The Astrophysical Journal, 604:596–603, 2004 April 1

WEAK-LENSING MASS RECONSTRUCTION OF THE INTERACTING CLUSTER 1E 0657−558: DIRECT EVIDENCE FOR THE EXISTENCE OF DARK MATTER[1]

Douglas Clowe[2]
Institut für Astrophysik und Extraterrestrische Forschung der Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany; dclowe@as.arizona.edu

Anthony Gonzalez
Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611-2055

and

Maxim Markevitch
Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138
Received 2003 October 28; accepted 2003 December 11

ABSTRACT

We present a weak-lensing mass reconstruction of the interacting cluster 1E 0657−558, in which we detect both the main cluster and a subcluster. The subcluster is identified as a smaller cluster that has just undergone initial infall and pass-through of the primary cluster and has been previously identified in both optical surveys and X-ray studies. The X-ray gas has been separated from the galaxies by ram pressure–stripping during the pass-through. The detected mass peak is located between the X-ray peak and galaxy concentration, although the position is consistent with the galaxy centroid within the errors of the mass reconstruction. We find that the mass peak for the main cluster is in good spatial agreement with the cluster galaxies and is offset from the X-ray halo at 3.4 σ significance, and we determine that the mass-to-light ratios of the two components are consistent with those of relaxed clusters. The observed offsets of the lensing mass peaks from the peaks of the dominant visible mass component (the X-ray gas) directly demonstrate the presence, and dominance, of dark matter in this cluster. This proof of dark matter existence holds true even under the assumption of modified Newtonian dynamics (MOND); based on the observed gravitational shear–optical light ratios and the mass peak–X-ray gas offsets, the dark matter component in a MOND regime would have a total mass that is at least equal to the baryonic mass of the system.

Subject headings: dark matter — galaxies: clusters: individual (1E 0657−556) — gravitational lensing

1. INTRODUCTION

It has long been established that the velocity dispersions and X-ray gas temperatures of clusters of galaxies are too high to be explained solely by the amount of visible matter in the clusters using a physical model with Newtonian gravity and general relativity. This observation led to the introduction of a dark matter component of the mass, which interacts with normal matter and light via only gravity. Recent observations of clusters suggest that the virialized mass is made of ∼1% baryons observable in optical and infrared data, ∼11% baryons observable in X-ray data (e.g., Allen, Schmidt, & Fabian 2002), and the remaining ∼88% in a dark matter component.

An alternative explanation has been that the gravitational force only follows the Newtonian r^{-2} law at the level of the force observed in the solar system and that for smaller forces the decline with distance is less (Milgrom 1983). This idea of modified Newtonian dynamics (MOND) has been used to reproduce the observed rotation velocities of spiral galaxies without inclusion of any dark matter (e.g., McGaugh & de Blok 1998) and could also explain the observed velocity dispersions and X-ray temperatures of clusters without needing any additional matter beyond the observed baryons and a small neutrino mass (Sanders 2003). Several arguments against MOND have been made using gravitational lensing observations of galaxies and clusters (Gavazzi 2002; Hoekstra, Yee, & Gladders 2002; Mortlock & Turner 2001), but because of the lack of a general relativistic theory giving the strength of the interaction between light and gravity in the MOND regime, these observations can be explained by alterations in the MOND formalism (Sanders & McGaugh 2002).

A definitive test of MOND, however, can be made with interacting clusters of galaxies. In the standard cold dark matter (CDM) paradigm, during the initial pass-through the dark matter particles and the galaxies are effectively collisionless, while the X-ray halo is affected by ram pressure. As a result, one expects the galaxies and dark matter halo to remain spatially coincident following the interaction, while the X-ray halo is displaced toward the center of mass of the combined system (e.g., Tormen, Moscardini, & Yoshida 2003). In the CDM paradigm the mass of the X-ray halo is a small component of the total mass, and therefore the mass maps created from weak lensing should have the primary mass peaks in good spatial agreement with the galaxies. In a MOND regime, however, the X-ray gas is the dominant component of the total mass. The weak-lensing mass reconstruction would therefore detect a primary mass peak coincident with the gas, which is spatially offset from the galaxy distribution.

The $z = 0.296$ interacting cluster 1E 0657−558 provides the ideal case in which to test this theory. First discovered by Tucker, Tananbaum, & Remillard (1995), subsequent analysis of *ROSAT* HRI data revealed that the system is comprised of two merging subclusters (Tucker et al. 1998). More recent *Chandra* and spectroscopic observations further indicate that

[1] Based on observations made with ESO telescopes at the Paranal Observatories under program IDs 60.A-9203 and 64.O-0332.
[2] Currently at Steward Observatory, University of Arizona.

this merger is nearly in the plane of the sky (Barrena et al. 2002; Markevitch et al. 2002), with the lower mass subcluster having recently exited the core of the main cluster (Markevitch et al. 2002).

The *Chandra* observations by Markevitch et al. (2002; M. Markevitch et al. 2004, in preparation) have been particularly valuable in elucidating the dynamical state and geometry of this unique system. These data reveal the presence of a prominent bow shock leading the lower mass subcluster ($T \sim 6$ keV), which is exiting the core of the main cluster ($T \sim 14$ keV) with a relative velocity of 4500 km s^{-1}, determined from the gas density jump at the bow shock. Coupled with the current 0.66 Mpc separation between the two components, this velocity requires the closest approach of the two components to have occurred 0.1–0.2 Gyr ago. The merger is constrained to be nearly in the plane of the sky by the sharpness of the shock front, a result consistent with the small line-of-sight component of the subcluster velocity derived from the spectroscopic data by Barrena et al. (2002). Finally, a comparison of the *Chandra* data with optical imaging reveals that the X-ray gas associated with the bullet trails the galaxy distribution. This latter result, coupled with the simple geometry of the system, enables the definitive test of MOND that we describe below. In a related paper, Markevitch et al. (2004) utilize the combination of *Chandra* and weak-lensing data to also constrain the collisional cross section for self-interacting dark matter.

This is not the only system that is known to have a spatial offset between the galaxies and X-ray gas in a subcomponent. The high-redshift cluster MS 1054−0321 has a double-peaked X-ray halo in which the western peak is offset from the nearby galaxy overdensity (Jeltema et al. 2001). Unlike 1E 0657−558, however, no shock front is observed in the X-ray data, and as a result the relative velocities and geometry of the merging components are unknown. Furthermore, while a weak-lensing mass peak has been measured near the galaxy overdensity (Hoekstra, Franx, & Kuijken 2000; Clowe et al. 2000), the uncertainty in the position of this mass peak is quite large (Marshall et al. 2002).

In this paper we use B and I images obtained from the ESO archive, taken with the FORS1 instrument in direct imaging mode on the 8 m Very Large Telescope 1 (VLT1) during 1998 and 2000. These include the images used in Barrena et al. (2002), but we have independently created the final images from the raw data. In § 2 we present the weak-lensing analysis of the image and discuss the significance and uncertainties in the positions of the detected mass peaks. We analyze the photometry in § 3 and give mass-to-light ratios for the detected mass peaks. Discussion of the results and our conclusions are presented in § 4. Throughout this paper we assume an $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1} universe unless stated otherwise.

2. WEAK-LENSING ANALYSIS

Weak gravitational lensing is a method that can be used to measure the surface mass in a region by utilizing the fact that the path of a light bundle passing a gravitational potential will be bent by the potential. As a result, images of background galaxies that are near a massive structure, such as a cluster of galaxies, are deflected away from the structure, becoming enlarged while preserving the surface brightness and distorted such that they are stretched tangentially to the center of the potential. This third effect, known as gravitational shear (γ), causes the background galaxies' ellipticities to deviate from an isotropic distribution, and the magnitude and direction of these deviations is used to measure the mass of the structure(s) causing the lensing. This technique of measuring the mass does not make any assumptions about the dynamical state of the mass and is therefore one of the few methods that can be used to measure the mass of a dynamically disturbed system.

The first step in the weak-lensing analysis is to detect background galaxies, measure their ellipticities, and correct the ellipticities for smearing due to the point-spread function (PSF). We used the prescription given in Clowe & Schneider (2002) for performing this step, in which the objects are detected and have their photometry measured using SExtractor (Bertin & Arnouts 1996), their shapes measured using the IMCAT software package,[3] and the PSF smearing correction performed using the KSB technique (Kaiser, Squires, & Broadhurst 1995). Background galaxies were selected using the criteria of having $I > 20$, $B-I < 3.2$, a detection significance in I greater than 11, and a 50% encircled light radius larger than that of a star. This selection resulted in a catalog with a density of 12.3 galaxies arcmin^{-2} over a box 6.′7 on a side, which is complete, as measured by the departure of the number counts from a power law, to $I \sim 24.5$, with the faintest galaxy having $I = 25.97$.

The next step in weak-lensing analysis is to convert the measured shear into a measurement for the convergence κ, which is related to the surface density of the lens Σ via

$$\kappa = \frac{\Sigma}{\Sigma_{\rm crit}}, \tag{1}$$

where $\Sigma_{\rm crit}$ is a scaling factor,

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}}, \tag{2}$$

D_s is the angular distance to the source (background) galaxy, D_l is the angular distance to the lens (cluster), and D_{ls} is the angular distance from the lens to the source galaxy. Using the same magnitude and color selections on the Hubble Deep Field–South photometric redshift catalog from Fontana et al. (1999) as were used to create the background galaxy catalog gives a mean-lensing redshift of $z_{\rm bg} = 0.85$ and a $\Sigma_{\rm crit} = 3.1 \times 10^9\ M_\odot$ kpc^{-2}.

Shown in Figure 1 in solid dark contours is a map of κ for this field created by using the KS93 algorithm (Kaiser & Squires 1993), which uses the fact that both the shear and the convergence are combinations of various second derivatives of the surface potential, and therefore the Fourier transform of the shear can be converted into the Fourier transform of κ by the multiplication of the appropriate wavenumbers. Because we are reconstructing a small field around a massive cluster, however, we actually measure the reduced shear $g = \gamma/(1-\kappa)$ from the background galaxy ellipticities. Therefore, we must perform an iterative solution to the KS93 algorithm in which an initial κ map is assumed (in this case, $\kappa = 0$ everywhere), g is corrected with this map to γ, which is then transformed to a new κ map, which is then used in turn to correct g, etc. (Seitz & Schneider 1995). This technique typically converges in a few iterations (in this case, six) and gives a measurement of κ in the field relative to the level of κ at the edge of the image, which is unknown.

[3] Available at http://www.ifa.hawaii.edu/~kaiser/imcat.

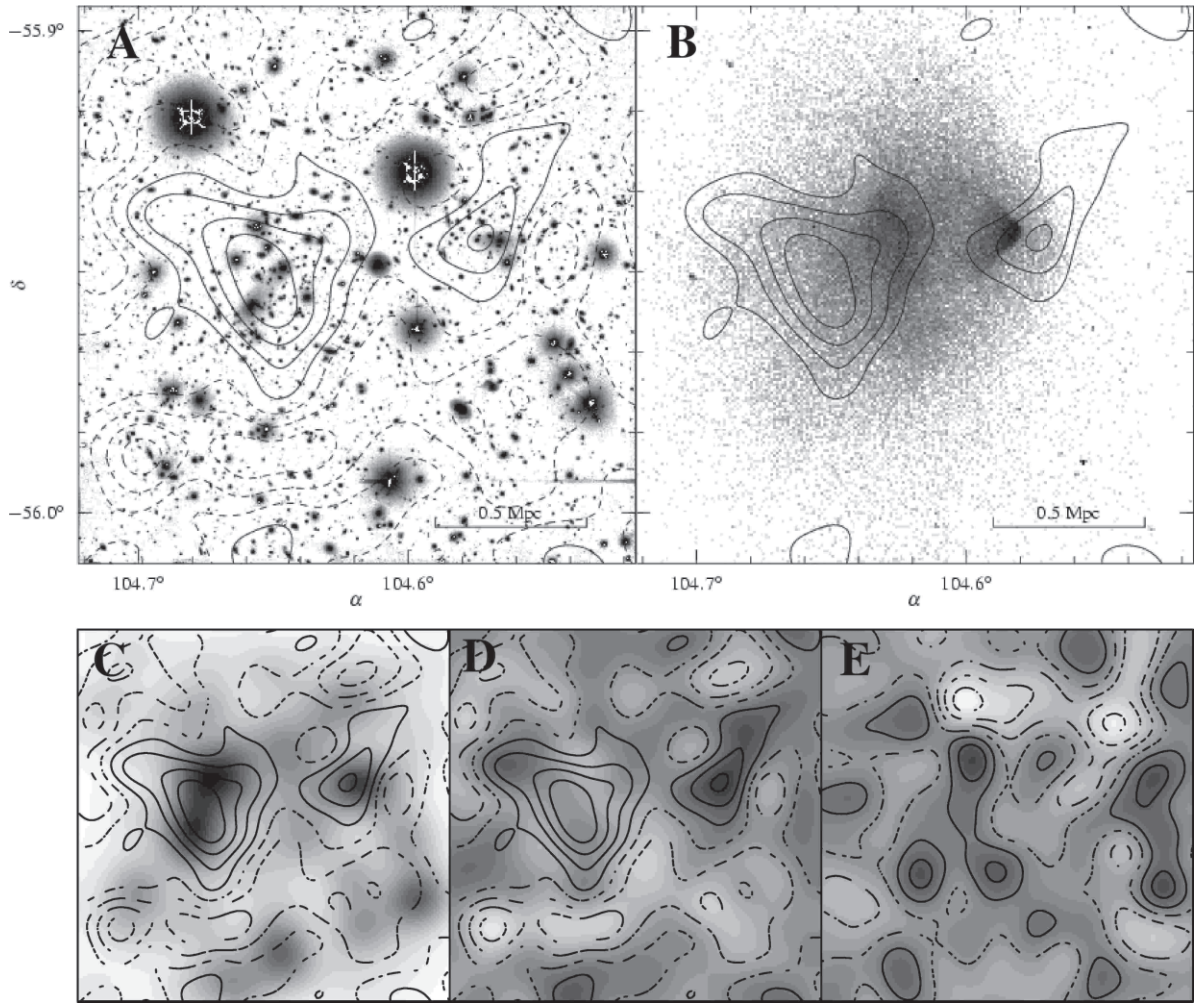

FIG. 1.—(*a*) Gray-scale *I*-band VLT image used to measure the galaxy shapes for the background galaxies. Overlain in black contours is the weak-lensing mass reconstruction, with solid contours for positive mass, dashed contours for negative mass, and the dash-dotted contour for the zero-mass level, which is set such that the mean mass at the edge of the image is zero. Each contour represents a change in the surface mass density of $2.8 \times 10^8\ M_\odot\ \mathrm{kpc}^{-2}$. (*b*) Gray-scale *Chandra* X-ray image from Markevitch et al. (2002), with the same weak-lensing contours as in (*a*). (*c*) Gray-scale luminosity distribution of galaxies with the same $B-I$ colors as the primary cluster's red sequence. Overlain are the same mass contours as in (*a*). (*d*) Gray-scale mass reconstruction of the field after subtraction of the best-fit King shear profile for the primary cluster. Overlain are the same mass contours as in (*a*). (*e*) Gray-scale mass reconstruction of the field after the background galaxies have been rotated by 45°, with the same color stretch as in (*d*). This provides a good indication of the level of the noise in the reconstruction. The contours for the noise are drawn at the same values of κ as for the mass reconstruction in (*a*).

As can be seen in Figure 1, two distinct mass peaks are found in the field, each of which is spatially coincident with an overdensity of galaxies. Spectra for galaxies in both structures have been published in Barrena et al. (2002), and the two groups have the same redshift. The peaks have significances of 6.4 σ for the larger eastern peak (hereafter referred to as "the cluster") and 3.0 σ for the smaller western peak (hereafter referred to as "the subclump"). The significances were measured by convolving the mass maps with Mexican-hat filters and comparing the filtered value at the peak position with those of randomizations of the mass maps. The randomizations were performed by first subtracting a smoothed value of the shear (smoothed using a 22.″4 Gaussian weighted average of the surrounding galaxy ellipticities) from the galaxy shear estimates to obtain an estimate of the intrinsic ellipticity of the galaxies, then applying a random spin to the orientation of each background galaxy while preserving their positions and intrinsic ellipticities, and creating mass maps from the catalogs.

An X-ray luminosity map from *Chandra* data (Markevitch et al. 2002) is overlain in gray contours in Figure 1. As can be seen, both peaks are also visible in the X-ray data but are offset in position from both the galaxies and the mass peaks. From the shape, strength, and location of the shock visible in the X-ray peak for the subclump, Markevitch et al. (2002) have concluded that this system has just undergone initial infall and pass-through and that the two clusters are now moving away from one another. The separation between the galaxies, which are effectively collisionless particles in such a pass-through event, and the X-ray gas is a result of the ram pressure of the interacting gas halos slowing down the X-ray halos during the interaction. As a result, a separation between the mass peak and the X-ray peak and an agreement in position between the mass peak and galaxy overdensity would suggest that the dark matter component of the cluster must be relatively collisionless, as compared to the X-ray–emitting baryonic gas.

In order to place limits on the collisional cross section of dark matter from the displacement of the mass peak from the X-ray peak, we calculated the error on the centroid determination of the subclump by performing mass reconstructions on 10,000 bootstrap-resampled catalogs of the background

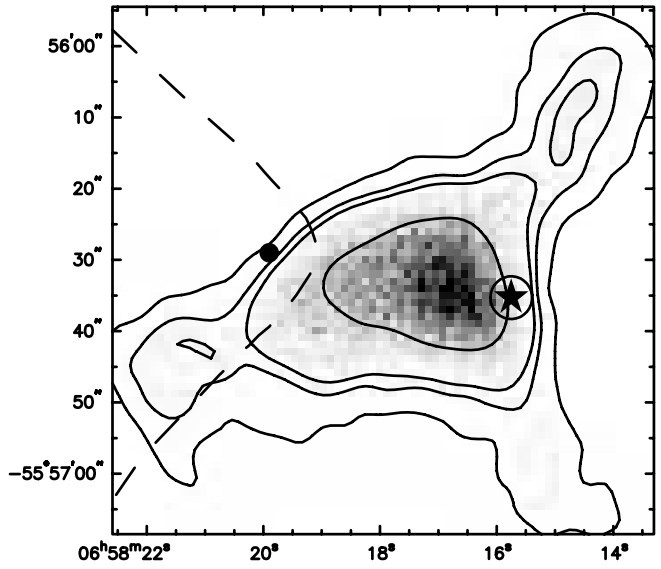

FIG. 2.—Error map for the centroid of the weak-lensing mass peak associated with the subclump, generated using 10,000 bootstrap resamplings of the background galaxy catalog. The thick black contours indicate the regions containing 68%, 90%, 95%, and 99% of the centroid positions after smoothing with a 2″ FWHM Gaussian kernel. The filled circle shows the position of the X-ray peak associated with the subclump, and the dashed contour shows the boundary of the gas associated with the subclump. The star shows the centroid of the galaxies in the subclump, with the encompassing circle showing the 1 σ error contour of the centroid.

galaxies. For each reconstruction, the κ map was convolved with a Mexican-hat filter to detect the peak nearest to the position of the subclump and measure its significance, imposing a minimum significance of 1 σ. The resulting distribution of positions is plotted in Figure 2. Since approximately 2.5% of the reconstructions should have the mass peak associated with the subclump at less than 1 σ significance, we eliminated the 250 peaks most distant from the position of the subclump in the data. The remaining peaks have an rms positional offset of 12.″1. The separation between the mass peak and X-ray peak in the data is 22.″6, which is significant at a 1.9 σ confidence level. This one-dimensional error analysis, however, is at some level incorrect, since the distribution of the peak positions is not a circular Gaussian and the resampled peak distribution has a larger rms error in right ascension than in declination. To measure the significance in the two-dimensional peak position distribution, we binned the data into $1'' \times 1''$ bins and drew contours of decreasing number of peaks until a contour intersected the position of the X-ray peak. Located inside this contour were 95.5% of the resampled peaks. We discuss constraints that this system gives on the collisional cross section of dark matter in a related paper (Markevitch et al. 2004).

The X-ray gas of the cluster is also offset from the cluster galaxies and associated dark matter peak. The dark matter peak is in good spatial agreement with the cluster galaxies, and the difference in the shape of the dark matter peak relative to the galaxy luminosity distribution seen in Figure 1 is consistent with being caused by the noise in the mass reconstruction (Clowe et al. 2000). Using the same bootstrap-resampled catalogs described above and looking for the peak nearest the position of the cluster gives the significance of the offset between the X-ray gas and dark matter as $\sim$3.4 σ. The offset gas, however, is a combination of the gas from the cluster and gas stripped from the outskirts of the subclump, and therefore requires more complicated physics to interpret.

Because the KS93 mass reconstruction can measure only the mass relative to the mean mass at the edge of the field and the images are smaller than the expected dynamical size of the cluster (6.′7 = 1770 kpc), one cannot measure the mass of the cluster reliably with the mass reconstruction in Figure 1. Instead, we have measured the mass of the cluster using radial shear profile fitting, in which one assumes a surface mass model for the cluster and converts this into a κ profile, and then into a profile for the reduced shear, which is compared with the azimuthally averaged shear profile from the data, as shown in Figure 3. We tried fitting a singular isothermal sphere (SIS), an NFW model (Navarro et al. 1996), and a King model to the data and found that the King model was marginally preferred over the NFW model, as measured by the $\delta\chi^2$ between the model reduced shear profile and the data. Using an F-test (Bevington & Robinson 1992) to compare the one-parameter SIS with the two-parameter NFW and King models resulted in both the NFW and King models being preferred to an SIS at 91% confidence. We excluded a 1′ diameter region around the subclump from the shear profile in order to minimize any contamination of the profile from the subclump. Even with this excluded area, however, the fit still includes the mass of the subclump in the total mass of the cluster for radii larger than the subclump-cluster separation, which will have the effect of overestimating the total mass of just the cluster itself, as well as underestimating the concentration of the cluster. It should also be noted that at small smoothing lengths, the mass reconstruction of the field shows two mass peaks for the main cluster, and thus the fact that the King core-model profile is the preferred mass profile may be due to the blending of two peaks in the radial profile rather than a core in a single peak.

The King model has a mass density profile

$$\rho(r) = \frac{\rho_0}{\left(1 + r^2/r_c^2\right)^{3/2}}, \tag{3}$$

which integrates to have a surface density profile

$$\Sigma(x) = \frac{2\rho_0 r_c}{1 + x^2/r_c^2}, \tag{4}$$

where ρ_0, the central mass density, and r_c, the core radius, are the fitting parameters, r is the three-dimensional radius, and x is

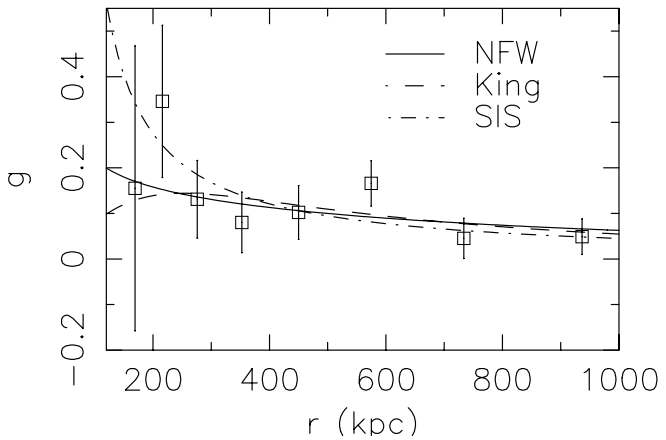

FIG. 3.—Reduced shear profile for the main cluster in the 1E 0657−558 system. Also shown are the reduced shear profiles for the best-fit NFW (*solid line*), King (*dashed line*), and SIS (*dash-dotted line*) models.

the two-dimensional projected radius. The integrated NFW profile can be found in Bartelmann (1996). The best-fit parameters were $\rho_0 = 3.85 \times 10^6\ M_\odot\ \mathrm{kpc}^{-3}$, $r_c = 214$ kpc for the King model and $r_{200} = 2250$ kpc, $c = 3.0$ for the NFW profile. Both models have the two parameters degenerate in the fits, with poor constraints on both c and r_c. The significances of the fits, as measured by the $\delta\chi^2$ between the model fit and a zero-mass model fit, are 6.48 for the King model and 6.37 for the NFW model. The King model has surface mass measurements of $(9.5 \pm 1.5) \times 10^{13}$, $(2.0 \pm 0.3) \times 10^{14}$, and $(4.4 \pm 0.7) \times 10^{14}\ M_\odot$ for 150, 250, and 500 kpc, respectively. The NFW model has surface mass measurements of $(1.02 \pm 0.16) \times 10^{14}$, $(2.1 \pm 0.3) \times 10^{14}$, and $(5.3 \pm 0.8) \times 10^{14}\ M_\odot$ for the same radii. These masses are in good agreement with the velocity dispersion for early-type galaxies given by Barrena et al. (2002).

In order to measure the mass of the subclump, we first had to remove the mass of the main cluster, which would otherwise provide a large positive bias to the mass measurements. This was accomplished by subtracting the reduced shear profile of the best-fit King model for the cluster from the background galaxy catalog. A mass reconstruction of this catalog shows that the main cluster has been effectively removed from the lensing signal, as can be seen in Figure 1. We fitted the three mass model profiles to the subclump shear profile, and while the King model was the preferred model, its χ^2 indicated that it was not a good fit to the data. Instead, we measured the mass of the subclump using aperture densitometry (Clowe et al. 2000; Fahlman et al. 1994), which measures the mean surface density inside a cylinder of a given radius minus the mean surface density in an annular region.

Using a 150 kpc radius for the disk and an annular region with a radial extent of 640–706 kpc resulted in a mass measurement of $(7.3 \pm 2.1) \times 10^{13}\ M_\odot$ when centered on the middle of the subclump mass peak seen in Figure 1 and $(6.6 \pm 1.9) \times 10^{13}\ M_\odot$ when centered on the centroid of the red subclump galaxies. The choice of a 150 kpc radius for the mass measurement was made for two reasons: First, the mass profile for the subclump, shown in Figure 4, shows evidence for a plateau in mass associated with the subclump between 150 and 250 kpc. This might indicate that the dark matter at larger radii has been tidally stripped from the core during the interaction. Second, the noise level in the aperture densitometry increases dramatically with radius, so the signal-to-noise ratio of the mass measurement changes from more than 3 σ at 100 kpc $\lesssim r \lesssim$ 150 kpc to less than 1.5 σ at $r \geq 300$ kpc.

The observed shear and derived mass of the subclump are significantly higher than could be produced by an SIS with a 212 km s^{-1} velocity dispersion, as measured by Barrena et al. (2002). This velocity dispersion, however, is measured from only seven galaxies and could be biased low by their method for distinguishing a cluster galaxy from a subclump galaxy. Furthermore, the conversion of a velocity dispersion to a mass measurement requires the assumption of virial equilibrium, which is unlikely to apply to the subclump. The weak-lensing mass is in good agreement with the 7 keV X-ray temperature for the cold gas blob, under the assumption that this was the temperature of the subclump prior to the interaction.

3. PHOTOMETRY

3.1. *Centroid of the Galaxy Distribution*

For the standard CDM paradigm, we expect the dark matter distribution to be coincident with the galaxy distribution if the dark matter particles are collisionless. To test this expectation, we employ adaptive kernel smoothing to determine the centroid of the galaxy distribution associated with the lower mass subcluster. The method is similar to that described in Gonzalez et al. (2002). An Epanechikov kernel with $h = 30''$ is used for the adaptive smoothing, and the highest density peak within a $50''$ radius is identified as the subcluster. We restrict the input catalog to objects with $m_I = 18$–25, $B-I$ color within 0.5 mag of the brightest cluster galaxy (BCG) ($B-I = 3.9$), and SExtractor stellarity indexes less than 0.5 in both bands. We compute the number-weighted and flux-weighted centroids, finding that both weighting schemes yield indistinguishable results. The uncertainty is calculated by recomputing the centroid for 1000 bootstrap-resampled catalogs.

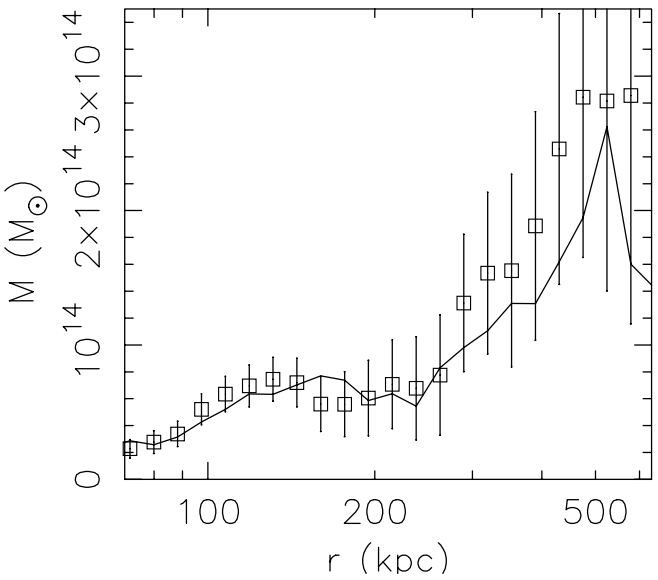

FIG. 4.—Minimum surface mass profile for the subclump, generated using aperture densitometry centered on the centroid of the galaxy distribution. The solid line shows the mass profile when centered on the mass peak in the two-dimensional mass reconstruction. The error bars for aperture densitometry are correlated such that every data point has knowledge of the values of the points at larger radius.

We find that the peak of the galaxy distribution is coincident with the location of the BCG. The number-weighted centroid is R.A. = $06^{\mathrm{h}}58^{\mathrm{m}}15\overset{\mathrm{s}}{.}66$, decl. = $-55^\circ56'35\overset{''}{.}3$ (J2000.0) with a $1\overset{''}{.}9$ 1 σ Gaussian uncertainty, while the flux-weighted centroid is R.A. = $06^{\mathrm{h}}58^{\mathrm{m}}15\overset{\mathrm{s}}{.}75$, decl. = $-55^\circ56'35\overset{''}{.}3$ (J2000.0) with a $3\overset{''}{.}0$ 1 σ uncertainty. The net separation between the galaxy and weak-lensing centroids is $12\overset{''}{.}3$, which has a significance of 1 σ in the one-dimensional error analysis and 70% in the two-dimensional error analysis described in § 2.

3.2. *Luminosity*

The luminosity is determined using two approaches. We first derive the luminosity from the galaxy catalog, which is the typical method used in cluster M/L measurements. We then directly compute the luminosity by integrating the total flux within the aperture, which places a hard upper limit on the luminosity. For both techniques we correct the derived luminosities for extinction (Schlegel, Finkbeiner, & Davis 1998) and apply $e + K$ corrections to convert to rest-frame passbands. The applied corrections are based on the Bruzual A. and Charlot evolutionary models for a passively evolving elliptical galaxy (Bruzual A. & Charlot 1993; Charlot, Worthey, &

TABLE 1
MASS-TO-LIGHT RATIOS

Region	R (kpc)	M ($10^{14}\ M_\odot$)	L_B ($10^{11}\ L_\odot$)	L_I ($10^{11}\ L_\odot$)	M/L_B ($M_\odot\ L_\odot^{-1}$)	M/L_I ($M_\odot\ L_\odot^{-1}$)
Integrated Flux Technique						
Subcluster.....	150	0.66 ± 0.19	2.1	4.9	314 ± 90	135 ± 39
Main	150	0.95 ± 0.15	3.5	7.8	271 ± 43	122 ± 19
Main	250	2.0 ± 0.3	8.5	15.8	235 ± 35	127 ± 19
Main	500	4.4 ± 0.7	17.4	32.5	253 ± 40	135 ± 22
Catalog Technique						
Subcluster.....	150		2.4	3.8	275 ± 79	174 ± 50
Main	150		3.2	6.3	297 ± 47	151 ± 19
Main	250		8.1	12.9	247 ± 37	155 ± 23
Main	500		21.4	28.8	206 ± 32	152 ± 24

NOTE.—Because of the different methods in which the main cluster and subclump masses were measured, the subclump M/L ratios must be considered a bound when comparing with the main cluster.

Bressan 1996). The integrated $B-I$ color is consistent with the prediction for passive evolution, indicating that this approximation is reasonable. Absolute magnitudes are converted to units of solar luminosity using the solar absolute magnitudes given in Allen & Cox (2000).

For the catalog approach, we first cull the input photometric catalog to minimize contamination from stars and foreground galaxies. We exclude all objects that are brighter than the BCG, are more than 1 mag redder than the BCG (i.e., $B-I > 4.9$), or have stellarity index greater than 0.8 and $m_I < 20$. Two foreground spirals are also removed from the input catalog. The flux from the remaining galaxies is then summed, with the resulting luminosity shown in Table 1. We caution that there are two caveats with this approach. First, we lack sufficient spatial coverage to employ background subtraction. The impact of background contamination is expected to be minor because of the large density contrast within our physical apertures, but such contamination will yield a positive bias in the derived luminosity. This effect will be greatest for the largest apertures. Second, incompleteness at the faint end of the luminosity function yields a negative bias in the derived luminosity. Because our data are complete to roughly 5 and 6 mag below L_* in I and B, respectively, the magnitude of this effect should be 4.5% and 2.5%, respectively, for a faint-end slope $\alpha = -1.3$.

Directly integrating the flux within the apertures provides a useful cross-check on the above technique. For this approach we mask only stars with stellarity index greater than 0.8 and $m_I < 18$. We then integrate the flux within the aperture, using two additional apertures located $2'$ north and south of the cluster to quantify the background sky level. The sky levels in these apertures were computed after masking detected objects in the region, and thus the detected background level does not include flux associated with resolved galaxies (except that scattered onto the extreme wings of the PSF). Background-galaxy contamination is expected to yield a slight positive bias in the derived luminosity, as with the catalog approach. The results are shown in Table 1. The luminosities derived via this method typically agree with the catalog results to within 15%.

3.3. *Mass-to-Light Ratios*

We determine the mass-to-light ratios of both the main cluster and the subcluster in rest-frame B and I. If the subcluster has suffered significant mass loss during the passage through the core of the main cluster, this should be reflected by a decreased mass-to-light ratio. We find no evidence for such mass loss, with the M/L ratios for both components consistent with one another to within the 1 σ uncertainties. This result implies, under the assumption that the initial M/L ratios for the two structures were similar, that the dark matter interaction cross section must be small, a topic that is explored in greater detail in a related paper (Markevitch et al. 2004). We further note that the derived M/L values are consistent with other recent lensing-derived mass-to-light ratios. Dahle (2000), for instance, finds $M/L_B = 259 \pm 12$ for a sample of 40 low-redshift clusters. The low dispersion of the M/L ratios in the Dahle (2000) sample also suggests that the assumption of similar intrinsic M/L ratios for the two components in this cluster is not unreasonable.

For these mass-to-light ratio calculations for the subclump, we use an aperture centered on the centroid of the galaxy light distribution, assuming that the offset in the mass peak seen in Figure 1 is a result of noise in the shear field. If instead we use an aperture centered on the observed mass peak for the subclump, the mass-to-light ratio of the subclump increases by $\sim$10%. In addition, the mass estimates for the subclump are created by subtracting the mean surface density in a 640–760 kpc annulus from the mean surface density within the 150 kpc disk, and since no similar subtraction is performed on the light or the cluster mass, the mass-to-light ratio of the subclump must be considered a minimum value in making comparisons with the main cluster.

4. DISCUSSION AND CONCLUSIONS

In a CDM universe, one would expect that the mass peaks for the cluster and subclump would agree with the centroid of the galaxy distributions, since both galaxies and dark matter particles are collisionless in such an interaction (Tormen et al. 2003). One would also expect that the mass-to-light ratios would decrease by $\sim$10%–15% as compared to relaxed systems because of the baryonic X-ray halo mass being removed from the structures. Such a scenario is, within errors, in good agreement with the data.

In a purely baryonic MOND universe, the X-ray and galaxy centroids would still be separated, since the galaxies still act as collisionless particles in the interaction. However, because the

X-ray halo is the dominant mass component of the visible baryons in the cluster, in the absence of a dark mass component the vast majority, ~85%–90%, of the mass of the subclump would be with the X-ray gas. Thus, any direct method to measure the mass of the system would detect a higher mass about the stripped X-ray halo than around the galaxies. This is not what is observed in this system. In order to quantify how much these observations disagree with MOND, however, we first need to determine a method to measure the masses of the clusters in a MOND cosmology.

Unfortunately, because there is not a derivation of MOND from general relativity, there is not a definitive way to measure a mass with weak lensing from a measured shear field. If one assumes that the relation from general relativity between the deflection of a photon or a massive particle moving at the speed of light and a static gravitational field is unchanged by MOND, then it can be shown that the shear field caused by a point mass is

$$\gamma(\theta) = \frac{\theta_{\rm E}^2}{\theta_0 \theta^2}\left[\frac{\theta}{2} + \theta_0 - \frac{\theta^3}{2(\theta + \theta_{\rm out})^2}\right], \qquad (5)$$

where θ is the distance from the point mass, $\theta_{\rm E}$ is the Einstein radius for the lens, θ_0 is the distance at which the gravitational acceleration changes from Newtonian to MOND, and $\theta_{\rm out}$ is the distance at which the gravitational acceleration changes back to a θ^{-2} law (Mortlock & Turner 2001; Hoekstra et al. 2002). Since both $\theta_{\rm E}$ and θ_0 scale as the square root of the point mass, the resulting shear profile scales linearly with the mass for $\theta \ll \theta_0$ and $\theta \gg \theta_{\rm out}$, as the square root of the mass for $\theta_0 \ll \theta \ll \theta_{\rm out}$, and somewhere between the two extremes for the transitional regions, $\theta \sim \theta_0$ and $\theta \sim \theta_{\rm out}$.

Calculating the expected shear profile for an extended source in the MOND regime is complicated by the lack of a thin-lens approximation, which is used to simplify the equations with Newtonian gravity (Mortlock & Turner 2001). However, it is reasonable to assume that the same general relation between the gravitational shear field and overall mass of a halo exists as per the point mass relation. As such, the level of weak shear produced by a cluster of galaxies at a radius of a few hundred kiloparsecs from the cluster core (which would be between θ_0 and $\theta_{\rm out}$ for 10^{13}–10^{16} $M_\odot$ clusters) should scale with the mass of the cluster, probably somewhere between a linear scale and a scale with the square root of the mass.

From observations, we know that the shear fields produced by individual galaxies (Hoekstra et al. 2003; McKay et al. 2002) are an order of magnitude lower than those produced by galaxy groups (Hoekstra et al. 2001), which are an order of magnitude lower than those produced by poor clusters (Wittman et al. 2000), which are in turn significantly lower than those produced by rich clusters (Dahle et al. 2002; Clowe et al. 2000; Clowe & Schneider 2001, 2002). Since the amount of visible baryons in these structures scale in a similar manner, then from these observations we have support for the above assumptions. Thus, in a MOND universe, one should still observe a change in the shear field of a structure with a change in the mass of the structure.

As a result, if the mass of a cluster of galaxies is limited solely to visible baryons, then by removing the X-ray halo from the cluster, one should reduce the gravitational shear centered on the galaxies by at least a factor of 3 if the shear scales as the square root of the mass and up to a factor of 10 if the shear scales linearly with the mass. In this system, however, we find that the ratio of the gravitational shear to visible light for the two components, where the X-ray halo is stripped from the galaxies, is consistent with that found in normal clusters, which have the X-ray halo and galaxies spatially coincident. This is inconsistent with the shear scaling as the square root of the mass in the MOND model at roughly a 2 σ confidence level and with the shear scaling linearly with the mass at roughly a 3 σ confidence level.

In order to reduce the inconsistency with the data to a ~1 σ confidence level, one would need to add a nonluminous mass component to the clusters that is equal to the mass of luminous matter for the case of the shear scaling as the square root of the mass and 2.5 times the mass of luminous matter for the case of the shear scaling linearly with the mass. This extra mass component would also reduce the problem with the detected mass peak for the subclump being closer to the galaxies than the X-ray halo, since the detected signal would be a blend of the two components, because of the required smoothing of the mass map.

The more significant offset between the cluster mass peak and X-ray halo would require a greater amount of dark mass to explain if the two components were cleanly separated. The X-ray halo, however, is extended over the cluster galaxies, which may indicate that some fraction of the X-ray gas has already been drawn back to the galaxy position.

Any dark mass component of the system must be relatively collisionless so it can undergo the pass-through without loss of velocity or mass and be able to clump on scales smaller than 100 kpc (the smallest aperture for which we can reliably measure the shear about the subclump). Adopting big bang nucleosynthesis limits on the mean baryonic mass of the universe excludes most of this mass from being baryons in cold, condensed structures. The clumpiness limit excludes the matter from being massive neutrinos with masses less than 4.5 eV (Sanders 2003; Tremaine & Gunn 1979). Since neutrinos more massive than 2.2 eV have been ruled out experimentally (Bonn et al. 2002; Lobashev et al. 2001), neutrinos thus cannot explain this mass.

4.1. *Summary*

We have shown above that the cluster 1E 0657−558 has a lower mass subclump visible in X-ray and optical observations, as well as in a weak-lensing mass reconstruction. The X-ray and optically luminous components are spatially separated at high significance, as one would expect for a system that has just undergone an initial infall and transit of a larger mass system (e.g., Tormen et al. 2003). The observed mass peak in the weak-lensing reconstruction lies between the X-ray and optical components but is closer to, and consistent with, the optical component. The centroid of the subclump mass peak has a fairly large error, resulting in the offset of the mass peak from the centroid of the galaxy distribution with a ~70% confidence level and the offset of the mass peak from the X-ray peak with a ~95% confidence level.

The primary cluster has also been detected in the weak-lensing mass reconstruction and has a mass peak that is spatially coincident with the cluster galaxies. The X-ray gas from the main cluster is offset from the mass peak at a 3.4 σ significance.

We have also measured the mass-to-light ratio for the subclump at a 150 kpc radius and for the main cluster at 150, 250, and 500 kpc radii. We find that the subclump has a

mass-to-light ratio that is consistent with the mass-to-light ratio of the main cluster and that both are consistent with mass-to-light ratios for relaxed clusters. The dominant source of error in the mass-to-light ratios and the mass–X-ray gas offsets is the weak-lensing mass reconstructions, which can be improved by obtaining shear information on a wider field than the $7' \times 7'$ VLT field and/or by obtaining deeper imaging on the same field with a smaller PSF in order to greatly increase the number density of background galaxies usable for the measurement of the shear field.

Finally, we have argued that even in a MOND universe, a significant fraction of the original mass of the subclump must exist in the form of dark matter, which furthermore should be nonbaryonic and nonneutrino. The exact amount of extra mass cannot be calculated because of the lack of a MOND derivation from general relativity, but phenomenological arguments suggest that it is at least equal to the baryonic mass of the cluster. While these observations cannot disprove MOND, or alternatively prove that gravity is Newtonian on small acceleration scales, they remove its primary motivation of avoiding the notion of dark matter.

We wish to thank Oliver Czoske and Alexey Vikhlinin for useful discussions. This work was supported by the Deutsche Forschungsgemeinschaft under the project SCHN 342/3-1 (DC). A. H. G. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST 04-07485. M. M. received support through NASA contract NAS8-39073, *Chandra* grant GO2-3165X, and the Smithsonian Institution.

REFERENCES

Allen, C. W., & Cox, A. N. 2000, Allen's Astrophysical Quantities (4th ed.; New York: Springer)
Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2002, MNRAS, 334, L11
Barrena, R., Biviano, A., Ramella, M., Falco, E. E., & Seitz, S. 2002, A&A, 386, 816
Bartelmann, M. 1996, A&A, 313, 697
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bevington, P. R., & Robinson, D. K. 1992, Data Reduction and Error Analysis for the Physical Sciences (New York: McGraw-Hill)
Bonn, J., et al. 2002, Nucl. Phys. B Suppl., 110, 395
Bruzual A., G., & Charlot, S. 1993, ApJ, 405, 538
Charlot, S., Worthey, G., & Bressan, A. 1996, ApJ, 457, 625
Clowe, D., Luppino, G. A., Kaiser, N., & Gioia, I. M. 2000, ApJ, 539, 540
Clowe, D., & Schneider, P. 2001, A&A, 379, 384
———. 2002, A&A, 395, 385
Dahle, H. 2000, in The NOT in the 2000's, ed. N. Bergvall, L. O. Takalo, & V. Piirola (Piikkiö: Univ. Turku), 45
Dahle, H., Kaiser, N., Irgens, R. J., Lilje, P. B., & Maddox, S. J. 2002, ApJS, 139, 313
Fahlman, G., Kaiser, N., Squires, G., & Woods, D. 1994, ApJ, 437, 56
Fontana, A., et al. 1999, A&A, 343, L19
Gavazzi, R. 2002, NewA Rev., 46, 783
Gonzalez, A. H., Zaritsky, D., Simard, L., Clowe, D., & White, S. D. M. 2002, ApJ, 579, 577
Hoekstra, H., Franx, M., & Kuijken, K. 2000, ApJ, 532, 88
Hoekstra, H., Franx, M., Kuijken, K., Carlberg, R. G., & Yee, H. K. C. 2003, MNRAS, 340, 609
Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, NewA Rev., 46, 767
Hoekstra, H., et al. 2001, ApJ, 548, L5
Jeltema, T. E., Canizares, C. R., Bautz, M. W., Malm, M. R., Donahue, M., & Garmire, G. P. 2001, ApJ, 562, 124
Kaiser, N., & Squires, G. 1993, ApJ, 404, 441 (KS93)
Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460
Lobashev, V. M., et al. 2001, Nucl. Phys. B Suppl., 91, 280
Markevitch, M., et al. 2002, ApJ, 567, L27
———. 2004, ApJ, 606, in press
Marshall, P. J., Hobson, M. P., Gull, S. F., & Bridle, S. L. 2002, MNRAS, 335, 1037
McGaugh, S. S., & de Blok, W. J. G. 1998, ApJ, 499, 66
McKay, T. A., et al. 2002, ApJ, 571, L85
Milgrom, M. 1983, ApJ, 270, 365
Mortlock, D. J., & Turner, E. L. 2001, MNRAS, 327, 557
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563 (NFW)
Sanders, R. H. 2003, MNRAS, 342, 901
Sanders, R. H., & McGaugh, S. S. 2002, ARA&A, 40, 263
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Seitz, C., & Schneider, P. 1995, A&A, 297, 287
Tormen, G., Moscardini, L., & Yoshida, N. 2003, MNRAS, in press
Tremaine, S., & Gunn, J. E. 1979, Phys. Rev. Lett., 42, 407
Tucker, W. H., Tananbaum, H., & Remillard, R. A. 1995, ApJ, 444, 532
Tucker, W., et al. 1998, ApJ, 496, L5
Wittman, D., dell'Antonio, I., Tyson, T., Bernstein, G., Fischer, P., & Smith, D. 2000, in Constructing the Universe with Clusters of Galaxies, ed. F. Durret & D. Gerbal (Paris: IAP), 60