A comparison of weak-lensing masses and X-ray properties of galaxy clusters

Henk Hoekstra†
Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 5C2, Canada

Accepted 2007 May 6. Received 2007 April 24; in original form 2007 February 6

ABSTRACT
We present measurements of the masses of 20 X-ray luminous clusters of galaxies at intermediate redshifts, determined from a weak-lensing analysis of deep archival R-band data obtained using the Canada–France–Hawaii Telescope. Compared to previous work, our analysis accounts for a number of effects that are typically ignored, but can lead to small biases, or incorrect error estimates. We derive masses that are essentially model-independent and find that they agree well with measurements of the velocity dispersion of cluster galaxies and with the results of X-ray studies. Assuming a power law between the lensing mass and the X-ray temperature, $M_{2500} \propto T^\alpha$, we find a best-fitting slope of $\alpha = 1.34^{+0.30}_{-0.28}$. This slope agrees with self-similar cluster models and studies based on X-ray data alone. For a cluster with a temperature of $kT = 5$ keV we obtain a mass $M_{2500} = (1.4 \pm 0.2) \times 10^{14} h^{-1} M_\odot$ in fair agreement with recent Chandra and XMM studies.

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

1 INTRODUCTION
Clusters of galaxies have been the focus of intense study for many decades for a variety of reasons. They are the largest gravitationally bound objects in the Universe and provide a large reservoir of baryons, mainly in the form of the hot intracluster medium (ICM) gas. The study of galaxy clusters has been transformed recently with the advent of powerful X-ray telescopes such as Chandra and XMM, which enable us to determine the properties of the gas in unprecedented detail and accuracy. Such studies will hopefully lead us to a better understanding of the complex physics of the ICM and its interaction with the various constituents of the cluster (e.g. dark matter, galaxies). The dynamical state of the ICM might also reveal information about the recent merger history, thus providing an interesting way to test the concept of hierarchical cluster formation, as predicted in the cold dark matter (CDM) paradigm.

As the most massive objects in the Universe, galaxy clusters are readily found out to large redshifts in optical, X-ray or millimetre (Sunyaev–Zel’dovich, hereafter SZ) surveys. This makes them excellent cosmological probes, because the number of clusters as a function of mass is very sensitive to a range of cosmological parameters, such as the matter density, the normalization of the matter power spectrum and the equation of state of the dark energy (e.g. Evrard 1989; Eke et al. 1998; Henry 2000; Haiman, Mohr & Holder 2001; Levine, Schultz & White 2002; Allen et al. 2004). These constraints are complementary to measurements of large-scale structure and the cosmic microwave background.

Given the great interest in the use of galaxy clusters as a cosmological probe, and the advent of efficient SZ telescopes in the coming years, it is important to derive accurate masses for these systems. To estimate masses from the motion of cluster galaxies, one needs to make assumptions about the orbital structure and the geometry of the cluster in addition to the assumption of equilibrium. Similarly, X-ray and SZ measurements of the total mass assume that the cluster is in hydrostatic equilibrium. These assumptions are not always valid, thus complicating systematic studies of cluster properties. Finally, it is becoming increasingly clear that non-gravitational physics complicates matters further.

Fortunately, there exists a direct way to determine the cluster mass. The gradient in the gravitational potential of the cluster causes differential deflection of light rays coming from distant galaxies. This causes small, systematic distortions in the shapes of these faint sources, an effect known as weak gravitational lensing. The amplitude of the signal provides a direct measurement of the projected mass along the line of sight in a given aperture, which in turn can be compared directly to numerical simulations, a crucial step if one wants to use clusters for cosmology. However, to compare the weak-lensing results to other mass indicators one typically has to make assumptions regarding the cluster geometry.

Weak lensing is now a well-established technique to study the distribution of (dark) matter in the Universe and the applications are numerous (e.g. see Hoekstra, Yee & Gladders 2002a;
This camera consists of an array of 6400 pixels. The pixel scale is 0.206 arcsec, which ensures good sampling of the lensing signal caused by large-scale structure (cosmic shear) and the density of sources is often too low to warrant stable results. Instead, we retrieve red images from the second generation Digital Sky Survey (POSS II) for each cluster. These observations have small geometric distortions. This can be taken care of by calibrating the astrometry of the POSS II images using the USNO-A2 catalogue. SEXTRACTOR (Bertin & Arnouts 1996) is used to generate a catalogue of sources with accurate astrometry, with a number density significantly higher than that of the USNO-A2 catalogue. In addition, the POSS II images have been taken more recently, thus reducing the effects of proper motions of the stars.

This new astrometric catalogue is matched to each of the CFH12k exposures. We combine the matched catalogues for each exposure into a master catalogue, which contains the average positions of the matched objects. This master catalogue is used to derive the final second-order astrometric solution for each chip. This procedure ensures that in the overlapping areas the objects are accurately matched to the same position.

The images with the improved astrometry are stacked using the SIRVAP routine into a large mosaic image. As discussed below, for certain regions, data from different chips contribute to the image. If the point spread function (PSF) properties are discontinuous between chips, this may result in complicated behaviour of the PSF. Fortunately, for most of the data, the offsets between exposures are small. In addition, the CFH12k PSF shows no evidence of ‘jumps’, because the chips are very well aligned (e.g. see Hoekstra et al. 2002b).

As mentioned earlier, the ELIXIR-processed images contain photometric calibrations based on observations of standard stars during the observing run. These zero-points are only valid under photometric conditions, a condition that needs to be checked for each exposure. We therefore examine the magnitudes of a large number of objects in the images, which enables us to examine the variation in the photometric zero-point. We found that most data were taken under photometric conditions. For a few clusters only part
of the data were photometric and we use these images to scale the non-photometric data.

As additional checks we compared galaxy counts (with the cluster region excised) to those provided by P. Hsieh (private communication), and the expected distribution of $B - R$ colours of stars. In both cases we found good agreement between our data and the reference data. Where available we also compared our magnitudes to those observed by the Sloan Digital Sky Survey, and found excellent agreement. As a final check, we note that colours of the cluster red sequence are in good agreement with the expected values (the scatter around the mean is 0.03 in $B - R$).

3 WEAK-LENSING ANALYSIS

In this section, we discuss the details of the weak-lensing analysis and how we interpret the resulting lensing signal. In Section 3.1, we discuss how we measure the shapes of galaxies used in the weak-lensing analysis. A useful way of quantifying the lensing signal is presented in Section 3.2. There are various ways to estimate the cluster mass from the observed lensing signal. For instance, one can adopt a parametrized model for the density profile and fit this to the observations. We consider two such models. The singular isothermal sphere (SIS) is discussed in Section 3.3 and the CDM NFW profile (Navarro, Frenk & White 1997, NFW hereafter) is presented in Section 3.4. Another approach is to determine the projected mass within an aperture of a given radius, with a minimal dependence on the density profile at large radii. This approach is discussed in Section 3.5.

Various effects that complicate a simple interpretation of the lensing signal are reviewed in Section 3.6. In Section 3.7, we discuss the contamination of the lensing signal by cluster members, which are included in the catalogue of source galaxies. The conversion of the lensing signal into a mass requires knowledge of the redshift distribution of the source galaxies, which is addressed in Section 3.8. Finally, lensing is sensitive to all matter along the line of sight, which needs to be accounted for. This is reviewed in Section 3.9.

3.1 Shape measurements

The various steps in the weak-lensing analysis have been described in detail in Hoekstra et al. (1998) and Hoekstra et al. (2000). Our method is based on the original procedure developed by Kaiser, Squires & Broadhurst (1995) and Luppino & Kaiser (1997) with modification presented by Hoekstra et al. (1998). The key step in the lensing analysis is to accurately measure the shapes of the faint background galaxies, correcting for the various observational distortions, such as seeing and PSF anisotropy.

For the weak-lensing analysis we only use the stacked $R$-band images. As can be seen from Table 1 the $R$ exposures have better image quality than the $B$-band data. Good image quality is crucial for an accurate measurement of the mass using weak lensing.

Although the PSF anisotropy varies coherently over the total field of view, we follow van Waerbeke et al. (2005) and split the co-added images into subsets which correspond approximately to the original chips. This enables us to better characterize the spatial variation in the PSF anisotropy. The various images used to obtain the deep co-added images have been taken with (typically) small offsets which, in principle, complicates the analysis. However, discontinuities in the PSF anisotropy between chips are small and can be neglected in our analysis. This is supported by the results from the

Table 1. A summary of the observational data for the cluster sample.

<table>
<thead>
<tr>
<th>Name</th>
<th>RA</th>
<th>Dec.</th>
<th>z</th>
<th>$t_{\text{exp}}$ (B) (s)</th>
<th>Seeing (B) (arcsec)</th>
<th>$t_{\text{exp}}$ (R) (s)</th>
<th>Seeing (R) (arcsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2390</td>
<td>$21^h53^m36^s$</td>
<td>+17°41'44&quot;</td>
<td>0.2280</td>
<td>8100</td>
<td>0.90</td>
<td>9600</td>
<td>0.66</td>
</tr>
<tr>
<td>MS 0016+16</td>
<td>$00^h18^m33^s$</td>
<td>+16°26'16&quot;</td>
<td>0.5465</td>
<td>6180</td>
<td>0.87</td>
<td>9423</td>
<td>0.68</td>
</tr>
<tr>
<td>MS 0906+11</td>
<td>$09^h09^m12^s$</td>
<td>+10°58'28&quot;</td>
<td>0.1704</td>
<td>6000</td>
<td>0.93</td>
<td>13800</td>
<td>0.84</td>
</tr>
<tr>
<td>MS 1224+20</td>
<td>$12^h27^m13^s$</td>
<td>+19°50'56&quot;</td>
<td>0.3255</td>
<td>8640</td>
<td>1.01</td>
<td>11400</td>
<td>0.86</td>
</tr>
<tr>
<td>MS 1231+15</td>
<td>$12^h33^m55^s$</td>
<td>+15°25'58&quot;</td>
<td>0.2353</td>
<td>3900</td>
<td>0.90</td>
<td>9000</td>
<td>0.71</td>
</tr>
<tr>
<td>MS 1358+62</td>
<td>$13^h59^m50^s$</td>
<td>+62°31'05&quot;</td>
<td>0.3290</td>
<td>11880</td>
<td>1.04</td>
<td>6300</td>
<td>0.82</td>
</tr>
<tr>
<td>MS 1455+22</td>
<td>$14^h57^m15^s$</td>
<td>+22°20'35&quot;</td>
<td>0.2568</td>
<td>4800</td>
<td>1.03</td>
<td>9000</td>
<td>0.67</td>
</tr>
<tr>
<td>MS 1512+36</td>
<td>$15^h14^m22^s$</td>
<td>+36°36'21&quot;</td>
<td>0.3727</td>
<td>12540</td>
<td>0.90</td>
<td>9900</td>
<td>0.71</td>
</tr>
<tr>
<td>MS 1621+26</td>
<td>$16^h23^m35^s$</td>
<td>+26°34'14&quot;</td>
<td>0.4275</td>
<td>15300</td>
<td>0.99</td>
<td>11520</td>
<td>0.72</td>
</tr>
<tr>
<td>A68</td>
<td>$00^h37^m06^s$</td>
<td>+09°09'24&quot;</td>
<td>0.255</td>
<td>8100</td>
<td>1.05</td>
<td>7200</td>
<td>0.67</td>
</tr>
<tr>
<td>A209</td>
<td>$01^h31^m52^s$</td>
<td>-13°36'40&quot;</td>
<td>0.206</td>
<td>7200</td>
<td>0.98</td>
<td>5400</td>
<td>0.70</td>
</tr>
<tr>
<td>A267</td>
<td>$01^h52^m42^s$</td>
<td>+01°00'26&quot;</td>
<td>0.230</td>
<td>3000</td>
<td>0.93</td>
<td>4800</td>
<td>0.72</td>
</tr>
<tr>
<td>A383</td>
<td>$02^h48^m03^s$</td>
<td>-03°31'44&quot;</td>
<td>0.187</td>
<td>7200</td>
<td>0.93</td>
<td>4800</td>
<td>0.90</td>
</tr>
<tr>
<td>A963</td>
<td>$10^h17^m03^s$</td>
<td>-39°05'21&quot;</td>
<td>0.206</td>
<td>7200</td>
<td>0.88</td>
<td>4800</td>
<td>0.77</td>
</tr>
<tr>
<td>A1689</td>
<td>$13^h11^m30^s$</td>
<td>-01°20'30&quot;</td>
<td>0.1832</td>
<td>3600</td>
<td>0.88</td>
<td>3000</td>
<td>0.81</td>
</tr>
<tr>
<td>A1763</td>
<td>$13^h35^m20^s$</td>
<td>+41°00'04&quot;</td>
<td>0.2323</td>
<td>3600</td>
<td>0.94</td>
<td>6000</td>
<td>0.85</td>
</tr>
<tr>
<td>A2218</td>
<td>$16^h35^m48^s$</td>
<td>+66°12'51&quot;</td>
<td>0.1756</td>
<td>3378</td>
<td>1.06</td>
<td>3300</td>
<td>0.84</td>
</tr>
<tr>
<td>A2219</td>
<td>$16^h40^m19^s$</td>
<td>+46°42'41&quot;</td>
<td>0.2256</td>
<td>5400</td>
<td>0.91</td>
<td>6300</td>
<td>0.78</td>
</tr>
<tr>
<td>A370</td>
<td>$02^h39^m52^s$</td>
<td>-01°34'18&quot;</td>
<td>0.375</td>
<td>10408</td>
<td>0.92</td>
<td>10800</td>
<td>0.77</td>
</tr>
<tr>
<td>CL0024+16</td>
<td>$00^h26^m35^s$</td>
<td>+17°09'44&quot;</td>
<td>0.390</td>
<td>12900</td>
<td>0.98</td>
<td>9000</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Notes. Column 1: cluster name; columns 2 and 3: right ascension (RA) and declination (Dec.) (J2000.0) of the brightest cluster galaxy. Note that in the case of two dominant central galaxies this might differ from previous positions; column 4: cluster redshift; columns 5 and 7: exposure times in the $B$ and $R$ band, respectively; columns 6 and 8: seeing in $B$ and $R$, respectively. The top nine clusters have been studied as part of the CNOC cluster survey. The next nine systems are a subset of the systems studied by Bardeau et al. (2005) (although we note that A2390 is also part of that study). The last two clusters are well-known clusters because of their spectacular strong lensing.
The measurements of loci at the one presented in Fig. 1. The stars are defined by the vertical the apparent magnitude as a function of half-light radius, such as in the PSF, we select a sample of moderately bright stars from our to characterize the spatial variation of the galaxies for PSF anisotropy.

To measure the small, lensing-induced distortions, it is important to accurately correct the shapes for PSF anisotropy, as well as for the diluting effect of seeing. To characterize the spatial variation in the PSF, we select a sample of moderately bright stars from our images. The stars are selected based on their location in a plot of the apparent magnitude as a function of half-light radius, such as the one presented in Fig. 1. The stars are defined by the vertical locus at $r_b = 0.5$ arcsec, and allow for a clean sample of stars. We fit a second-order polynomial to the shape parameters of the selected stars for each chip, which provides us with an estimate of the PSF anisotropy at every position. These results are used to correct the shapes of the galaxies for PSF anisotropy.

Seeing circularizes the images, thus lowering the raw lensing signal. To correct for the seeing, we need to rescale the polarizations to their ‘pre-seeing’ value. This scale factor is the ‘pre-seeing’ shear polarizability $P'$ (Luppino & Kaiser 1997; Hoekstra et al. 1998). The measurements of $P'$ for each individual galaxy are noisy, and we therefore bin the measurements as a function of galaxy size, and use the ensemble average value as a function of size to correct for the effect of seeing. This approach has been tested on simulated images in great detail (e.g. Hoekstra et al. 1998; Heymans et al. 2006). These results suggest that we can recover the shear with an accuracy of $\sim 2$ per cent (Heymans et al. 2005).

Having corrected the shapes of the galaxies in each ‘chip’, we combine the catalogues into a master catalogue, which covers the full field of view. This catalogue is used for the weak-lensing analysis.

### 3.2 Tangential distortion

The azimuthally averaged tangential shear $\langle \gamma_t \rangle$ as a function of radius from the cluster centre is a useful measure of the lensing signal (e.g. Miralda-Escude 1991):

$$\langle \gamma_t \rangle(r) = \frac{\bar{\Sigma}(< r) - \bar{\Sigma}(> r)}{\Sigma_{\text{crit}}},$$  

where $\bar{\Sigma}(< r)$ is the mean surface density within an aperture of radius $r$, and $\bar{\Sigma}(> r)$ is the mean surface density on a circle of radius $r$. The convergence $\kappa$, or dimensionless surface density, is the ratio of the surface density to the critical surface density $\Sigma_{\text{crit}}$, which is given by

$$\Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_h}{D_l D_s}.$$  

where $D_t$ is the angular diameter to the lens, $D_s$ and $D_h$ are the angular diameter distances from the observer to the source and from the lens to the source, respectively. The parameter $\beta = D_h/D_s$ is a measure of how the amplitude of the lensing signal depends on the redshifts of the source galaxies.

Fig. 2(a) shows the observed lensing signal as a function of distance from the cluster centre for Abell 2390. The centre is taken to be the position of the brightest cluster galaxy. A significant signal is measured out to large radii. Panel b indicates a test for systematics: if the signal presented in panel (a) is caused by lensing, no signal
where the resulting values with their statistical errors are listed in Table 2.

Numerical simulations indicate that on large scales CDM gives rise to a particular density profile (e.g. Dubinski & Carlberg 1991; Navarro, Frenk & White 1995; NFW; Moore et al. 1999). In these simulations, the NFW profile, given by:

$$\rho(r) = \frac{M_{\text{vir}}}{4\pi r_s^3} \frac{1}{(r/r_s + 1)^2},$$

appears to be a good description of the radial mass distribution for haloes with a wide range in mass. Here, $M_{\text{vir}}$ is the virial mass, which is the mass enclosed within the radius $r_{\text{vir}}$. The virial radius is related to the scale radius through the concentration $c = r_{\text{vir}}/r_s$, and the function $f(c) = \ln(1 + c) - c/(1 + c)$.

One can fit the NFW profile to the measurements with $M_{\text{vir}}$ and concentration $c$ (or equivalently $r_s$) as free parameters. However, numerical simulations have shown that the average concentration depends on the halo mass and the redshift (Bullock et al. 2001):

$$c = \frac{9}{1 + z} \left( \frac{M_{\text{vir}}}{8.12 \times 10^{12} h M_{\odot}} \right)^{-0.14}.$$  

For the NFW mass estimates presented in Section 4, we will use this relation between mass and concentration, thus assuming that we can describe the cluster mass distribution with a single parameter. We note that the simulations show considerable scatter in the profiles from halo to halo. For instance, the values of $c$ for haloes of a given mass have a lognormal dispersion of approximately 0.14 around the median.

By definition, the virial mass and radius are related through:

$$M_{\text{vir}} = 4\pi \frac{\Delta_{\text{vir}}(z)}{3} \rho_0(z) r_{\text{vir}}^3,$$

where $\Delta_{\text{vir}}(z)$ is the virial overdensity and $\rho_0(z)$ is the mean density of the universe at redshift $z$.

### Table 2. Results for the SIS model.

<table>
<thead>
<tr>
<th>Name</th>
<th>$R_e$ (mag)</th>
<th>$\beta$</th>
<th>$\beta^2$</th>
<th>$D_{\text{f}}$ (h$^{-1}$ Gpc)</th>
<th>$r_{\text{E}}$ (arsec)</th>
<th>$\sigma_{\text{E}}$ (arsec)</th>
<th>$\sigma_{\text{LSS}}$ (arsec)</th>
<th>$\sigma_{\text{tot}}$ (arsec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2390</td>
<td>21–25</td>
<td>0.54</td>
<td>0.34</td>
<td>0.53</td>
<td>391</td>
<td>19.4</td>
<td>2.3</td>
<td>1.3</td>
</tr>
<tr>
<td>MS 0016+16</td>
<td>22–25.5</td>
<td>0.24</td>
<td>0.11</td>
<td>0.92</td>
<td>223</td>
<td>9.2</td>
<td>2.2</td>
<td>1.0</td>
</tr>
<tr>
<td>MS 0906+11</td>
<td>21–25</td>
<td>0.62</td>
<td>0.43</td>
<td>0.42</td>
<td>492</td>
<td>13.9</td>
<td>2.7</td>
<td>1.6</td>
</tr>
<tr>
<td>MS 1224+20</td>
<td>21–25</td>
<td>0.40</td>
<td>0.21</td>
<td>0.68</td>
<td>303</td>
<td>8.1</td>
<td>2.4</td>
<td>1.1</td>
</tr>
<tr>
<td>MS 1231+15</td>
<td>21–25.5</td>
<td>0.54</td>
<td>0.35</td>
<td>0.54</td>
<td>382</td>
<td>5.0</td>
<td>2.3</td>
<td>1.5</td>
</tr>
<tr>
<td>MS 1358+62</td>
<td>21–25</td>
<td>0.40</td>
<td>0.22</td>
<td>0.69</td>
<td>301</td>
<td>12.6</td>
<td>2.2</td>
<td>1.1</td>
</tr>
<tr>
<td>MS 1455+22</td>
<td>21–25.5</td>
<td>0.52</td>
<td>0.33</td>
<td>0.58</td>
<td>358</td>
<td>13.9</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td>MS 1512+36</td>
<td>22–25.5</td>
<td>0.40</td>
<td>0.22</td>
<td>0.74</td>
<td>278</td>
<td>5.9</td>
<td>2.2</td>
<td>1.2</td>
</tr>
<tr>
<td>MS 1621+26</td>
<td>22–25.5</td>
<td>0.34</td>
<td>0.18</td>
<td>0.81</td>
<td>255</td>
<td>9.7</td>
<td>2.3</td>
<td>1.1</td>
</tr>
<tr>
<td>A68</td>
<td>21–25</td>
<td>0.49</td>
<td>0.29</td>
<td>0.57</td>
<td>360</td>
<td>15.2</td>
<td>2.3</td>
<td>1.3</td>
</tr>
<tr>
<td>A209</td>
<td>21–25</td>
<td>0.57</td>
<td>0.38</td>
<td>0.49</td>
<td>423</td>
<td>13.3</td>
<td>2.3</td>
<td>1.4</td>
</tr>
<tr>
<td>A267</td>
<td>20–25</td>
<td>0.53</td>
<td>0.33</td>
<td>0.53</td>
<td>389</td>
<td>15.5</td>
<td>2.4</td>
<td>1.3</td>
</tr>
<tr>
<td>A383</td>
<td>21–24.5</td>
<td>0.58</td>
<td>0.38</td>
<td>0.45</td>
<td>456</td>
<td>8.2</td>
<td>3.2</td>
<td>1.3</td>
</tr>
<tr>
<td>A963</td>
<td>21–25</td>
<td>0.57</td>
<td>0.37</td>
<td>0.49</td>
<td>423</td>
<td>11.6</td>
<td>2.4</td>
<td>1.4</td>
</tr>
<tr>
<td>A1689</td>
<td>21–24.5</td>
<td>0.60</td>
<td>0.40</td>
<td>0.44</td>
<td>464</td>
<td>32.2</td>
<td>2.7</td>
<td>1.3</td>
</tr>
<tr>
<td>A1763</td>
<td>21–25</td>
<td>0.54</td>
<td>0.33</td>
<td>0.52</td>
<td>398</td>
<td>17.3</td>
<td>2.5</td>
<td>1.4</td>
</tr>
<tr>
<td>A2218</td>
<td>21–24.5</td>
<td>0.61</td>
<td>0.42</td>
<td>0.43</td>
<td>480</td>
<td>19.1</td>
<td>2.9</td>
<td>1.3</td>
</tr>
<tr>
<td>A2219</td>
<td>21–25</td>
<td>0.54</td>
<td>0.34</td>
<td>0.52</td>
<td>394</td>
<td>17.9</td>
<td>2.4</td>
<td>1.4</td>
</tr>
<tr>
<td>A370</td>
<td>22–25</td>
<td>0.37</td>
<td>0.19</td>
<td>0.75</td>
<td>277</td>
<td>19.7</td>
<td>2.4</td>
<td>1.0</td>
</tr>
<tr>
<td>CL0024+16</td>
<td>22–25.5</td>
<td>0.36</td>
<td>0.19</td>
<td>0.76</td>
<td>270</td>
<td>13.6</td>
<td>2.4</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Notes. Column 1: cluster name; column 2: adopted range in apparent magnitude for the source galaxies; column 3: effective value of $\beta$ as described in the text. The derived statistical error is 0.01; column 4: value for $\beta^2$; column 5: angular diameter distance to the cluster; column 6: angular size corresponding to $1 h^{-1}$ Mpc; column 7: value for the Einstein radius $r_{\text{E}}$ obtained from an SIS model fit to the tangential distortion from 0.25 to 1.5 $h^{-1}$ Mpc; column 8: statistical error for the measurement of the Einstein radius; column 9: uncertainty in $r_{\text{E}}$ caused by uncorrelated large-scale structure along the line of sight; column 10: total uncertainty in $r_{\text{E}}$, combining the statistical and LSS errors.

3.3 Singular isothermal sphere

A simple model for the cluster mass distribution is the SIS for which the convergence and tangential shear are given by:

$$\kappa = \langle \chi \rangle = \frac{r_{\text{E}}}{2\pi},$$

where $r_{\text{E}}$ is the Einstein radius. For the clusters in our sample we fit this model to the observed lensing signal (taking into account the complicating factors discussed below) from 0.25 to 1 $h^{-1}$ Mpc, and the resulting values with their statistical errors are listed in Table 2.

Under the assumption of isotropic orbits and spherical symmetry, the Einstein radius (in radians) is related to the line-of-sight velocity dispersion through:

$$r_{\text{E}} = 4\pi \left( \frac{\sigma_v}{c} \right)^2 \beta,$$

which allows for a direct comparison with measurements of the line-of-sight velocity dispersion of cluster galaxies from redshift surveys. We compare our lensing results with these dynamical measurements in Section 4.2.

3.4 NFW profile

Collisionless CDM provides a good description for the observed structure in the Universe. Numerical simulations indicate that on...
where \( \rho_v = 3H_0^2\Omega_m(1 + z)^3/(8\pi G) \) is the mean density at the cluster redshift and the virial overdensity \( \Delta_{vir} \approx (18\pi^2 + 82\xi - 39\xi^2)/\Omega(z) \), with \( \xi = \Omega(z) - 1 \) (Bryan & Norman 1998). For the \( \Lambda \)CDM cosmology considered here, \( \Delta_{vir}(0) = 337 \).

Cluster masses are also often quoted in terms of \( M_\Delta \), which is the mass contained within the radius \( r_\Delta \), where the mean mass density of the halo is equal to \( \Delta \) times the critical density at the redshift of the cluster. For reference with other work, we list results for a number of commonly used values for \( \Delta \). Note that \( M_{200} \) is often referred to as the virial mass, but that our definition for \( M_{200} \) is different.

The expressions for the tangential shear and surface density for the NFW profile have been derived by Bartelmann (1996) and Wright & Brainerd (2000) and we refer the interested reader to these papers for the relevant equations.

### 3.5 Aperture mass

Another approach to determine the cluster mass is known as aperture mass densitometry. It uses the fact that the shear can be related directly to a density contrast. We use the statistic of Clowe et al. (1998), which is related to the \( \xi \)-statistic of Fahlman et al. (1994).

In terms of the dimensionless surface density \( \xi(r) \) is equal to the mean surface density interior to \( r \) relative to the mean surface density in an annulus from \( r_1 \) to \( r_\text{max} \):

\[
\xi(r_1) = \frac{\bar{\kappa}(r' < r_1) - \bar{\kappa}(r_2 < r' < r_\text{max})}{\bar{\kappa}(r_2 < r_\text{max})}. \tag{8}
\]

It is related to the (observed) shear through

\[
\xi(r_1) = 2 \int_{r_1}^{r_2} \frac{d\ln r}{\bar{\kappa}} + \frac{2r_\text{max}^2 - r_2^2}{r_\text{max}^2 - r_2^2} \int_{r_2}^{r_\text{max}} d\ln r. \tag{9}
\]

Equation (8) shows that we can determine the average surface density within a given aperture up to a constant. If we ignore the surface density in the annulus, \( \xi_c \) provides a lower limit to the mass. This demonstrates the usefulness of wide-field imaging data: at large radii, the mean surface density in the annulus should be small and its contribution relative to the mean surface density within \( r_1 \) decreases.

We use the surface density from the NFW model to convert the observed signal into an estimate of the shear. In this case the mass estimate is based on the signal measured at large radii, and the correction is small.

For the mass estimates presented in Section 4, we adopt \( r_2 = 600 \) arcsec and \( r_\text{max} = 1000 \) arcsec. We estimate the mean surface density in the annulus based on the best-fitting NFW model. Thanks to our ability to measure the lensing signal out to large radii, the (model-dependent) correction to the mass is only \( \sim 10 \) per cent.

### 3.6 Complications in weak-lensing mass estimates

Hoekstra et al. (2000) lists various effects which complicate a simple interpretation of the lensing signal. Some of these have typically been included in published weak-lensing studies, but others are often ignored. We therefore briefly review them here.

First of all, the observed lensing signal is invariant under the transformation \( \kappa' = (1 - \lambda)k + \lambda \), which is usually referred to as the mass sheet degeneracy (Gorenstein, Shapiro & Falco 1988): we can determine the surface density up to a constant. This can also be seen from equation (8), where one needs to estimate the mean surface density in an annulus to derive the enclosed mass. At very large radii, however, the cluster surface density should be small, and the mass sheet degeneracy is less relevant. As mentioned above, we estimate the mass in the annulus from a mass model, and find that the correction to the enclosed mass is typically less than 10 per cent.

Hence, the uncertainty in the mass due to the mass sheet degeneracy is a few per cent at most for the wide-field imaging data used here.

The images of the distant galaxies are not only distorted, but also magnified. The flux is increased by a factor of \( \mu = ((1 - \kappa)^2 - \gamma^2)^{-1} \). This changes the source redshift distribution as intrinsically fainter galaxies are included as sources. This effect is small, even in the inner regions of the cluster where the surface density is large. It is therefore safe to ignore this effect in our analysis.

A complication which is relevant for our measurements is the fact that we do not measure the shear \( \gamma \) from the data, but the distortion (or reduced shear) \( g = \gamma / (1 - \kappa) \). In the weak-lensing limit (\( \kappa \ll 1 \)), the distortion is equal to the shear, but even at the large radii probed here, the amplitude of this effect is a few per cent. Hence, it is important to account for the \( (1 - \kappa) \) factor in our cluster mass estimates.

### 3.7 Contamination by cluster galaxies

Weak-lensing observations typically lack redshift information for the galaxies used in this analysis. As a result we cannot distinguish between unlensed cluster galaxies and lensed sources. The contamination by cluster galaxies will lower the lensing signal (where we assume their orientations are random) and therefore needs to be accounted for, that is, we need to increase the observed shear by a factor of \( 1 + f_{cg}(r) \), where \( f_{cg} \) is the fraction of cluster galaxies at that radius. Note that foreground field galaxies also dilute the lensing signal, but this is naturally accounted for in the calculation of the critical surface density (see equation 2 and the discussion in Section 3.8).

We used our data to examine the excess counts of galaxies as a function of distance from the cluster. We combined the measurements of the clusters in our sample. The amount of contamination at a given angular or physical scale depends on the mass (or richness) of the cluster. Although the range in mass is relatively small, we found that the profiles match best if we use the radius in units of \( r_{200} \) (or equivalently the virial radius). Note that this does not require a scaling of the counts themselves.

We also examined whether the contamination depends on cluster redshift. One might expect this, because the red sequence occupies different regions in the colour–magnitude diagram as a function of redshift. However, we did not find a significant trend with redshift and we therefore assume that the correction is redshift-independent.

The results are presented in Fig. 3. The shaded area indicates the fraction of cluster galaxies as a function of radius if no colour selection is used for the sources. At \( r_{2500} \) (which roughly corresponds to \( \sim 2 \) arcmin for the clusters considered here) we find that about 14 per cent of the galaxies are cluster members. Thanks to our colour information, we can define a cluster red sequence and reject galaxies that lie on this sequence. To do so we identify the red ridge of the cluster colour–magnitude relation and select all galaxies that are up to \( \sim 0.3 \) mag bluer. This removes the bright elliptical galaxies quite effectively, but at fainter magnitudes, the red sequence is not well defined and many cluster members are actually blue. Consequently the colour selection reduces the contamination by a modest \( \sim 30 \) per cent to 10 per cent at \( r_{2500} \). This remaining contamination is typically ignored in the literature.

A comparison to the overdensity of red galaxies shows that the observed contamination traces the distribution of red galaxies well. At large radii the measurements are noisy and suffer from the fact that it is difficult to determine the background level. We used the counts in an annulus from 10 to 15 arcmin (more than five times...
and use this to correct the profile and redetermine it is small at large radii. It is clear, however, that the contamination at large radii. We assume that the corrections for contamination by cluster members. The corrections are applied. The level of contamination is reduced by only \(\sim 30\) per cent by applying a colour-cut. However, it is clear that some contamination remains at small radii. The solid line is the best-fitting \(1/r\) model.

\(r_{2500}\) for most clusters) from the cluster centre. This may not be sufficiently far out, and as a result we may underestimate the contamination at large radii. It is clear, however, that the contamination is small at large radii.

To estimate the level of contamination as a function of radius, we assume that \(f_{cg}\) \(\propto r^{-1}\) (as the data do not allow a good estimate of the slope). The best-fitting result is indicated by the solid line in Fig. 3. We use this model to correct the tangential shear measurements for contamination by cluster members. The corrections are small, and we determine \(r_{2500}\) from the uncorrected shear profile, and use this to correct the profile and redetermine \(r_{2500}\). This iteration scheme converges rapidly. We find that our correction for the residual contamination by cluster members increases the measured Einstein radii by \(\sim 7\) per cent (see Table 2). The aperture masses listed in Table 3 are affected less, and the mass within an aperture of radius of \(0.5\) \(h^{-1}\) Mpc increases by \(\sim 4\) per cent.

### 3.8 Source redshift distribution

As mentioned above, the interpretation of the lensing signal requires knowledge of the redshifts of the source galaxies. Based on our data alone, we do not know the redshifts of the individual sources. The observed lensing signal, however, is an ensemble average of many different galaxies, each with their own redshift. As a result, it is sufficient to know the source redshift distribution to compute \(\beta\).

The source galaxies are typically too faint to be included in spectroscopic redshift surveys, although much progress is expected in the coming years. Instead we use the photometric redshift distributions determined from the Hubble Deep Fields (HDFs) (Fernández-Soto, Lanzetta & Yahil 1999). The clusters in our sample are at relatively low redshifts, which reduces the uncertainty in the mass measurements caused by errors in the redshift distribution. The HDF photometric redshift distributions match redshift distributions from other (shallower) photometric redshift surveys, such as COMBO-17 (Wolf et al. 2004) and the Red-Sequence Cluster Survey (P. Hsieh, private communication).

As discussed in the previous section, we identify the cluster colour–magnitude relation and remove galaxies with similar colours as the cluster galaxies from the source catalogues. This increases the lensing signal in the central regions. Consequently, the colour-cut changes the redshift distribution of the source galaxies, but we find that the effect is very small. It is negligible for the higher redshift clusters, but lowers the value for \(\langle \beta \rangle\) by \(\sim 2\) per cent for clusters at \(z \sim 0.2\).

The noise in the shape measurements for the faintest galaxies is large, which is further amplified by the large correction for seeing. We therefore limit the range in apparent magnitude for the sources used in the analysis. The adopted range is listed in Table 2. Following Hoekstra et al. (2000) we weigh the shapes proportional to the inverse square of the measurement error and adjust the redshift distribution accordingly. All the considerations listed above lead to an effective value for \(\langle \beta \rangle\), which is listed in Table 2. We estimate a statistical error in \(\langle \beta \rangle\) \(\sim 0.01\). This value is similar for all the clusters in our sample (note, however, that the relative error is larger for the higher redshift clusters).

When computing the ensemble averaged distortion, one uses an average value of \(\beta\) for the sources. In doing so, one effectively assumes that the redshift distribution can be approximated by a sheet of galaxies at a redshift corresponding to the mean value of \(\beta\). As shown by Hoekstra et al. (2000) this results in an overestimate of the shear by a factor

\[
1 + \frac{\langle \beta^2 \rangle_\beta \langle \beta \rangle_\beta^2 - 1}{\langle \beta \rangle_\beta^2} \kappa.
\]

For high-redshift clusters, this effect can be of comparable size as the correction for the fact that we measure the reduced shear. It is, however, typically ignored in the literature. We computed the values for \(\langle \beta^2 \rangle\) using the HDF photometric redshift distributions and the results are listed in Table 2. We include this correction for the mass estimates presented in Section 4.

### 3.9 Projection effects

The fact that lensing is sensitive to all matter along the line of sight complicates the direct comparison with other mass estimates. One complication is the three-dimensional structure of clusters (e.g. Corless & King 2006). Although the measurement of the weak-lensing signal does not require any assumptions about the geometry of the cluster, one does need to make such assumptions when comparing to other mass indicators, which depend on the cluster mass distribution in a different way. Therefore, some of the scatter between the various mass estimates presented in Section 4 will be caused by this effect.

Large-scale structure gives rise to two distinct contributions, both of which have been studied in detail. Structures associated with the cluster, such as filaments, have been studied by Metzler et al. (1999) and Metzler et al. (2001) using numerical simulations. Unfortunately, these studies focused on the use of aperture mass measurements at large radii, which makes it difficult to estimate the effect for our mass model fits. Nevertheless, it is clear from Metzler et al. (2001) that the effect is significantly reduced by focusing on the central regions of the cluster, which are much denser than the cosmic web. Alternatively, provided photometric redshifts of the sources are available, one can reconstruct the three-dimensional mass distribution (e.g. Taylor et al. 2004) and correct for additional structures along the line of sight.
The other contribution arises from distant (uncorrelated) large-scale structure. The observed aperture mass $M_{\text{obs}}(\theta)$ is the sum of the actual mass of the cluster $M_l(\theta)$ and the contribution from all other structure along the line of sight $M_{\text{LSS}}(\theta)$. The expectation value for the latter contribution vanishes, although the distribution is slightly skewed. Therefore, on average, the distant large-scale structure does not bias the lensing mass, but it introduces an additional uncertainty in the mass measurement of size $(M_{\text{LSS}}^2)^{1/2}$. This effect was first studied in detail in Hoekstra (2001) and Hoekstra and Mackey (2002) using a semi-analytic approach which has been verified using numerical simulations by White, van Waerbeke & Mackey (2002).

An expression introduced by the large-scale structure can be written as an integral of the projected convergence power spectrum $P_{\kappa}(l)$ multiplied by a filter function $g(l, \theta)$, which depends on the adopted statistic (Hoekstra 2001):

$$\langle M_{\text{LSS}}(\theta)^2 \rangle = 2\pi \int_0^\infty dl l P_{\kappa}(l) g(l, \theta)^2.$$  \hspace{1cm} (11)

The expressions for $P_{\kappa}(l)$ and $g(l, \theta)$ can be found in Hoekstra (2001). For the purpose of this paper, it suffices to note that the amplitude of the contribution of distant large-scale structure depends on the adopted cosmology and source redshift distribution. For the former, we adopt a $\Lambda$CDM cosmology with $\sigma_8 = 0.85$, while the source redshift distribution is the same as the one used in the cluster weak-lensing analysis.

The resulting additional uncertainty in the value of the Einstein radius is listed in Column 9 in Table 2. This additional error should be added in quadrature to the statistical error (as the two sources of uncertainty are uncorrelated), resulting in the total error listed in Column 10. Including the contribution from large-scale structure increases the uncertainty in the determination of the Einstein radius by $10\% - 15\%$. As shown in Hoekstra (2001) the effect of distant large scale is more important for aperture mass measurements, which is not surprising, given that such a mass determination depends on the lensing signal at large radii (e.g. see equation 9). We find that for the aperture mass measurements, the contribution from large-scale structure increases the true uncertainty by $20\% - 30\%$ over the statistical error. Similarly, we follow Hoekstra (2003) to estimate the errors for the NFW model fits. We do not list the additional uncertainties separately, but note that the large-scale structure contribution is included in the errors listed in Table 3.

Weak-lensing masses of clusters

4 MASS ESTIMATES

Table 3 lists the various weak-lensing mass estimates. Column 2 shows the inferred line-of-sight velocity dispersions from the SIS model fit discussed in Section 3.3. These results can be compared directly to velocity dispersions of galaxies from spectroscopic observations, which is done in Section 4.2. The error bars include both the statistical and distant large-scale structure errors.

Column 3 shows the projected mass within an aperture of 0.5 $h^{-1}$ Mpc using the aperture mass technique. This measurement requires no assumptions about the geometry or mass profile (apart from the small correction for the mean surface density in the annulus). However, it is difficult to compare this result to other mass indicators such as X-ray properties because the latter typically assume spherical symmetry.

The virial mass as defined by equation (7) is of interest, because it has some physical meaning. Given the low density in the ΛCDM model, the virial radius is rather large for the clusters considered here, and cannot be computed from the aperture mass method for two reasons. The first is that the area covered by the observations is not large enough. The second reason is more fundamental: as described in the previous section, projection effects limit the accuracy of the weak-lensing mass determination. At the virial radius these errors are too large for a useful measurement.

Instead, we list the values for $M_{500}$ and $M_{500}$ in Table 3, where the overdensities are measured with respect to the critical density at the redshift of the cluster. For these values of $\Delta$ we can derive reasonably accurate masses in a model-independent manner. For the comparison with the ASCA data in Section 4.3, $M_{2500}$ is most relevant as $r_{500}$ is close to the radius out to which X-ray temperatures are measured.

We also fitted the NFW profile to the lensing signal at radii $0.25 < r < 1.5 h^{-1} \text{Mpc}$, which is the same range as was used for the SIS model fit. The resulting virial masses for this model are presented in Column 9 of Table 3. For completeness, we also list the corresponding masses for a range of other values of $\Delta$ which are commonly found in the literature (see Section 3.4 for details).

We can measure $M_{500}$ in two different ways, which are almost independent. The aperture mass is predominantly based on data at large radii, whereas the NFW model fit uses measurements at small scales. The comparison of these mass estimates provides an additional check of the weak-lensing analysis. We compare these two mass estimates in Fig. 4, where we limited the NFW fit to radii $0.25 < r < 1 h^{-1} \text{Mpc}$, to minimize the overlap with the scales used for the aperture mass measurement (the mean value for $r_{500}$ for the clusters in our sample is $870 h^{-1} \text{kpc}$). We find that the mass estimates agree quite well.

4.1 Comparison with published weak-lensing results

Weak-lensing masses have been published for a fair number of the clusters in our sample, based on observations using a wide range of telescopes and cameras. In this section, we compare our results to the literature values. Before proceeding, we want to stress that in many cases the comparison is crude at best, because of differences in the assumed source redshift distribution, removal of cluster members, assumed mass model, etc.

The first detection of a weak-lensing signal was presented in Tyson et al. (1990) based on the analysis of A1689. Since then, several mass estimates for this cluster have been published, and we focus on two recent ones. Based on wide-field imaging data from the ESO 2.2-m telescope, Clowe & Schneider (2001) estimate a velocity dispersion of $\sigma = 1162 \pm 40 \text{ km s}^{-1}$ (using a mean source redshift of $\sim 0.5$), which implies a mass which is about 30 per cent lower than our value. The analysis presented in Clowe & Schneider (2001) lacked colour information, and hence a possible explanation for the lower signal could be contamination by cluster members. The comparison with the measurement of Bardeau et al. (2005) is of particular interest because it is based on the same data used here, although the data processing and weak-lensing analysis are completely independent. Bardeau et al. (2005) list an Einstein radius of 22 $\pm$ 3 arcsec, using a fit to the lensing signal from 70 to 1100 arcsec. Matching our measurement to their range, and ignoring contamination by cluster members, lowers our estimate for the Einstein radius from 32, $\pm$ 2.6 to 25.9 $\pm$ 1.7 arcsec, which is still 18 per cent higher than that estimated by Bardeau et al. (2005). More recently, Limousin et al. (2006) used a combined strong- and weak-lensing analysis. They list a value of $M_{500} = (19 \pm 3) \times 10^{14} \text{ h}^{-1} M_\odot$, which is in excellent agreement with our result. On small scales, however, the weak-lensing analysis appears to underestimate the mass. Limousin et al. (2006) find a projected mass of $(2.7 \pm 0.4) \times 10^{14} h^{-1} M_\odot$ within a radius of 45 arcsec. Extrapolating the weak-lensing aperture mass measurements inwards to this small scale, we find a mass of $(2 \pm 0.3) \times 10^{14} h^{-1} M_\odot$.

Another well-studied cluster is MS1224+20, because of its apparent high mass-to-light ratio. Fahlman et al. (1994) list a projected mass within a 2.76 arcmin radius of $3.5 \times 10^{14} h^{-1} M_\odot$, which is only slightly larger than our result of $(3.0 \pm 0.6) \times 10^{14} h^{-1} M_\odot$. The cluster was also observed by Fischer (1999) who inferred a velocity dispersion of $\sim 1300 \text{ km s}^{-1}$, considerably higher than our estimate. The lensing signal presented in Fischer (1999) barely decreases with distance, which is not observed in our data, and we cannot find an obvious explanation for this difference.

Squires et al. (1996a) studied A2390 and list a projected mass enclosed within a 100-arcsec aperture of $1.8 \times 10^{14} h^{-1} M_\odot$, in excellent agreement with our estimate of $2.1 \times 10^{14} h^{-1} M_\odot$. For A2218, Squires et al. (1996b) list a mass $M(<3.5 \text{ arcmin}) = (4.6 \pm$
1.4) \times 10^{14} h^{-1} M_{\odot}, which is higher than our value of 3.4 \times 10^{14} h^{-1} M_{\odot}.

Dahle et al. (2002) studied a sample of 38 X-ray luminous clusters, six of which overlap with our sample. We compared the estimates for the velocity dispersion listed in table 2 of Dahle et al. (2002) to our results and found fair agreement for A209, A267, A963, A1763 and A2219. However, Dahle et al. (2002) list a velocity dispersion of 1650 \pm 220 for A68, which is significantly higher than our measurement. Unfortunately, we were unable to identify a reason for this difference.

The mass distribution of MS1358+62 was studied by Hoekstra et al. (1998) using a mosaic of WFPC2 pointings. Hoekstra et al. (2002c) provide an updated value for the velocity dispersion of \sigma = 835_{-56}^{+52}. This number is lower than the value of 1048_{-112}^{+102} km s^{-1} found in the CFH12k analysis, but could be due to differences in the range of scales that were fitted.

CL0024+16 was also studied using HST observations. Kneib et al. (2003) were able to probe the lensing signal out to \sim 10 arcmin by sparsely covering the area with WFPC2 pointings. A direct comparison is complicated by the fact that Kneib et al. (2003) do not list the total mass of the cluster, but give the masses of two mass concentrations instead. The masses are obtained from NFW fits to the data. Adding the masses yields a value of M_{200} = 5.7 \pm 1.3 \times 10^{14} h^{-1} M_{\odot}. Our data imply a significantly higher mass of M_{200} = 15_{-9}^{+12} \times 10^{14} h^{-1} M_{\odot}. By considering only two clumps, Kneib et al. (2003) might have underestimated the mass, as a more diffuse mass distribution would not have been accounted for. To avoid this problem, we also fitted an SIS model to the tangential distortions presented in fig. 7 of Kneib et al. (2003). This yields an Einstein radius of r_{\text{E}} \sim 18 \text{ arcsec}. Kneib et al. (2003) use sources with 23 < I < 26, which yields a value of \beta = 0.54. Hence their signal corresponds to a velocity dispersion of 1075 km s^{-1}, which is in excellent agreement with our findings.

In summary, the agreement between our measurements and the literature values is fair, but there are a number of cases where the results are discrepant without a clear cause. However, based on the excellent quality of the CFH12k data used here, the fact that our weak-lensing technique is well tested, and the comparison with other mass estimators (see following sections), we are confident that our results are robust and accurate.

### 4.2 Comparison with dynamical mass estimates

The measurement of the line-of-sight velocity dispersion clearly benefits from observing a large number of galaxies. Such observations decrease the statistical error, but also allow for a better rejection of interlopers and for the identification of substructures. These studies are time-consuming, and as a result only a relatively small number of clusters have been studied in sufficient detail.

The clusters studied in this paper are well-studied clusters and therefore have relatively good spectroscopic coverage. Nine of the clusters are part of the CNOC1 redshift survey sample (e.g. Carlberg et al. 1996; Yee et al. 1996). Borgani et al. (1999) (re-)derived velocity dispersions for the clusters in this sample using an interloper rejection scheme which is more sophisticated than the one used by Carlberg et al. (1996), which also enabled them to measure separate velocity dispersions for the two clusters in the field of MS0906. The velocity dispersions derived by Borgani et al. (1999) are listed in Column 4 of Table 4. The dynamical data for the remaining clusters in our sample are also listed in Table 4.

Fig. 5 shows the velocity dispersion of the cluster galaxies versus the weak-lensing estimate using the SIS fit to the data. The solid points correspond to the clusters in the CNOC1 sample. The agreement between the weak-lensing and dynamical measurements is excellent for these data. The open points correspond to measurements from Girardi & Mezzetti (2001), which also agree with the lensing estimates, albeit with a larger scatter. Compared to the CNOC1 sample, these four clusters show evidence of more complicated dynamics.

### 4.3 Comparison with X-ray properties

Much work has been devoted to the X-ray properties of galaxy clusters. In particular scaling relations between the X-ray luminosity, temperature and mass are of great interest as they provide clues about cluster formation. For instance, simple self-similar models (e.g. Kaiser 1986; Bryan & Norman 1998) predict power-law relations between the mass, temperature and luminosity. More detailed numerical simulations (e.g. Evrard, Metzler & Navarro 1996; Bryan & Norman 1998) also provide evidence for simple scaling relations. If the gas is virialised, the mass $M_{\Delta}$ is given by

$$E(z)M_{\Delta} \propto T^{3/2},$$

\begin{table}
\centering
\begin{tabular}{|l|c|c|c|c|}
\hline
Name & $L_X$ & $kT_X$ & $\sigma_{\text{dyn}}$ & Reference \\
& (10^{44} h^{-2} \text{ erg s}^{-1}) & (keV) & (\text{km s}^{-1}) & \\
\hline
A2390 & 30.5 & 9.2 & $\pm 0.6$ & 1262_+89_-68 & 1 \\
MS 0016+16 & 39.3 & 8.7 & $^{+0.8}_{-0.7}$ & 1127_+166_-112 & 1 \\
MS 0906+11 & 6.65 & 6.1 & $\pm 0.4$ & 886_+78_-68 & 1 \\
MS 1224+20 & 3.3 & 4.8 & $^{+1.2}_{-1.0}$ & 831_+129_-57 & 1 \\
MS 1231+15 & – & – & – & 868_+65_-50 & 1 \\
MS 1358+62 & 9.3 & 6.7 & $^{+0.5}_{-0.4}$ & 1003_+61_-52 & 1 \\
MS 1455+22 & 12.7 & 4.5 & $\pm 0.2$ & 1032_+130_-95 & 1 \\
MS 1512+36 & 4.1 & 3.6 & $^{+0.9}_{-0.7}$ & 575_+138_-90 & 1 \\
MS 1621+26 & 8.1 & 6.5 & $^{+1.3}_{-1.0}$ & 839_+167_-53 & 1 \\
A68 & 11.3 & 8.0 & $^{+0.8}_{-0.6}$ & – & \\
A209 & – & – & – & – & \\
A267 & 8.6 & 5.9 & $^{+0.5}_{-0.4}$ & – & \\
A963 & 9.4 & 6.6 & $\pm 0.4$ & – & \\
A1689 & 25.1 & 9.2 & $^{+0.4}_{-0.3}$ & 1172_+123_-99 & 2 \\
A1763 & 15.6 & 7.7 & $^{+0.5}_{-0.4}$ & 1222_+147_-109 & 2 \\
A2218 & 8.3 & 7.0 & $^{+0.3}_{-0.2}$ & 1222_+147_-109 & 2 \\
A2219 & 32.8 & 9.8 & $^{+0.7}_{-0.6}$ & – & \\
A370 & 13.4 & 7.2 & $\pm 0.8$ & 859_+118_-112 & 2 \\
CL0024+16 & 3.5 & 5.2 & $^{+2.0}_{-1.3}$ & 911_+81_-107 & 2 \\
\hline
\end{tabular}
\caption{X-ray and dynamical properties.}
\end{table}

Notes. Column 1: cluster name; column 2: bolometric X-ray luminosity from Horner (2001). The values listed here are corrected for galactic absorption, and transformed to the cosmology adopted in this paper; column 3: X-ray temperature from Horner (2001) based on ASCA observations. The errors indicate the 90 per cent confidence intervals; column 4: velocity dispersion of cluster galaxies; column 5: references for velocity dispersions. References: (1) Borgani et al. (1999); (2) Girardi & Mezzetti (2001).
one expects that \(L\), which is the radius we will focus on.

The relevant temperature is the mean mass-weighted temperature within the radius \(r\). As mentioned earlier, X-ray observations typically measure temperatures out to \(r_{200}\), which is the radius we will focus on.

The source of the X-ray emission is bremsstrahlung and therefore one expects that \(L_X/E(z) \propto T^2\). The observed slope is found to be steeper (e.g. Edge & Stewart 1991; Markevitch 1998; Arnaud & Evrard 1999). For clusters with temperatures \(\geq 2\) keV the slope is \(\sim 3\), while even steeper slopes are observed for lower temperature galaxy groups (e.g. Helsdon & Ponman 2000). Allen & Fabian (1998) have argued that the steeper slope, compared to the self-similar case, of the hotter clusters is caused by the effects of cool central components. However, studies that attempt to account for this still find steeper slopes, with values \(\sim 2.7\) (e.g. Markevitch 1998; Lumb et al. 2004).

When deriving X-ray temperatures, various groups employ different approaches to deal with temperature gradients and cool cores. To avoid introducing scatter caused by variations in the analysis method, we use the results from Horner (2001). The temperatures are based on ASCA observations. Although these measurements are not necessarily the most recent and accurate, they do provide a large sample, analysed homogeneously. As a caveat, we note that Horner (2001) did not attempt to correct for the cool centres of clusters. As part of the Canadian Cluster Comparison Project\(^1\) (CCCP) we will derive X-ray properties in a consistent manner from modern Chandra and XMM observations, accounting for temperature variations.

The large open points in Fig. 6 indicate the temperature as a function of bolometric X-ray luminosity for the clusters studied in this paper. With the exceptions of MS1455+22 and MS1512+36 (both strong ‘cooling flow’ clusters), the clusters follow a very tight relation. If we exclude these outliers, we find a slope of \(3.57 \pm 0.23\) for our sample.

The aim of the CCCP is not only to derive X-ray temperatures in a consistent manner, but also to involve the systematic study of the scatter in the relation between the weak-lensing mass and the X-ray properties. This paper is a first attempt, but a more comprehensive study requires a larger sample of clusters. As part of the CCCP we therefore complement the sample studied here with an additional 30 massive clusters, resulting in a sample of \(\sim 50\) clusters with \(T > 5\) keV, with lensing masses derived from deep CFHT imaging.

The small points in Fig. 6 correspond to the additional clusters in the CCCP sample. Although many of these clusters lie on the same relation as the clusters studied here, the overall scatter is significantly larger! Part of the scatter is likely to be caused by the fact that Horner (2001) did not correct for the presence of cool cores. For instance, Markevitch (1998) has shown that a considerable fraction of the scatter in the \(L-T\) relation stems from scatter in the strength of the cool core. This is supported by theoretical models which indicate that a cool core lowers the average temperature and raises the X-ray luminosity of a cluster of a given mass (e.g. Voit et al. 2002; McCarthy et al. 2004). However, the sample of clusters studied here contains a typical mix of cooling and non-cooling flow systems similar to other studies (e.g. Peres et al. 1998; Bauer et al. 2005) and thus we do not expect a bias towards the selection of clusters with relatively weak cool cores.

Nevertheless, to understand the wide range of cluster properties, larger samples of well-studied clusters are needed. Such samples will also improve our estimates of the normalization and slope of the mass–temperature relation agree well with theoretical predictions (e.g. Evrard et al. 1996). However, models that lack feedback processes do not fare so well for the amplitude. Compared to these models, X-ray observations find lower masses (e.g. Horner, Mushotzky & Scharf 1999; Nevalainen, Markevitch &

---

\(^1\) http://www.astro.uvic.ca/~hoekstra/CCCP.html
The origin of the differences between theory and observations is not fully resolved, but several explanations have been proposed, many of which are related to non-gravitational physics. Cooling and feedback processes tend to alter the mass–temperature relation in a systematic way by raising the gas entropy level (as compared to the purely gravitational heating scenario). The net result is an increased temperature, or lower normalization (e.g. Voit et al. 2002; Borgani et al. 2004). Because the X-ray mass estimates are also affected by the complex ICM physics, an independent weak-lensing mass estimate is of great importance if we want to understand the mass–temperature relation.

Using the expected relation between luminosity and temperature, these simple models also predict that the mass and luminosity are related through

$$E(z)M_{\Delta} = \left( \frac{L_X}{E(z)} \right)^{3/4}.$$  \hfill (14)

Fig. 7(a) shows the observed mass–luminosity relation. We use the value for $M_{2500}$ derived from the aperture mass method. The best-fitting power-law model is indicated by the dashed line. Fig. 7(b) shows the likelihood contours for the parameters of this power-law model for the $M_{2500}$–luminosity relation. The best-fitting model has a $\chi^2 = 38.1$ for 15 degrees of freedom. Hence, the mass–luminosity relation shows evidence of intrinsic scatter. We find a slope of $\alpha = 0.43^{+0.09}_{-0.10}$ and a mass of $M_{2500} = 1.94^{+0.11}_{-0.12} \times 10^{14} h^{-1} M_\odot$ for a cluster with a luminosity of $10^{45} h^{-2} \text{ erg s}^{-1}$. The slope expected from simple self-similar models is ruled out by our measurements at the $\sim 3\sigma$ level. This is not surprising, as this reflects the fact that the observed $L-T$ relation is steeper than the one expected from self-similar models: if we assume $M \propto T^{3/2}$ and $L \propto T^3$, we expect $M \propto L^{1/2}$, as is observed.

Fig. 7(c) shows the observed mass–temperature relation. We find $\chi^2 = 25.3$ for the best-fitting power-law model, which is somewhat high, but significantly smaller than for the mass–luminosity relation. The likelihood contours for the power-law model parameters are presented in Fig. 7(d). In this case we find a slope $\alpha = 1.34^{+0.30}_{-0.28}$, which agrees well with the self-similar slope of 1.5. The slope also agrees well with studies based on X-ray data alone (e.g. Nevalainen et al. 2000; Allen, Schmidt & Fabian 2001; Arnaud et al. 2005; Vikhlinin et al. 2006).

Our best-fitting power-law model yields a mass of $M_{2500} = (1.4 \pm 0.2) \times 10^{14} h^{-1} M_\odot$ for a cluster with a temperature of 5 keV. Allen et al. (2001) list a mass of $M_{2500} = (3.8 \pm 0.4) \times 10^{14} h^{-1} M_\odot$ for a cluster with a temperature of 5 keV at $z = 0$. The contours indicate the 68.3, 95.4 and 99.7 per cent confidence limits on two parameters jointly. The side panels show the probability density distribution for each parameter (while marginalizing over the other).

**Figure 7.** Panel (a): $M_{2500}$ as a function of the bolometric X-ray luminosity. To account for the range in redshift of the clusters, the mass and luminosity have been rescaled using the corresponding value for $E(z)$. Panel (b): $M_{2500}$ as a function of the X-ray temperature. The dashed lines indicate the best-fitting power-law relations. Panel (a): likelihood contours for the slope of the $M_{2500}-L_X$ relation and the mass of a cluster with a luminosity of $L_X = 10^{45} \text{ erg s}^{-1}$ at $z = 0$. Panel (b): likelihood contours for the slope of the $M_{2500}-T_X$ relation and the mass of a cluster with a temperature of $T_X = 5 \text{ keV}$ at $z = 0$. The contours indicate the 68.3, 95.4 and 99.7 per cent confidence limits on two parameters jointly. The side panels show the probability density distribution for each parameter (while marginalizing over the other).
10-keV cluster (assuming a slope of 1.5). If we fix the slope to 1.5, we find $M_{2500} = (3.6 \pm 0.2) \times 10^{14} h^{-1} M_\odot$ for a 10-keV cluster, which is in excellent agreement with Allen et al. (2001).

Arnaud et al. (2005) studied the $M$–$T$ relation of a sample of six nearby relaxed clusters ($T > 3.5$ keV), using XMM–Newton. They find a mass of $(1.3 \pm 0.04) \times 10^{14} h^{-1} M_\odot$ for a 5-keV cluster, in excellent agreement with our findings. Finally, Vikhlinin et al. (2006) list a mass of $(0.9 \pm 0.05) \times 10^{14} h^{-1} M_\odot$ for a 5-keV cluster. This result is marginally consistent with our result. Note, however, that the ASCA temperatures used here may be biased low, because they have not been corrected for cool cores. A rough comparison between the Vikhlinin et al. (2006) and Horner (2001) temperatures suggests that the ASCA temperatures used here are about 10 per cent lower. This suggests that our normalization may need to be reduced by about 15 per cent.

So far, we have assumed that the evolution of the X-ray properties follows the expected evolution, parametrized by $E(z)$. To examine whether the cluster properties show evidence of additional evolution, we plot the ratio of the lensing mass and the best mass as a function of redshift in Fig. 8(a). Similarly, Fig. 8(b) shows the results for the $M$–$T$ relation. The clusters with $z > 0.35$ seem to have somewhat larger-than-expected lensing masses, but a larger sample is required to confirm this suggestion.

We find that the residuals in the $M$–$L_X$ and $M$–$T$ relations are highly correlated. This is not too surprising, given that the clusters in our sample follow such a tight $L$–$T$ relation. The implications, however, are interesting. The largest outliers in Fig. 8 are CL0024+16 and A370, for which we find lensing masses that are approximately two times larger than that expected from the X-ray properties. Yet, these clusters lie on the $L$–$T$ relation. This discrepancy could indicate a problem with the weak-lensing mass determination, but we were unable to identify an obvious measurement error. However, these clusters have been studied in detail because of their extreme strong-lensing properties. Therefore, these clusters may be considered ‘lensing-selected’, which could explain their larger masses than that expected when compared to their X-ray properties. A related observation is that both clusters show evidence of recent, ongoing merging. The kinematics of CL0024+16 have been studied in detail by Czoske et al. (2002), who conclude that it might have experienced a high-speed collision. A370 consists of two distinct mass concentrations (e.g. Kneib et al. 1993), which enhances its strong-lensing cross-section.

5 CONCLUSIONS

Tremendous progress in measurement techniques, in conjunction with new field-imaging capabilities on large telescopes, has led us to the next step in the study of galaxy clusters using weak gravitational lensing: the study of the mass distribution of large samples of clusters. This paper presents the first results of such a systematic, multiwavelength study of rich clusters of galaxies.

In this paper, we have presented new measurements of the masses of 20 X-ray luminous clusters of galaxies at intermediate redshifts. The results are based on a careful analysis of deep archival $R$-band data obtained using the CFHT. In particular, we accounted for a number of subtle effects that, if ignored, can lead to small biases or incorrect error estimates. Thanks to the wide field of view of the CFH12k camera, we were able to derive masses that are essentially model-independent.

A comparison of the lensing results with measurements of the velocity dispersion of cluster galaxies shows good agreement. We typically find good agreement between our results and weak-lensing mass estimates in the literature. Assuming a power law between the lensing mass and the X-ray temperature, $M_{2500} \propto T^\alpha$, we find a best-fitting slope of $\alpha = 1.34_{-0.28}^{+0.30}$. This slope agrees with self-similar cluster models and studies based on X-ray data alone (Nevalainen et al. 2000; Allen et al. 2001; Arnaud et al. 2005; Vikhlinin et al. 2006). For a cluster with a temperature of $kT = 5$ keV we obtain a mass $M_{2500} = (1.4 \pm 0.2) \times 10^{14} h^{-1} M_\odot$, in fair agreement with recent Chandra and XMM studies (e.g. Allen et al. 2001; Arnaud et al. 2005; Vikhlinin et al. 2006).

The comparison to X-ray properties is complicated by the fact that the analysis pipelines employed by different groups can yield quite different luminosities and temperatures. We therefore used measurements by Horner (2001) which are based on ASCA observations and not corrected for the presence of cool cores. As part of the Canadian Cluster Comparison Project (CCCP), we are re-analysing available modern X-ray data in a consistent manner. Furthermore, to improve constraints on the normalization and slope of the scaling relations between mass and X-ray properties, as well as quantifying the scatter in these relations, even larger samples of clusters, with accurate weak-lensing masses, are required. To achieve this goal, we have recently augmented the sample of clusters studied in this paper with deep $g'$ and $r'$ CFHT Megacam imaging of an additional 30 massive clusters.

ACKNOWLEDGMENTS

We thank Andisheh Mahdavi and Arif Babul for many useful discussions. We are also grateful to Pat Henry and Mark Voit for comments on this manuscript. This research was supported by the National Science and Engineering Research Council (NSERC), the Canadian Foundation for Innovation (CFI), and the Canadian Institute for Advanced Research (CIAR). 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.
