	7.07	$1.46 \pm 0.43$
933	41.76–64.46	29.90	50.37	8.10	$2.47 \pm 0.59$
320	64.46–115.55	52.95	88.44	9.83	$4.13 \pm 0.82$
49	115.55–274.71	85.14	146.91	12.44	$10.57 \pm 1.44$
Bins in $L_{\text{BCG}}$					
4091	0.66–3.95	13.47	16.84	2.95	$0.71 \pm 0.31$
5164	3.95–6.56	16.13	24.88	5.15	$1.00 \pm 0.32$
2055	6.56–8.90	18.56	32.35	7.57	$1.69 \pm 0.42$
933	8.90–11.73	21.56	40.73	10.02	$2.40 \pm 0.55$
320	11.74–16.68	25.31	50.45	13.40	$3.28 \pm 0.83$
49	16.68–29.05	34.78	74.61	19.74	$6.77 \pm 1.57$

galaxies over the range  $7 < R < 50 h^{-1}$  kpc. Since we have used photometric redshift estimates to separate source galaxies from lenses, and we only use the lensing signal data in the range  $R = 0.5\text{--}4.0 h^{-1}$  Mpc, our results should not be affected by contamination from intrinsic alignments.

None the less, we present constraints on intrinsic alignments contamination of the lensing signal for the transverse separations used here. To do so, we use the formalism and results on intrinsic alignments in LRG lenses with our source catalogue from Mandelbaum et al. (2006). For the ‘bright’ lens sample in that work (corresponding to halo masses of  $\sim 7 \times 10^{13} h^{-1} M_{\odot}$ ), the intrinsic alignment signal was not detected at  $0.5\text{--}0.6 h^{-1}$  Mpc, and was constrained to contaminate the lensing signal by  $< 3 h M_{\odot} \text{pc}^{-2}$  at 95 per cent CL. This constraint is in fact conservative, since there is reason to believe that the sample of red ‘source’ galaxies that we used to place the constraint is more strongly intrinsically aligned than the general galaxy population. Given that the typical lensing signal for the maxBCG clusters on these scales is more than 20 times larger than this conservative bound, we conclude that it is not an important contaminant for this work. For larger scales, it is also not important, since the effect is expected to decrease with transverse separation, as does the fraction of physically associated ‘source’ galaxies.

## 5 RESULTS

In this section, we calibrate and assess the scatter in the relation between several cluster properties and cluster mass, as outlined in Section 3.5. In Section 5.1, we consider three main observable parameters — cluster richness in red galaxies  $N_{200}$ , cluster luminosity in red galaxies  $L_{200}$  and luminosity of the BCG  $L_{\text{BCG}}$ . In Section 5.2, we consider power-law combinations of  $N_{200}$  and  $L_{200}$  with  $L_{\text{BCG}}$ , with the aim of finding improved mass tracers for galaxy clusters.

### 5.1 $N_{200}$ , $L_{200}$ and $L_{\text{BCG}}$ as mass tracers

#### 5.1.1 Calibration of mean mass–observable relations

We begin by calibrating the mean relation between cluster mass and three cluster properties:  $N_{200}$  (cluster richness in red galaxies),  $L_{200}$  (cluster luminosity in red galaxies) and  $L_{\text{BCG}}$  (luminosity of the BCG). We rank order the clusters in each property and divide them into six individual bins, keeping the same number of clusters in each bin (Table 1). We measure the stacked weak lensing signal around clusters in each bin, and determine the best-fitting mass  $M_{200\bar{\rho}}$  using the procedure described in Section 3.4. We do this analysis for the full redshift range  $0.1 < z < 0.3$ . The results are shown in Table 1 and Fig. 5.

The scaling of mean cluster mass with  $N_{200}$ ,  $L_{200}$  and  $L_{\text{BCG}}$  are well described by power laws. To determine the normalization and slope in these relations, we minimize  $\chi^2$  simultaneously for the six sets of measured lensing signals. We determine uncertainties on the parameters by repeating the fitting procedure for the 2500 bootstrap-resampled data sets. The best-fitting relations are

$$M_{14}(N_{200}) = (1.42 \pm 0.08)(N_{200}/20)^{1.16 \pm 0.09}, \quad (15a)$$

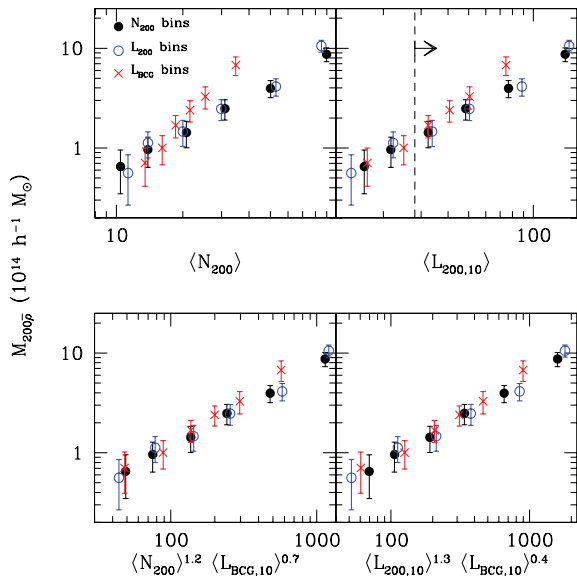
$$M_{14}(L_{200}) = (1.76 \pm 0.17)(L_{200,10}/40)^{1.40 \pm 0.19}, \quad (15b)$$

$$M_{14}(L_{\text{BCG}}) = (1.07 \pm 0.07)(L_{\text{BCG},10}/5)^{1.10 \pm 0.13}, \quad (15c)$$

where  $M_{14}$  is  $M_{200\bar{\rho}}$  in units of  $10^{14} h^{-1} M_{\odot}$ , and  $L_{200,10}$  and  $L_{\text{BCG},10}$  are in units of  $10^{10} h^{-2} L_{\odot}$ . From the covariance matrix of the best-fitting parameters, we find that the slope and normalization are uncorrelated for the  $N_{200}$  relation, and anticorrelated at the  $\sim 50\text{--}60$  per cent level for the  $L_{200}$  and  $L_{\text{BCG}}$  relations.

The mass– $L_{200}$  relation equation (15b) is derived using only clusters with  $L_{200,10} > 28$ , where the sample is complete in  $L_{200}$  (Fig. 1). Thus, it is not affected by the selection effect introduced by the





**Figure 5.** Scaling of cluster mass  $M_{200\bar{p}}$  with various cluster mass tracers. Masses are determined from the stacked weak lensing signal around clusters in individual bins in  $N_{200}$  (filled circles),  $L_{200}$  (open circles) and  $L_{BCG}$  (crosses). The upper panels show the scaling of cluster mass with mean parameters ( $N_{200}$ ) and ( $L_{200}$ ); the dashed line on the upper right-hand panel shows the  $L_{200}$  value above which the sample is complete. The lower panels show the scaling of cluster mass with a combination of the mean parameters, with exponents taken from Table 2. The tighter scaling of cluster mass with the combined tracers, regardless of whether  $N_{200}$ ,  $L_{200}$  or  $L_{BCG}$  is used for the binning, suggests that these combined quantities trace mass more faithfully than either  $N_{200}$  or  $L_{200}$  taken alone.

$N_{200} \geq 10$  cut. If we include the full sample in the analysis, we find  $M_{200\bar{p},14}(L_{200}) = (1.96 \pm 0.11)(L_{200,10}/40)^{1.13 \pm 0.08}$ . The derived slope is shallower than that in equation (15b), consistent with the effect of missing lower mass clusters at low luminosity. The sample is incomplete at all values of  $L_{BCG}$ , so the shallow slope of equation (15c) is partly due to this selection effect; it should therefore be kept in mind that this relation is valid only for the richness-selected sample. For a sample complete in  $L_{BCG}$ , the slope would likely be steeper and the mean mass at low  $L_{BCG}$  would be lower (e.g. Lin & Mohr 2004 find that BCG luminosity scales with halo mass with an exponent of  $0.33 \pm 0.06$ , which implies a much steeper relation than that in equation 15c).

The slopes we find for the scaling of mass with  $N_{200}$  and  $L_{200}$  are roughly consistent with the results of Johnston et al. (2007) (tables 10 and 11 list 1.30 and 1.25, respectively, for the mass definition closest to ours  $M_{180b}$ ), though it should be noted that they use additional clusters (with  $N_{200} < 10$ ). Our normalization for the mass–richness relation is higher by  $\sim 18$  per cent, but this can be explained by our use of different methods for determining photometric redshifts of source galaxies, which leads to different amounts of bias in the estimated lensing signals. Mandelbaum et al. (2008a) tested for calibration bias in the lensing signal due to use of different methods of determining source redshifts, using source galaxies with spectroscopy from zCOSMOS and DEEP2 as a reference. They concluded that for the maxBCG lens redshift distribution and the methods used here for determining source redshifts, the calibration bias in the lensing signal is small (consistent with zero within our quoted systematic error), whereas for the SDSS DR6 neural net photometric redshifts used by Johnston et al. (2007), it is approximately  $-18$  per cent.

### 5.1.2 Scatter in the mass–observable relations

In this section, we assess the relative amount of scatter in the various mass–observable relations derived above. As discussed in Section 3.5, an observable threshold that yields a higher best-fitting mass has a mass relation with lower scatter. We define thresholds corresponding to cluster comoving number densities of  $\bar{n} = \{20, 10, 5, 2.5\} \times 10^{-7} (h^{-1} \text{Mpc})^{-3}$ . This translates to taking the top  $\{384, 192, 96, 48\}$  clusters for the  $0.10 < z < 0.23$  sample, and the top  $\{456, 233, 116, 58\}$  clusters for the  $0.23 < z < 0.30$  sample. We measure the stacked weak lensing signal for each threshold in  $N_{200}$ ,  $L_{200}$  and  $L_{BCG}$  and compare the derived best-fitting masses in Fig. 6.

Out of the three parameters considered, we find that  $L_{BCG}$  is the poorest tracer of cluster mass. This statement is robust to the selection effect introduced by the  $N_{200} \geq 10$  cut, since the inclusion of poorer, low-mass clusters into the threshold would further decrease the lensing signal. We note that the scatter in the mass relation is a combination of intrinsic and observational scatter, and the contribution from the latter may be significant because of the difficulty in measuring accurate BCG luminosities. For example, systematic errors from sky subtraction are important for BCGs because they have large, diffuse envelopes, and deblending issues are also important because BCGs are located in dense environments.

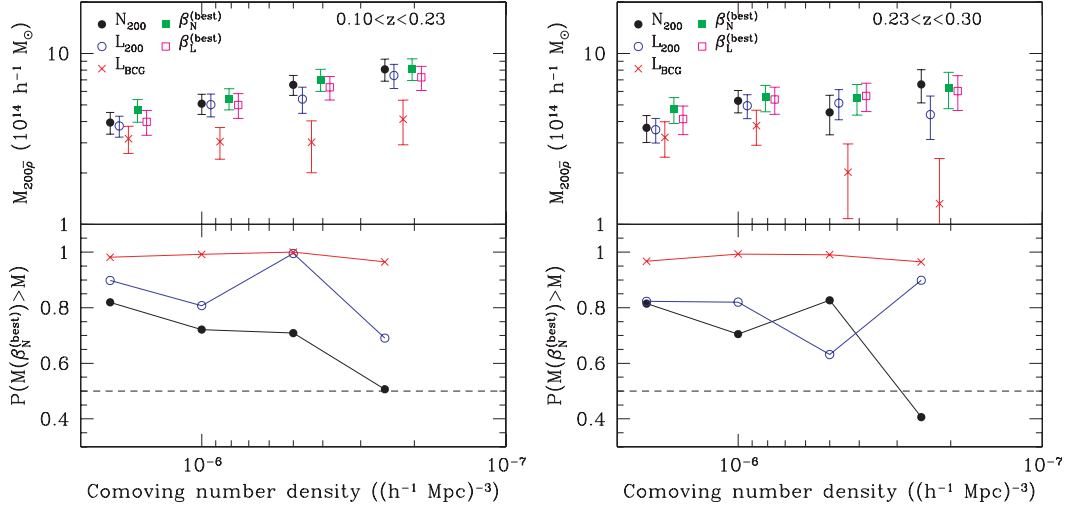
Fig. 6 shows that the best-fitting masses  $M_{200\bar{p}}$  for clusters in  $N_{200}$  and  $L_{200}$  thresholds at the same number density tend to be comparable. However, about 70–80 per cent of the clusters selected by the  $N_{200}$  threshold is also selected by the corresponding  $L_{200}$  threshold. Thus, the error bars in these data points are tightly correlated, and the differences in the masses are more significant than what one would estimate by eye. We therefore assess the statistical significance of these differences using results from many bootstrap data sets. We find that the  $N_{200}$  threshold yields a higher mass than the  $L_{200}$  threshold in  $\{72, 45, 93, 68\}$  per cent of the cases (for the  $0.10 < z < 0.23$  sample), and for  $\{37, 68, 27$  and  $90\}$  per cent of the cases (for the  $0.23 < z < 0.30$  sample), in order of decreasing number density. These high values indicate that  $N_{200}$  picks out more of the most massive clusters most of the time, and therefore has smaller scatter than  $L_{200}$  at this range of masses. Fig. 6 also shows for comparison the masses obtained from thresholds in the combined mass tracers, which we discuss in Section 5.2.2.

## 5.2 Combined mass tracers

In this section, we consider whether adding information from BCG luminosity can provide improved estimates of cluster masses. Our previous analysis shows that  $L_{BCG}$  by itself does not trace mass as well as  $N_{200}$  or  $L_{200}$ . However, the scatter in  $L_{BCG}$  at a fixed  $N_{200}$  or  $L_{200}$  suggests that there may be residual scaling of mass with  $L_{BCG}$ . Fig. 5 shows that the scaling of mass with a combination of  $N_{200}$  (or  $L_{200}$ ) and  $L_{BCG}$  (lower panels) is tighter than that with  $N_{200}$  or  $L_{200}$  taken alone (upper panels), regardless of the parameter used for binning the clusters. This suggests that the additional information in  $L_{BCG}$  reduces the scatter in the mass relation. Here, we consider power-law combinations of  $L_{BCG}$  with  $N_{200}$  (or  $L_{200}$ ) as mass tracers. We calibrate the mass relation in Section 5.2.1 and assess the scatter in this relation in Section 5.2.2.

### 5.2.1 Calibration of mean mass–observable relations

To consider the scaling of mass with both  $N_{200}$  and  $L_{BCG}$  simultaneously, we divide the cluster sample into five bins in  $N_{200}$  and further split these bins in  $L_{BCG}$ , for a total of 22 bins in the two-dimensional



**Figure 6.** Comparison of the relative amount of scatter in the various mass tracers. Higher values of the best-fitting cluster mass,  $M_{200\bar{\rho}}$ , indicate a lower scatter in the mass relation. Upper panels: cluster masses  $M_{200\bar{\rho}}$  from stacked weak lensing signals around clusters satisfying thresholds in the various tracers, for comoving number densities  $\bar{n} = \{20, 10, 5 \text{ and } 2.5\} \times 10^{-7} (h^{-1} \text{ Mpc})^{-3}$ . We compare the mass tracers  $N_{200}$ ,  $L_{200}$ ,  $L_{BCG}$ ,  $N_{200}L_{BCG}^{\beta_N^{(best)}}$  and  $L_{200}L_{BCG}^{\beta_L^{(best)}}$ , the combined tracers that yield the highest masses at each number density (Section 5.2.2). Left- and right-hand plots are for the two-redshift ranges; the data points in the figure are slightly offset horizontally for clarity. The  $1\sigma$  error bars shown here are tightly correlated, so the differences in the masses are more significant than apparent by eye. Lower panels: probability that the  $\beta_N^{(best)}$  tracer yields a higher mass than  $N_{200}$  (filled circles/black),  $L_{200}$  (open circles/blue) or  $L_{BCG}$  (crosses/red) taken alone, defined to be the percentage of cases among 1000 bootstrap-resampled data sets. High values of this quantity suggest that the combined tracers have comparable or lower scatter than either  $N_{200}$  or  $L_{200}$  taken alone, for this range of cluster abundances.

**Table 2.** Best-fitting parameters for the scaling of cluster mass with  $N_{200}$  and  $L_{BCG}$  (equation 16a), and with  $L_{200}$  and  $L_{BCG}$  (equation 16b). The  $1\sigma$  errors and correlation coefficients  $r$  in the table are derived from 1000 bootstrap-resampled data sets.

	$M_N^0$	$\alpha_N$	$\gamma_N$	$r(M_N^0, \alpha_N)$	$r(M_N^0, \gamma_N)$	$r(\alpha_N, \gamma_N)$
$0.10 < z < 0.23$	$1.27 \pm 0.08$	$1.20 \pm 0.09$	$0.71 \pm 0.14$	$-0.24$	$-0.40$	$0.03$
$0.23 < z < 0.30$	$1.57 \pm 0.14$	$1.12 \pm 0.15$	$0.34 \pm 0.24$	$-0.07$	$-0.18$	$0.09$
	$M_L^0$	$\alpha_L$	$\gamma_L$	$r(M_L^0, \alpha_L)$	$r(M_L^0, \gamma_L)$	$r(\alpha_L, \gamma_L)$
$0.10 < z < 0.23$	$1.81 \pm 0.15$	$1.27 \pm 0.17$	$0.40 \pm 0.23$	$-0.34$	$-0.17$	$0.34$
$0.23 < z < 0.30$	$1.76 \pm 0.22$	$1.30 \pm 0.29$	$0.26 \pm 0.41$	$-0.42$	$-0.35$	$0.41$

$N_{200}$ – $L_{BCG}$  space. We make a similar division in  $L_{200}$ – $L_{BCG}$  space for clusters with  $L_{200} > 28$  (for which the sample is complete) resulting in nine bins. We then measure the stacked weak lensing signal around clusters in each bin. We do this analysis for two-redshift ranges,  $0.10 < z < 0.23$  and  $0.23 < z < 0.30$ .

We parametrize the scaling of mass as a power law in  $N_{200}$  (or  $L_{200}$ ) with an additional scaling with  $L_{BCG}$  at fixed  $N_{200}$  (or  $L_{200}$ ):

$$M_{14}(N_{200}, L_{BCG}) = M_N^0 (N_{200}/20)^{\alpha_N} \left( L_{BCG}/\bar{L}_{BCG}^{(N)} \right)^{\gamma_N}, \quad (16a)$$

$$M_{14}(L_{200}, L_{BCG}) = M_L^0 (L_{200,10}/40)^{\alpha_L} \left( L_{BCG}/\bar{L}_{BCG}^{(L)} \right)^{\gamma_L}, \quad (16b)$$

where  $M_{14}$  is  $M_{200\bar{\rho}}$  in units of  $10^{14} h^{-1} M_{\odot}$ ,  $L_{200,10}$  is the cluster luminosity in units of  $10^{10} h^{-2} L_{\odot}$  and the BCG luminosity dependence is pivoted at the mean  $L_{BCG}$  at the given  $N_{200}$  (or  $L_{200}$ ). Parametrizing this mean relation as a power law, the best-fitting relations are

$$\bar{L}_{BCG}^{(N)} \equiv \bar{L}_{BCG}(N_{200}) = a_N N_{200}^{b_N}, \quad (17a)$$

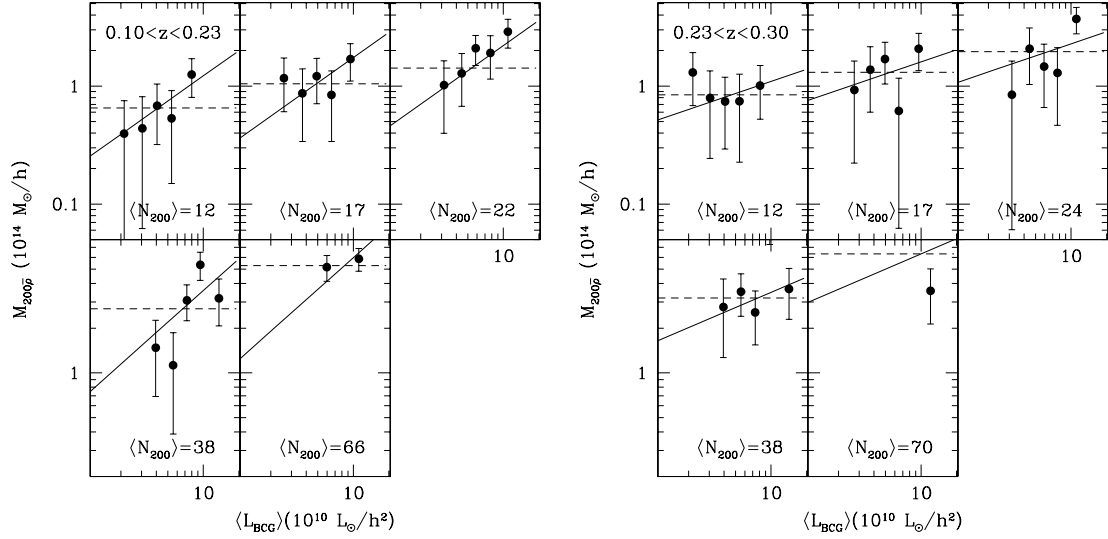
$$\bar{L}_{BCG}^{(L)} \equiv \bar{L}_{BCG}(L_{200}) = a_L L_{200,10}^{b_L}, \quad (17b)$$

where  $a_N = (1.54, 1.64) \times 10^{10} h^{-2} L_{\odot}$ ,  $b_N = (0.41, 0.43)$  and  $a_L = (7.77, 7.92) \times 10^{10} h^{-2} L_{\odot}$ ,  $b_L = (0.67, 0.66)$  for the two-redshift ranges ( $0.10 < z < 0.23$ ,  $0.23 < z < 0.30$ ). Combining equations

(16) and (17) gives a cluster mass estimate for any cluster with measured  $N_{200}$  (or  $L_{200}$ ) and  $L_{BCG}$ .

We derive best-fitting parameters  $M^0$ ,  $\alpha$  and  $\gamma$  (shown in Table 2) by minimizing  $\chi^2$  simultaneously for the set of measured lensing signals. To obtain confidence intervals on these fits, we repeat the fitting procedure for the 1000 bootstrap-resampled data sets, using the analytical covariance matrix (rather than the full bootstrap covariance matrix, which is too noisy to use to weight the fits). The bootstrap-resampled data sets yield Gaussian probability distributions in  $M^0$ ,  $\alpha$  and  $\gamma$ ; the  $1\sigma$  errors and correlation coefficients for these parameters are also shown in Table 2.

Comparison of the best-fitting mass relations for the two-redshift ranges suggests an increase in cluster mass with redshift at fixed richness. Using the 1000 bootstrap-resampled data sets, we find that the mass normalization for the higher redshift sample is larger than that for the lower redshift sample at  $\sim 97$  per cent CL. We note however that the redshift dependence may result from systematic effects due to photo- $z$  errors, which have a larger dispersion at lower redshifts, and/or from evolution in the richness estimator  $N_{200}$  (e.g. due to an incorrect assumption of the evolution of the luminosity cut  $0.4L_*$ ). Disentangling these effects from ‘true’ evolution requires a more careful control of the systematics. Hints of an increase in cluster mass with redshift at fixed  $N_{200}$  have been found in



**Figure 7.** Scaling of mean cluster mass  $M_{200\bar{p}}$  with  $L_{\text{BCG}}$  within narrow bins in  $N_{200}$ . The best-fitting mass relation  $M(N_{200}, L_{\text{BCG}})$  (given by equation 16a) are shown in solid lines. The mass relation without the  $L_{\text{BCG}}$  dependence (i.e. with  $\gamma_N = 0$ ) are shown in dashed lines. We find residual scaling with  $\gamma_N = 0.71 \pm 0.14$  in the lower redshift sample (left), and with  $\gamma_N = 0.34 \pm 0.24$  in the higher redshift sample (right).

measurements of X-ray luminosities (Rykoff et al. 2008) and velocity dispersions (Becker et al. 2007) of clusters in the maxBCG catalogue, but no evidence of evolution had been detected in a previous analysis of their weak lensing signal (Sheldon et al. 2007a).

Figs 7 and 8 show the scaling of cluster mass  $M_{200\bar{p}}$  with  $L_{\text{BCG}}$  within narrow bins in  $N_{200}$  and  $L_{200}$ . These scalings are traced well by the best-fitting relations equations (16a) and (16b). At fixed  $N_{200}$ , residual scaling with  $L_{\text{BCG}}$  is seen with  $\gamma_N = 0.71 \pm 0.14$  ( $\sim 5\sigma$ ) for the lower redshift sample, and with  $\gamma_N = 0.34 \pm 0.24$

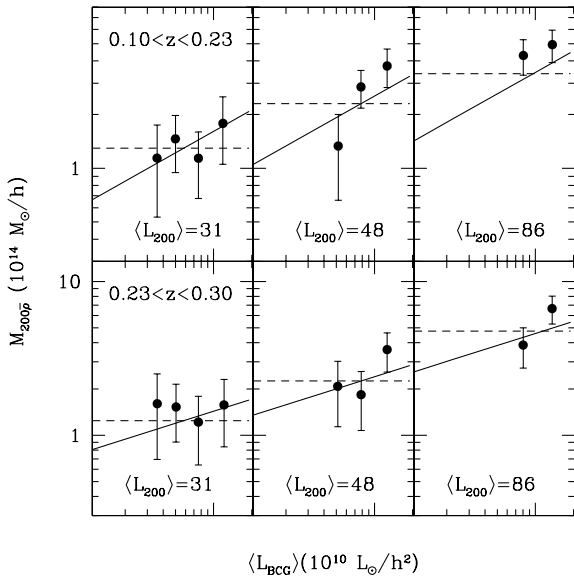
for the higher redshift sample. At fixed  $L_{200}$ , we find  $\gamma_L = 0.40 \pm 0.23$  ( $\sim 2\sigma$ ) for the lower redshift sample, and  $\gamma_L = 0.26 \pm 0.41$  for the higher redshift sample. Constraints for the scaling with  $L_{200}$  are relatively weaker because of the luminosity cut applied to the complete sample, which reduces the number of clusters to about one-third of the full sample. The scaling parameters are less well constrained for the higher redshift range because there are fewer lensed sources behind the high-redshift clusters.

### 5.2.2 Scatter in the mass–observable relations

We turn to the question of whether exploiting information about BCG luminosity in addition to either  $N_{200}$  or  $L_{200}$  reduces the scatter in the mass relation. Similar to Section 5.1.2, we rank clusters according to  $L_{200}L_{\text{BCG}}^{\beta_L}$  and  $N_{200}L_{\text{BCG}}^{\beta_N}$  and take the top  $N$  clusters to define thresholds with comoving number densities  $\bar{n} = \{20, 10, 5, 2.5\} \times 10^{-7} (h^{-1} \text{Mpc})^{-3}$ . We explore a set of values of exponents,  $\beta_N = \{0.25, 0.50, 0.75, 1.0, 1.5, 2.0\}$  and  $\beta_L = \{0.2, 0.4, 0.6, 0.8, 1.0\}$ , to find the one that maximizes  $M_{200\bar{p}}$ , or equivalently, minimizes the scatter in the mass–observable relation. We do this analysis for two-redshift ranges,  $0.10 < z < 0.23$  and  $0.23 < z < 0.30$ .

The exponents that yield the highest masses at each number density are (from highest to lowest number density):  $\beta_N^{(\text{best})} = \{1.5, 1.5, 0.25, 0.25\}$  and  $\beta_L^{(\text{best})} = 0.4$  for the lower redshift sample and  $\beta_N^{(\text{best})} = \{1.5, 1.0, 0.5, 0.5\}$  and  $\beta_L^{(\text{best})} = \{0.8, 0.6, 0.6, 0.2\}$  for the higher redshift sample. In general, the tracer with the minimal scatter is a combination of  $N_{200}$  and  $L_{\text{BCG}}$  [except for  $\bar{n} = 2.5 \times 10^{-7} (h^{-1} \text{Mpc})^{-3}$  in the higher redshift sample, where  $N_{200}$  alone yields the highest mass; one possible reason for this trend is that at higher redshifts, the large  $L_{\text{BCG}}$  bins are more likely to be contaminated by low-luminosity objects for which the photo- $z$  has been overestimated (Section 4.1)].

The error bars are tightly correlated between the combined and individual tracers, as well as between different  $\beta_N$  or  $\beta_L$  values, because a significant fraction of the clusters that satisfy the different thresholds are the same. For example, for the lowest number density bin  $\bar{n} = 2.5 \times 10^{-7} (h^{-1} \text{Mpc})^{-3}$ , there is substantial overlap



**Figure 8.** Scaling of mean cluster mass  $M_{200\bar{p}}$  with  $L_{\text{BCG}}$  within narrow bins in  $L_{200}$ . The mean  $L_{200}$  in each bin is shown in units of  $10^{10} h^{-2} L_{\odot}$ ; we restrict this analysis to  $L_{200} > 28 \times 10^{10} h^{-2} L_{\odot}$ , for which the sample is complete. We find residual scaling of  $M_{200\bar{p}}$  with  $L_{\text{BCG}}$  at fixed  $L_{200}$  in the lower redshift sample (upper panels), with  $\gamma_L = 0.40 \pm 0.23$  ( $\sim 2\sigma$ ) and no significant evidence for residual scaling in the higher redshift sample (lower panels), with  $\gamma_L = 0.26 \pm 0.41$ .

between clusters satisfying the threshold in  $\beta_N^{(\text{best})}$  and in  $N_{200}$  (94 per cent),  $L_{200}$  (83 per cent) and  $L_{\text{BCG}}$  (26 per cent). We assess the significance of the differences in the masses using the 1000 bootstrap-resampled data sets. We find that the combined tracers with exponents  $\beta_N^{(\text{best})}$  yield higher masses than  $N_{200}$ ,  $L_{200}$  or  $L_{\text{BCG}}$  in the majority of cases ( $>50$  per cent), for the range of number densities we consider.

We emphasize that this result is relevant even if we are not complete in  $L_{\text{BCG}}$  or  $L_{200}$ , in the sense that this is the estimate that minimizes the scatter among the clusters we have. This does not imply that we could not have an even better sample if we included clusters with  $N_{200} < 10$  for which  $L_{\text{BCG}}$  is high. However, from Fig. 1, we see that we are complete for  $N_{200} L_{\text{BCG},10}^{0.75} > 80$ , so our results are not affected by incompleteness for number densities below  $5 \times 10^{-6} (h^{-1} \text{Mpc})^{-3}$ .

Together with the results of Section 5.2.1, these findings suggest that additional information from  $L_{\text{BCG}}$  provides improved determination of cluster masses, both in the mean and the scatter of the mass–observable relation.

## 6 SUMMARY AND CONCLUSIONS

We considered optical parameters that are available in large samples of clusters of galaxies: cluster richness  $N_{200}$ , cluster luminosity  $L_{200}$  and the luminosity of the BCG  $L_{\text{BCG}}$ , as well as power-law combinations of  $N_{200}$  with  $L_{\text{BCG}}$ , and  $L_{200}$  with  $L_{\text{BCG}}$ , to determine which is the best mass tracer for clusters.

We calibrate the mean mass relation for these tracers by measuring the stacked weak lensing signal around clusters rank ordered according to a given parameter. Our best-fitting mass relations for  $N_{200}$  and  $L_{200}$  are given in equations (15a) and (15b). We then ask whether the weak lensing signal changes significantly when a second parameter is added to the first one. We can exploit any such residual scaling to derive improved, lower scatter mass tracers. We explore such tracers in the form  $N_{200}^{\alpha N} L_{\text{BCG}}^{\gamma N}$  and  $L_{200}^{\alpha L} L_{\text{BCG}}^{\gamma L}$ . The best-fitting mass relations are given in equations (16a) and (16b), with parameters given in Table 2. The best mass tracer  $M_{200\bar{p}}$  (in units of  $10^{14} h^{-1} M_{\odot}$ ) we find is (for the lower redshift sample)

$$M_{14} = (1.27 \pm 0.08) \left( \frac{N_{200}}{20} \right)^{1.20 \pm 0.09} \left[ \frac{L_{\text{BCG}}}{\bar{L}_{\text{BCG}}(N_{200})} \right]^{0.71 \pm 0.14},$$

where  $\bar{L}_{\text{BCG}}(N_{200}) = 1.54 N_{200}^{0.41} \times 10^{10} h^{-2} L_{\odot}$  is the mean BCG luminosity at a given richness.

Our results suggest that  $L_{\text{BCG}}$  is an important second parameter in addition to  $N_{200}$  and  $L_{200}$ . At fixed  $N_{200}$ , residual scaling with  $L_{\text{BCG}}$  is seen at the  $\sim 5\sigma$  level in the lower redshift sample ( $0.10 < z < 0.23$ ), and at the  $\sim 1.5\sigma$  level in the higher redshift sample ( $0.23 < z < 0.30$ ). The need for a second parameter is less evident when  $L_{200}$  is used as the primary variable instead of  $N_{200}$ ; we find that residual scaling with  $L_{\text{BCG}}$  is preferred at the  $\sim 2\sigma$  level in the lower redshift sample, and find no evidence for residual scaling in the higher redshift sample.

We assess the relative amount of scatter in the various mass–observable relations by measuring the stacked weak lensing signal around clusters satisfying thresholds in each parameter. For a given comoving number density of clusters, low-scatter mass tracers will select more of the most massive clusters in the sample and thus yield a stronger lensing signal, compared to a large-scatter mass tracer. Among the parameters  $N_{200}$ ,  $L_{200}$  and  $L_{\text{BCG}}$ , cluster richness is the best mass tracer for clusters, while  $L_{\text{BCG}}$  is the poorest tracer. We find that a combined tracer of the form  $N_{200} L_{\text{BCG}}^{\beta N}$  reduces the scatter

in the mass relation compared to cluster richness taken alone, for the most massive clusters in the sample.

From SDSS spectroscopy of clusters in the maxBCG catalogue, Becker et al. (2007) found residual scaling of velocity dispersions with BCG luminosity  $L_{\text{BCG}}$  at fixed richness  $N_{200}$ . Our results consequently confirm that this residual scaling also appears in the projected mass distributions.

Our results are consistent with the current picture of cluster formation from halo mergers.  $N$ -body simulations and semi-analytic models find that at a fixed mass, dark matter haloes which form earlier have brighter, redder central subhaloes (i.e. brighter, redder BCGs) and lower richness (Wechsler et al. 2006; Croton et al. 2007). This may result from the satellites having had more time to merge on to the BCG, lowering the richness from when the cluster formed while enhancing the BCG luminosity. This implies that  $N_{200}$  and  $L_{\text{BCG}}$  are anticorrelated at fixed mass, and provides an explanation for our result above, i.e. that a combination of these two observables yields a tighter relation with mass than either of them taken alone.

The weaker residual scaling with  $L_{\text{BCG}}$  when using  $L_{200}$  instead of  $N_{200}$  suggests that the anticorrelation between  $L_{200}$  and  $L_{\text{BCG}}$  at fixed mass is much weaker; this is also consistent with the above scenario, since the luminosity of the BCG is included in the cluster luminosity. Moreover, this result constrains the amount of light that has been lost to the intracluster medium due to the merging of red satellite galaxies with the BCG since the formation of the cluster. If this was a significant fraction of the cluster luminosity in red galaxies,  $L_{200}$  would be lower for earlier-forming clusters, and therefore anticorrelated with  $L_{\text{BCG}}$ . We do not detect such an effect, so our results are consistent with a scenario where the cluster luminosity in red galaxies remains approximately constant over time.

Independent of the underlying astrophysical mechanisms, the improved mass tracers we found can be used to obtain accurate mass estimates and define mass thresholds in cluster samples with optical data. These in turn can be used to provide more precise constraints on cosmological parameters, such as the amplitude of mass fluctuations  $\sigma_8$ , which will be the subject of future work.

## ACKNOWLEDGMENTS

We thank Ben Koester, Tim McKay, Erin Sheldon and Risa Wechsler for useful discussions regarding the maxBCG cluster catalogue. RM is supported by NASA through Hubble Fellowship grant # HST-HF-01199.02-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. CH is supported by the US Department of Energy under contract DE-FG03-02-ER40701. US is supported by the Packard Foundation and NSF CAREER-0132953.

## REFERENCES

- Abazajian K. et al., 2003, *AJ*, 126, 2081
- Abazajian K. et al., 2004, *AJ*, 128, 502
- Abazajian K. et al., 2005, *AJ*, 129, 1755
- Adelman-McCarthy J. K. et al., 2006, *ApJS*, 162, 38
- Adelman-McCarthy J. K. et al., 2007, *ApJS*, 172, 634
- Adelman-McCarthy J. K. et al., 2008, *ApJS*, 175, 297
- Agustsson I., Brainerd T. G., 2006, *ApJ*, 644, L25
- Albrecht A. et al., 2006, preprint (astro-ph/0609591)
- Arnaud M., Pointecouteau E., Pratt G. W., 2007, *A&A*, 474, L37
- Bahcall N. A., Cen R., 1992, *ApJ*, 398, L81
- Becker M. R. et al., 2007, *ApJ*, 669, 905

- Bernstein G. M., Jarvis M., 2002, *AJ*, 123, 583
- Blanton M. R. et al., 2003a, *AJ*, 125, 2276
- Blanton M. R. et al., 2003b, *AJ*, 125, 2348
- Bradač M. et al., 2005, *A&A*, 437, 49
- Cacciato M., Bartelmann M., Meneghetti M., Moscardini L., 2006, *A&A*, 458, 349
- Croton D. J., Gao L., White S. D. M., 2007, *MNRAS*, 374, 1303
- Eisenstein D. J. et al., 2001, *AJ*, 122, 2267
- Finkbeiner D. P. et al., 2004, *AJ*, 128, 2577
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748
- Gunn J. E. et al., 1998, *AJ*, 116, 3040
- Gunn J. E. et al., 2006, *AJ*, 131, 2332
- Haiman Z., Mohr J. J., Holder G. P., 2001, *ApJ*, 553, 545
- Hansen S. M., McKay T. A., Wechsler R. H., Annis J., Sheldon E. S., Kimball A., 2005, *ApJ*, 633, 122
- Hansen S. M., Sheldon E. S., Wechsler R. H., Koester B. P., 2007, preprint (astro-ph/0710.3780)
- Hernquist L., 1990, *ApJ*, 356, 359
- Hirata C., Seljak U., 2003, *MNRAS*, 343, 459
- Hoekstra H., 2003, *MNRAS*, 339, 1155
- Hoekstra H., 2007, *MNRAS*, 379, 317
- Hogg D. W., Finkbeiner D. P., Schlegel D. J., Gunn J. E., 2001, *AJ*, 122, 2129
- Ivezić Ž. et al., 2004, *Astron. Nachr.*, 325, 583
- Johnston D. E. et al., 2007, preprint (astro-ph/0709.1159)
- Kaiser N., Squires G., Broadhurst T., 1995, *ApJ*, 449, 460
- Koester B. P. et al., 2007a, *ApJ*, 660, 221
- Koester B. P. et al., 2007b, *ApJ*, 660, 239
- Kravtsov A. V., Vikhlinin A., Nagai D., 2006, *ApJ*, 650, 128
- Lin Y.-T., Mohr J. J., 2004, *ApJ*, 617, 879
- Lupton R. H., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., Yasuda N., 2001, in Harnden F. R. Jr., Primini F. A., Payne H. E., eds, *ASP Conf. Ser. Vol. 238, Astronomical Data Analysis Software and Systems X*. Astron. Soc. Pac., San Francisco, p. 269
- Mahdavi A., Hoekstra H., Babul A., Sievers J., Myers S. T., Henry J. P., 2007, *ApJ*, 664, 162
- Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, *MNRAS*, 384, 1567
- Mandelbaum R., Tasitsiomi A., Seljak U., Kravtsov A. V., Wechsler R. H., 2005a, *MNRAS*, 361, 1287 (M05)
- Mandelbaum R., Tasitsiomi A., Seljak U., Kravtsov A. V., Wechsler R. H., 2005b, *MNRAS*, 362, 1451
- Mandelbaum R., Seljak U., Cool R. J., Blanton M., Hirata C. M., Brinkmann J., 2006, *MNRAS*, 372, 758
- Mandelbaum R., Seljak U., 2007, *J. Cosmol. Astropart. Phys.*, 6, 24
- Mandelbaum R. et al., 2008a, *MNRAS*, 386, 781
- Mandelbaum R., Seljak U., Hirata C. M., 2008b, *J. Cosmol. Astropart. Phys.*, 8, 6
- Metzler C. A., White M., Loken C., 2001, *ApJ*, 547, 560
- Nagai D., Kravtsov A. V., Vikhlinin A., 2007, *ApJ*, 668, 1
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Neto A. F. et al., 2007, *MNRAS*, 381, 4
- Padmanabhan N. et al., 2004, *New Astron.*, 9, 329
- Pedersen K., Dahle H., 2007, *ApJ*, 667, 26
- Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, *AJ*, 125, 1559
- Richards G. T. et al., 2002, *AJ*, 123, 2945
- Rykoff E. S. et al., 2008, *ApJ*, 675, 1106
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525
- Schmidt R. W., Allen S. W., 2007, *MNRAS*, 379, 209
- Seljak U., Warren M. S., 2004, *MNRAS*, 355, 129
- Sereno M., 2007, *MNRAS*, 380, 1207
- Sheldon E. S. et al., 2007a, preprint (arXiv:0709.1153)
- Sheldon E. S. et al., 2007b, preprint (arXiv:0709.1162)
- Sheth R. K., Tormen G., 1999, *MNRAS*, 308, 119
- Smith J. A. et al., 2002, *AJ*, 123, 2121
- Stoughton C. et al., 2002, *AJ*, 123, 485
- Strauss M. A. et al., 2002, *AJ*, 124, 1810
- Tucker D. L. et al., 2006, *Astron. Nachr.*, 327, 821
- Wang S., Haiman Z., Hu W., Khoury J., May M., 2005, *Phys. Rev. Lett.*, 95, 011302
- Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, *ApJ*, 652, 71
- Weller J., Battye R. A., 2003, *New Astron. Rev.*, 47, 775
- York D. G. et al., 2000, *AJ*, 120, 1579

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.