MC^2: CONSTRAINING THE DARK MATTER DISTRIBUTION OF THE VIOLENT MERGING GALAXY CLUSTER CIZA J2242.8+5301 BY PIERCING THROUGH THE MILKY WAY

M. JAMES JEE[1], ANDRA STROE[2], WILLIAM DAWSON[3], DAVID WITTMAN[1], HENK HOEKSTRA[2],
MARCUS BRÜGGEN[4], HUUB RÖTTGERING[2], DAVID SOBRAL[2,5,6], AND REINOUT J. VAN WEEREN[7]

[1] Department of Physics, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA
[2] Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands
[3] Lawrence Livermore National Laboratory, P.O. Box 808 L-210, Livermore, CA 94551, USA
[4] Hamburger Sternwarte, Gojenbergsweg 112, D-21029 Hamburg, Germany
[5] Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, PT1349-018 Lisboa, Portugal
[6] Centro de Astronomia e Astrofísica da Universidade de Lisboa, Observatório Astronómico de Lisboa, Tapada da Ajuda, 1349-018 Lisboa, Portugal
[7] Harvard Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
Received 2014 October 10; accepted 2014 December 26; published 2015 March 20

ABSTRACT

The galaxy cluster CIZA J2242.8+5301 at $z = 0.19$ is a merging system with a prominent ($\sim$2 Mpc long) radio relic, which together with the morphology of the X-ray emission provides strong evidence for a violent collision along the north–south axis. We present our constraints on the dark matter distribution of this unusual system using Subaru and Canada–France–Hawaii Telescope imaging data. Measuring a high signal-to-noise ratio lensing signal from this cluster is potentially a challenging task because of its proximity to the Milky Way plane ($|b| \sim 5°$). We overcome this challenge with careful observation planning and systematics control, which enables us to successfully map the dark matter distribution of the cluster with high fidelity. The resulting mass map shows that the mass distribution of CIZA J2242.8+5301 is highly elongated along the north–south merger axis inferred from the orientation of the radio relics. Based on our mass reconstruction, we identify two sub-clusters, which coincide with the cluster galaxy distributions. We determine their masses using Markov Chain Monte Carlo analysis by simultaneously fitting two Navarro–Frenk–White halos without fixing their centroids. The resulting masses of the northern and southern systems are $M_{200} = 11.0^{+3.7}_{-3.2} \times 10^{14} \, M_\odot$ and $9.8^{+3.8}_{-2.5} \times 10^{14} \, M_\odot$, respectively, indicating that we are witnessing a post-collision of two giant systems of nearly equal mass. When the mass and galaxy centroids are compared in detail, we detect $\sim$1′ ($\sim$190 kpc) offsets in both northern and southern sub-clusters. After investigating the statistical significance of the offsets by bootstrapping both mass and galaxy centroids, we find that the galaxy luminosity-mass offset for the northern clump is statistically significant at the $\gtrsim 2\sigma$ level whereas the detection is only marginal for the southern sub-cluster in part because of a relatively large mass centroid error. We conclude that it is yet premature to uniquely attribute the galaxy-mass misalignment to self-interaction of dark matter and discuss caveats.

Key word: gravitational lensing: weak

1. INTRODUCTION

Merging galaxy clusters are receiving growing attention because of their potential to provide constraints on properties of dark matter, which is gravitationally the most dominant constituent of galaxy clusters. The capability of weak lensing in constraining the mass distribution of a foreground lens based on its gravitational impact on the shapes of background galaxies provides a unique opportunity to learn from the difference in distribution between dark matter and baryonic components (galaxies and gas). If collisionless, dark matter particles will distribute nearly in the same way as galaxies do. The famous example supporting this standard paradigm is the "Bullet Cluster" 1E0657-56 (Clowe et al. 2006), which exhibits no significant difference between the locations of the cluster galaxies and mass peaks. However, there have been reports on some exceptional cases, where weak-lensing reveals mass distributions apparently at odds with the conventional paradigm of collisionless dark matter. For example, in A520 (Mahdavi et al. 2007; Jee et al. 2012, 2014a) and A2744 (Merten et al. 2011), significant ($\gtrsim 6\sigma$ in case of A520) mass peaks without any strong concentration of luminous cluster galaxies are detected. In addition, in the case of the Musket Ball cluster (Dawson et al. 2012), a marginal detection of a weak-lensing mass peak

trailing galaxies is claimed (Dawson 2013). Kahlhoefer et al. (2014) have shown that such observational anomalies may be produced by self-interacting dark matter (SIDM). Thus they open up interesting research opportunities to study a collection of these rare systems and put meaningful constraints on the properties of dark matter.

We recently launched a Merging Cluster Collaboration[8] (MC^2; PI: W. Dawson) project with aims to enhance our understanding of cluster physics and constrain properties of dark matter from systematically analyzing a number of prominent merging clusters. Combining data from different instruments (e.g., X-ray, spectroscopy, weak-lensing, radio relic observation, etc.) will greatly reduce the parameter space that we need to explore in setting up initial conditions. The simulation of the MC^2 collaboration will include various models of dark matter (e.g., Rocha et al. 2013). Iterative comparisons between the simulations and observations will enable us to constrain properties of dark matter with unprecedented precision.

This paper is part of a upcoming series of the MC^2 publications, presenting our first weak-lensing analysis of CIZA J2242.8+5301 at $z = 0.19$. The cluster is one of the most remarkable clusters in our sample, possessing a prominent giant

---

[8] http://www.mergingclustercollaboration.org
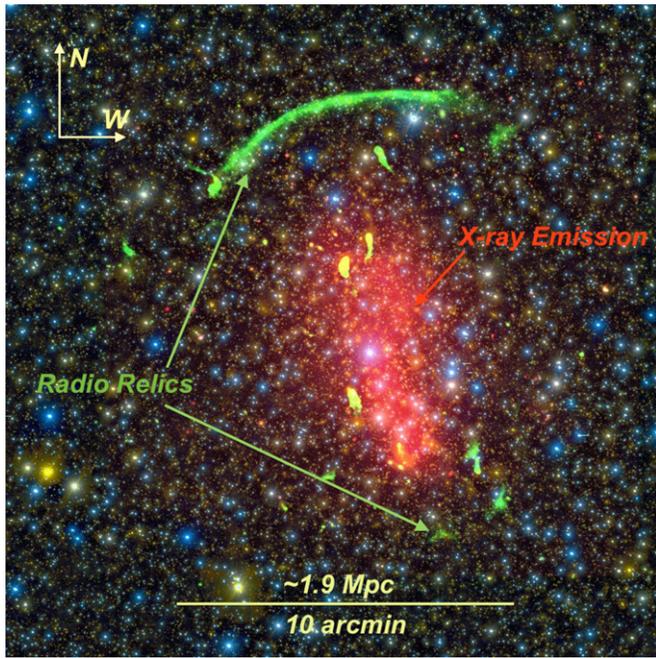
Figure 1. Illustration of different components in the merging cluster CIZA J2242.8+5301. The intensity in green represents the 610 MHz radio emission measured with GMRT (van Weeren et al. 2010). The intensity in red shows the X-ray emission observed with *Chandra*. The background color-composite is created using Subaru/Suprime Cam data with the g, $g + i$, and i filters depicting the intensity in blue, green, and red channels, respectively.

radio arc (van Weeren et al. 2010) stretched over ∼2 Mpc. Radio relics are discovered at the edge of merging galaxy clusters elongated perpendicular to the merger axis and are believed to trace shock fronts (Ensslin et al. 1998). Therefore, the observed morphology of the radio relics of CIZA J2242.8+5301 unambiguously indicates that we are observing a post-merger system colliding along the north–south axis. The north–south elongation of the X-ray emission from *XMM-Newton* and *Chandra* also supports this merger scenario (Ogrean et al. 2013, 2014). In Figure 1, we display the pseudo-color composite image of CIZA J2242.8+5301 with illustration of the radio relics and X-ray emission.

These rich data sets offer opportunities for detailed simulations of the system in order to understand the exact physical mechanism leading to the observed features of CIZA J2242.8+5301. Our weak-lensing study will provide reliable mass properties of the cluster, which are among the key input parameters of the simulation, but are currently missing. Because of the apparent disruption of the system, we need to derive the mass without relying on any dynamical assumption, which is an important merit of weak lensing probes.

We structure our paper in the following way. In Section 2, we briefly review the discovery and development of our understanding of CIZA J2242.8+5301. Section 3 describes the data and reduction method. A basic description of weak lensing theory is presented in Section 4. We present our mass reconstruction in Section 5. The cluster mass estimation is described in Section 6. Our results are discussed in Section 7 before we conclude in Section 8.

We assume a flat Λ cold dark matter (ΛCDM) cosmology with $H_0 = 70 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. At the redshift of CIZA J2242.8+5301 $z \sim 0.19$, the plate scale is ∼3.17 kpc arcsec^{-1}.

2. THE GALAXY CLUSTER CIZA J2242.8+5301: DEVELOPMENT OF UNDERSTANDING

The galaxy cluster CIZA J2242.8+5301 at $(\alpha, \delta) =$ (22$^\text{h}$42$^\text{m}$49$^\text{s}$.1, 53°00′51″) or $(l, b) = (104°.191, -5°.111)$ was discovered in the Clusters In the Zone of Avoidance (CIZA) survey as one of the 57 clusters located at low Galactic latitude (Kocevski et al. 2007). The clusters in the CIZA survey are in general difficult to identify based on galaxy overdensity because of the severe extinction and stellar obscuration. Kocevski et al. (2007) report that the X-ray luminosity from the original *ROSAT* All Sky Survey data is $L_X = 6.8 \times 10^{44} \, \text{erg} \, \text{s}^{-1}$, which is converted to $M_{500} \sim 5.5 \times 10^{14} \, M_\odot$ according to the scaling relation of Pratt et al. (2009).

The spectacular radio relics of the cluster were discovered by van Weeren et al. (2010) from Westerbork Synthesis Radio Telescope (Katgert et al. 1973) and Giant Metrewave Radio Telescope (GMRT; Swarup et al. 1991) data with a clear detection of a spectral index gradient toward the cluster center, which led the authors to conclude that the radio signal is arising from electrons accelerated at large-scale shocks due to a head-on binary merger of roughly equal masses. They also argued that the fraction of the polarization constrains the angle between the plane of the sky and the merger axis to be less than ∼30°.

Based on the constraints imposed by the radio data, van Weeren et al. (2011) carried out hydrodynamical simulations to determine the merger geometry. They concluded that CIZA J2242.8+5301 might be undergoing a merger with a mass ratio of 2:1 and an impact parameter less than 400 kpc nearly in the plane of the sky ($<10°$). The estimated core passthrough time is about 1 Gyr ago. As we will discuss in Section 7, we believe that the simulation set-up of van Weeren et al. (2011) should be revised because the total mass in the two systems there is assumed to be only $5.5 \times 10^{14} \, M_\odot$, which is based on the estimation of Kocevski et al. (2007), who converted the *ROSAT* X-ray luminosity to the mass with the scaling relation of Pratt et al. (2009). Our weak-lensing analysis shows that the total mass of CIZA J2242.8+5301 well exceeds ∼$2 \times 10^{15} \, M_\odot$ (Section 6).

Additional evidence supporting the north–south merger scenario is the presence of tailed radio sources. Stroe et al. (2013) found that the tailed radio sources in the CIZA J2242.8+5301 field are pointing either north or south, consistent with the direction of the merger hypothesis.

The *ROSAT* Position Sensitive Proportional Counter image already showed that the intracluster medium (ICM) of CIZA J2242.8+5301 is severely disturbed (van Weeren et al. 2010). A more detailed study of the ICM was presented by Ogrean et al. (2013) using *XMM-Newton* observations. The X-ray emission clearly shows a north–south elongation of the cluster ICM, which is consistent with the merger axis inferred from the radio relics and numerical simulations (van Weeren et al. 2010, 2011). However, no evidence for shock compression is found near the northern relic in the *XMM-Newton* data. A weak indication of shock is present near the southern relic, although the feature cannot be exclusively attributed to a shock. Ogrean et al. (2013) concluded that many features in their X-ray data including temperature gradient, surface brightness distribution, etc. suggest that perhaps the actual merger is much more complicated than a simple binary merger, which was the initial hypothesis of van Weeren et al. (2010, 2011).

With the *Suzaku* X-Ray Telescope, Akamatsu & Kawahara (2013) found significant jumps in temperature at the position

of the radio relics in CIZA J2242.8+5301, which is the first confirmation that the radio relics in the cluster indeed traces the location of the shocks. The implied Mach number from the temperature drop ($\sim$8 keV to $\sim$2 kev) is $\mathcal{M}_X = 3.15 \pm 0.52$, which is slightly lower than, but statistically consistent with the radio measurement $\mathcal{M}_{\mathrm{radio}} = 4.6^{+1.3}_{-0.9}$ derived from the spectral index (van Weeren et al. 2010).

Higher resolution X-ray data from *Chandra* were analyzed by Ogrean et al. (2014). Their investigation of the surface brightness profile does not provide any evidence of a shock near the northern relic in agreement with their previous study with *XMM-Newton* (Ogrean et al. 2013). Instead, Ogrean et al. (2014) report that multiple density discontinuities are present in other regions both on and off the hypothesized merger axis, speculating that the X-ray shock features arise, if real, from violent relaxation of the dark matter cores of the clusters. The presence of the large temperature discontinuity of Akamatsu & Kawahara (2013) was confirmed by the reanalysis of the *Suzaku* data by Ogrean et al. (2014).

Stroe et al. (2014a) report that the relic features of CIZA J2242.8+5301 are detected even at high-frequency ($\sim$16 GHz) using the Arcminute Microkelvin Imager (Zwart et al. 2008). Interestingly, the high-frequency data detects a clear trend of spectral steepening toward 16 GHz. However, this finding is somewhat inconsistent with the diffusive shock acceleration model, which is believed to be the main physical mechanism responsible for most radio relics in merging clusters. Another noteworthy feature in the high-frequency data is a flat-spectrum diffuse extension of the southern relic, which was not present in the lower frequency maps of van Weeren et al. (2010) and Stroe et al. (2013).

The Hα survey of the CIZA J2242.8+5301 field (Stroe et al. 2014b, 2014c) shows a significant excess of Hα sources with respect to the blank field. Possibly, this excess is linked to the enhanced star formation activities triggered by the passage of the shock wave.

3. OBSERVATIONS

3.1. Subaru/SuprimeCam Data Reduction

CIZA J2242.8+5301 was observed with Subaru/Suprime Cam on 2013 July 13 in *g* and *i* with integrations of 720 s (four visits) and 3313 s (nine visits), respectively (PI: D. Wittman). We rotated the field of each visit in order to distribute the bleeding trails and diffraction spikes from bright stars azimuthally and later removed them by median-stacking different visits. This scheme enables us to maximize the number of galaxies usable for lensing. The median seeing for *g* and *i* are 0$\farcs$72 and 0$\farcs$65, respectively.

We used the SDFRED (Ouchi et al. 2004; Yagi et al. 2002) package to subtract over-scan and bias, make flats, correct for geometric distortion (third order polynomial), and mask out regions affected by bad pixels and auto guide probe. Then, we removed residual distortion and refined the astrometric solution with the SCAMP software (Bertin 2006),[9] which compares the coordinates of the common objects in individual frames with those in a reference catalog. From our experience, we find that the Sloan Digital Sky Survey Data Release 7 (SDSS DR7; Abazajian et al. 2009) reference catalog in general provides the most reliable astrometric solution if the field is observed with SDSS. However, the CIZA J2242.8+5301 field is not within the

SDSS survey. Nevertheless, we were still able to achieve weak-lensing quality accuracy (rms $\sim$ 0.02 pixel) in image registration utilizing the Two Micron All Sky Survey (Skrutskie et al. 2006) catalog.

Finally, we created a large mosaic $\sim$48$'$ $\times$ 46$'$ image with the SWARP package (Bertin et al. 2002)[10] in two passes. In the first pass, we generated a median-stack image. In the second pass, we created a coadd stack by weight-averaging input frames. Because it is necessary to identify the pixels that should be excluded in co-addition, we masked them out by modifying the weight files of input frames after identifying them by comparing pixel values of input frames with those of the median stack; the current version (v2.19.1) of SWARP does not automate this procedure.

3.2. CFHT/MegaCam Data Reduction

The cluster image was taken with Canada–France–Hawaii Telescope (CFHT)/MegaCam during the 2013 July 3–12 period in *r* (PI: A. Stroe). The total integration is 24,000 s, consisting of 40 short (600 s) exposures. The median seeing is $\sim$0$\farcs$74. The field rotation mentioned above for the Subaru observation was not implemented because of the technical restriction of the instrument. Raw level data reduction was carried out with the Elixir pipeline.[11] However, similarly to the Subaru data, we used a combination of SCAMP and SWARP to generate our final mosaic.

3.3. Spectroscopic Data

We conducted two spectroscopic surveys with the Keck DEep Imaging Multi-Object Spectrograph (DEIMOS; Faber et al. 2003) and the William Hershel Telescope AF2. We refer readers to Dawson et al. (2014) for details of the observation and data processing. The Keck/DEIMOS spectroscopic data taken on 2013 July 14 and 2013 September 5 provided a total of 505 spectra, out of which we obtained reliable redshifts for 447 objects. We targeted 73 objects with AF2. These AF2 targets are Hα candidates belonging to the cluster redshift (Stroe et al. 2014c). Of those 73 objects, we determined redshifts for 19 objects within the $\sim$15$'$ radius of the cluster center. The two spectroscopic surveys provided a total of 217 spectroscopic cluster member redshifts.

3.4. Object Detection, Photometry, and Extinction Correction

We run SExtractor (Bertin & Arnouts 1996) in a dual image mode using the *i*-band image for detection. Prior to object detection, we choose to let SExractor convolve the *i*-band image with a Gaussian kernel whose size approximately matches the seeing. Objects are identified by looking for at least five connected pixels above two times the sky rms. The blending threshold parameter (BLEND_NTHRESH) is set to 64 with a minimal contrast of DEBLEND_MINCONT = 10^{-4}. We employ isophotal magnitudes (MAG_ISO) to estimate object colors whereas total magnitude (MAG_AUTO) is used to compute luminosity.

Given the low Galactic latitude of the cluster, the measured magnitudes are highly attenuated by dust extinction, which varies across the field (e.g., from $\sim$0.6 to $\sim$0.9 in the *i*-band). In order to correct for this effect and recover intrinsic magnitudes, we used the reddening values from Schlafly & Finkbeiner (2011), with a spatial resolution of $\sim$4$'$. We interpolate between the extinction pierce points using cubic interpolation to predict

---

[9] http://www.astromatic.net/software/scamp

[10] http://www.astromatic.net/software/swarp
[11] http://www.cfht.hawaii.edu/Instruments/Elixir/

the dust attenuation at each source position. We correct the g, r and i magnitudes by interpolating in wavelength to the effective wavelength of the Subaru and CFHT filters (see Stroe et al. 2014c for details). The 5σ limiting magnitudes are 24.9, 24.1, and 25.4 for the g, r, and i filters, respectively.

3.5. Point-spread Function Modeling

A point-spread function (PSF) distorts galaxy shapes, and thus it is important to carefully model it and remove its effect when we measure weak lensing signals. Our PSF modeling is based on the principal component analysis (PCA) described in-depth in Jee et al. (2007a) and Jee & Tyson (2011). The method has been successfully applied to a wide range of data from ground (e.g., Jee et al. 2013; Dawson et al. 2012) to space (e.g., Jee et al. 2014a, 2014b). In the most recent competition called the 3rd GRavitational lEnsing Accuracy Testing (Mandelbaum et al. 2014), the method (team name: sFIT) won the data challenge. The application to ground-based data was described in detail in Jee et al. (2013). Below, we briefly summarize the key steps.

The SWARP software allows us to save the intermediate files (RESAMP) to be used for the final co-add. These files are already resampled to match the final co-add exactly (only integer pixel shifts are required). Our PSF model starts with identifying stars from these intermediate resampled images. In the CIZA J2242.8+5301 field, the typical number of stars per CCD usable for PSF modeling is ~400 or higher. This is more than enough to obtain reliable principal components (we kept the most significant 20 principal components) and describe the coefficients along these principal components using a polynomial interpolation method. Although recent studies suggest that a more sophisticated PSF interpolation method is required for future cosmic shear surveys,[12] we find that a simple third order polynomial interpolation method provides sufficient accuracy in the current cluster weak-lensing analysis. Because we choose to measure shapes from final co-adds, we model the PSF in the co-adds by propagating the PSF models of individual CCD frames.

The fidelity of the result obtained from this procedure can be examined by comparing the PSF statistics directly measured from stars in the co-add with those inferred from our PSF model propagated from individual exposure data. In Figure 2, we present such comparisons for both Subaru and CFHT images. A few features are noteworthy here. First, the PSF pattern in the coadd is complicated and shows a large spatial variation. Second, conspicuous discontinuities in the PSF variation is present (often these features are found across the CCD boundaries). Third, the PSF pattern predicted by our PCA-based model closely matches the observed PSF pattern. This agreement verifies that the PSF model obtained after stacking PSFs from individual exposures is indeed valid in representing the PSF in the co-add. In addition, it also warrants that there is no major systematic error in our image registration. For example, if there is any frame misaligned by as few as ~0.5 pixels in the x-axis, the (observed) stars in that region would possess large ellipticities along the x-axis in the coadd image, which in turn would cause large discrepancies between observation and model.

3.6. Shape Measurement

We use forward-modeling to determine galaxy ellipticity. The details are described in our past publications (e.g., Jee et al.

2013). In brief, we first convolve elliptical Gaussian with a PSF profile and then fit the result to galaxy images. We define the ellipticity with

$$e = \frac{a - b}{a + b}, \tag{1}$$

where a and b are the semi-major and -minor, axes, respectively, that are determined from the fit. Because real galaxy profiles are not Gaussian, the result is biased. Also, photon noise causes non-negligible bias in the shape parameters, in particular, for faint objects. In Jee et al. (2013), we determined the correction factor by carrying out image simulations with real galaxy images and found that our ellipticity should by multiplied by ~1.1 to recover input shears. This procedure calibrates out the isotropic part of the shear bias. Typically, shear measurement is also subject to anisotropic bias often correlated with PSF elongation. We corrected for this additive bias in our cosmic shear study presented in Jee et al. (2013). However, this additive bias is several orders of magnitude smaller than the statistical error and thus can be safely ignored in the current cluster lensing study.

As the ellipticity has both direction and magnitude, it is convenient to represent it using the following two components:

$$e_+ = e\cos(2\theta) \tag{2}$$

$$e_\times = e\sin(2\theta), \tag{3}$$

where θ is the angle between the x-axis and the major axis of the ellipse.

### 3.7. Multiplicative Bias by Stellar Contamination

Isolated stars with a reasonably high signal-to-noise ratio (S/N) are easily identified by their shape parameters and reliably removed from source catalogs. However, inevitably, some stars are also blended with other stars or galaxies. These blended objects when included into a source catalog as single objects dilute lensing signals because stars do not contain any lensing signal. H. Hoekstra et al. (2014, in preparation) estimated this dilution effect using the stellar population model of the Galaxy by Robin et al. (2003) and the lensing image simulation tool GalSim.[13] They report that at $|b| \gtrsim 40°$ the dilution effect by star contamination is much less than ~1%. However, as the Galactic latitude decreases below $|b| < 40°$, the dilution effect is rapidly increasing, reaching ~2% at $b = 20°$. Because the simulation of H. Hoekstra et al. (2014, in preparation) did not cover the range that includes our case ($|b| \sim 5°$), we have to extrapolate their result to estimate the dilution effect at $|b| = 5°$, which gives ~5% compared to the case at $b = 90°$. We apply this correction in addition to our shear calibration mentioned in Section 3.6. However, this dilution effect is still small relative to the shot noise, which already causes as large as ~20% error in cluster mass.

3.8. Source Selection

We base our source selection mainly on the color–magnitude relation with some additional steps to minimize the contamination from stars misidentified as galaxies. Figure 3 shows the relation between $g - i$ and i. The color–magnitude relation of the spectroscopically confirmed red members is tight. The objects in the yellow polygon were explicitly rejected because a significant fraction of the objects there is either stars or cluster members. We choose the "red" edge to be 0.5 mag redder than the mean red-sequence color. We let the "blue" edge have
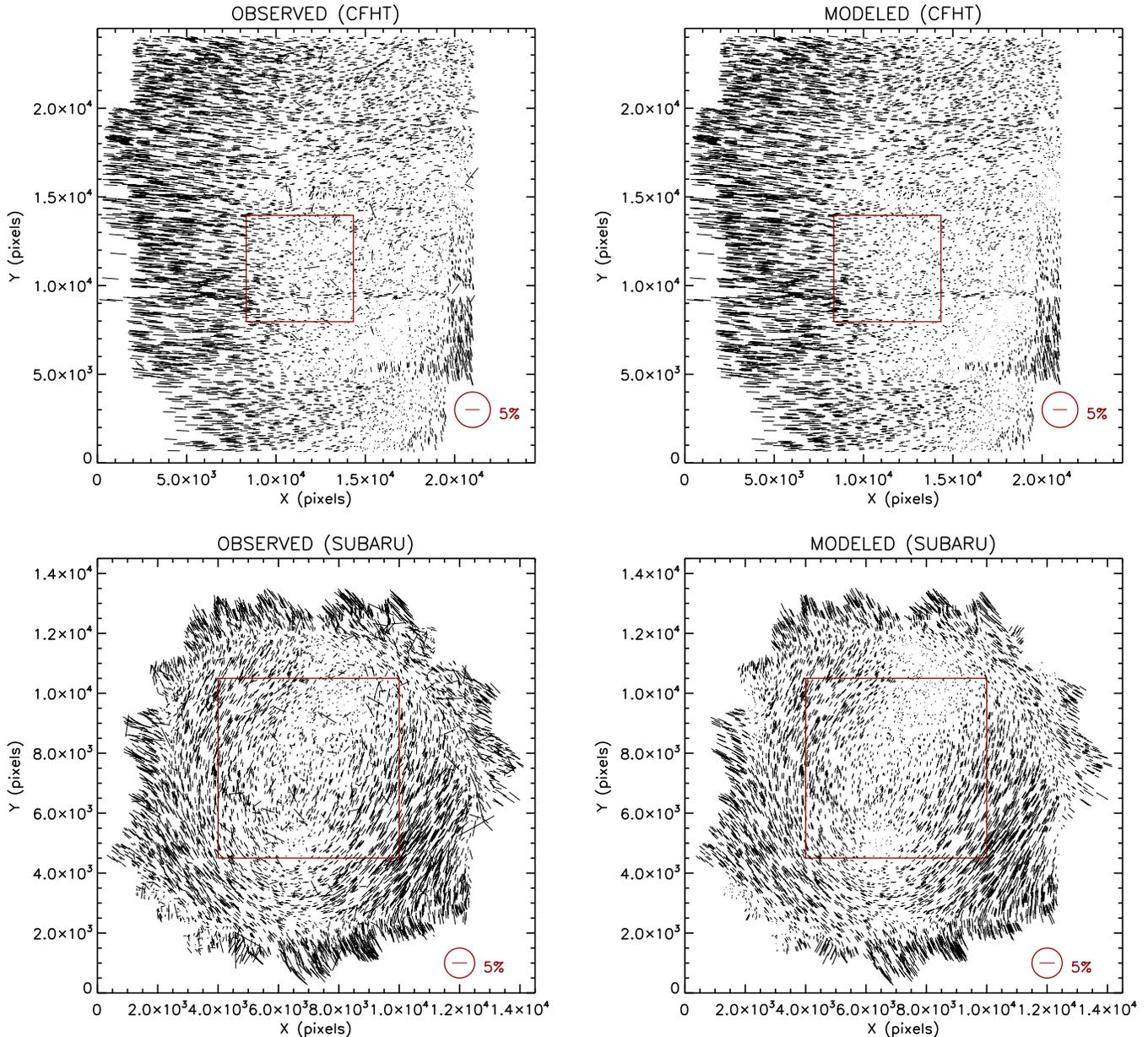

Figure 2. Comparison of ellipticities between observed stars and model PSFs after image coadding. We define ellipticity as $(a - b)/(a + b)$, where a and b are the semi-major and minor axes, respectively. In the top panel, we display the comparison in CFHT/MegaCam data (r band). We do not remove spurious ellipticities in the left panel, which are contributed by severely blended stars or galaxies mis-identified as stars. Except for these outliers, the agreement is excellent. This agreement also serves as a verification that no large systematic error is present in image registration, in which case we would detect systematic elongation of stars only in the left panel (the model in the right panel assumes that image alignment is perfect). The red box indicates the central $20' \times 20'$ region, where we perform our two-dimensional mass reconstruction. The encircled red stick on the lower right corner shows the size of 5% ellipticity. In the bottom panel, the result is for the Subaru/SuprimeCam data in the i band.

a larger margin(2 mag) to minimize stellar contamination. This color cut is applied to sources with $i \sim 22$ because beyond this limit the photometric scatter will make the color–magnitude relation-based source selection much less reliable. The blue dots represent the sources that we select for weak-lensing analysis. We do not use the r-band photometry in this selection because photometric redshift estimation with the three g, r, and i bands does not provide significant advantage in isolating the cluster members. In addition, it is difficult to obtain reliable r-band photometry for many sources because of the charge bleeding.

In addition to the above color–magnitude relation, we require the object half light radius (in the i band) to be greater than $0''\!.44$, which is the upper limit of the stellar locus in the current data. Also, the pre-seeing semi-minor axis (determined after PSF-convolved elliptical Gaussian fitting) should be significantly greater than zero. This last measure is needed to remove remaining stars and avoid the aliasing effect pertinent to small objects. We also applied S/N cuts by ensuring that the ellipticity error per component is less than 0.3 and the detection significance is above 5σ. From the Subaru i-band image, we obtain $\sim$16 galaxies per sq. arcmin, which is roughly a factor of two lower than the typical number that we achieve with Subaru in other fields (e.g., Dawson et al. 2012). We repeat that this low source density is due to both high stellar density and high
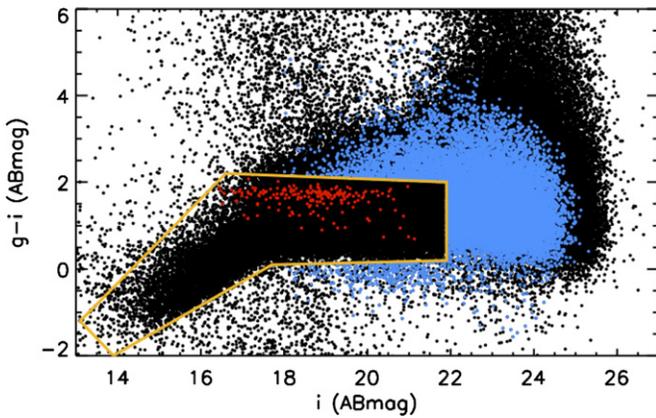
Figure 3. Source selection based on the color–magnitude relation in the CIZA J2242.8+5301 field. Extinctions are corrected for. The yellow polygon encloses an area that is mostly populated with stars and the cluster members. The red dots represent the 218 spectroscopically confirmed members of CIZA J2242.8+5301. The sources selected for weak-lensing analysis are shown in blue.

extinction. Because of both relatively shallow depth (the 5σ limiting mag is $r_{\mathrm{limit}} \sim 24.1$) and obscuration by bleeding trails, the source number density is a factor of two lower ($\sim$8 galaxies per sq. arcmin) in the CFHT r-band data. When objects appear both in Subaru and CFHT, we combine their shapes by weight-averaging their ellipticities. Weights are computed using ellipticity errors, which are in turn derived from a Hessian matrix. Although we merge the two shape catalogs from Subaru and CFHT to perform our final weak-lensing analysis, the result is consistent with the case when only Subaru shapes are used. This is not surprising because most of the objects detected in the CFHT image are also visible in the Subaru image. However, it is important to note that the results based on the CFHT shapes are also statistically consistent with the results obtained with the Subaru shapes. In particular, we emphasize that the two versions of mass maps from both Subaru and CFHT are in excellent agreement, which will be demonstrated in Section 5.

3.9. Cluster Member Selection

Detailed comparison of lensing results with cluster galaxies provides useful probes to the interplay between different constituents of galaxy clusters. To facilitate comparisons between the mass and member galaxy distributions in CIZA J2242.8+5301, we select a separate sample of likely cluster members as follows.

We use the color–magnitude relation of the red cluster members in conjunction with our spectroscopic survey data of the cluster field. The locus of the red-sequence is determined using the 217 spectroscopic members. Although this spectroscopic sample contains a significant number of blue members, the tight relation (Figure 3) allows us to define a narrow strip ($\delta(g-i) \sim 0.6$) in the magnitude range $16.5 < i < 20$ to isolate the red-sequence galaxies. Stars are removed from this initial selection by discarding objects if they are spectroscopically confirmed stars, their half light radius is greater than $r_h = 0\farcs44$, or their semi-minor axis after deconvolution is statistically consistent with zero. The cluster member catalog is further refined by visually inspecting the object to reject any conspicuously blended stars. Finally, we add the blue spectroscopic members that are located outside the initial narrow strip. We compare the smoothed luminosity and number density maps with our mass reconstruction in Section 5.

4. BASIC WEAK-LENSING THEORY AND SOURCE REDSHIFT ESTIMATION

Interested readers are referred to Bartelmann & Schneider (2001) and Hoekstra (2013) for more details on weak lensing and its application to galaxy clusters. Here we only provide a brief review.

In a weak-lensing regime, we can assume that averaging the ellipticity components defined in Equations (2) and (3) become reduced shears g_1 and g_2, respectively. That is,

$$\langle e_+ \rangle \simeq g_1 \qquad (4)$$

and

$$\langle e_\times \rangle \simeq g_2. \qquad (5)$$

The reduced shears comprise the matrix elements of the following weak-lensing transformation:

$$\mathbf{A} = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}, \qquad (6)$$

where κ is the surface mass density in units of the critical density:

$$\Sigma_c = \frac{c^2}{4\pi G D_l \beta}. \qquad (7)$$

In Equation (7), c is the speed of light, G is the gravitational constant, and D_l is the angular diameter distance to the lens. β is defined below.

The lensing signal strength depends on the distance ratios between lens, source, and the observer. In order to interpret and quantify the observed signal properly, it is necessary to estimate the source redshift. Often, a lensing efficiency is expressed in terms of β defined as

$$\beta = \max[D_{\mathrm{ls}}/D_s, 0], \qquad (8)$$

where D_{ls} and D_s are the angular diameter distances between the lens and the source, and between the observer and the source, respectively. Because sources in front of the lens do not contribute to the lensing signal, we assign zero to β in those cases (D_{ls}/D_s becomes negative). Therefore, we need to obtain $\langle \beta \rangle$ to characterize the effective lensing efficiency of the source population. We refer to the corresponding redshift as effective redshift z_{eff}. However, because the lensing efficiency is nonlinear, using $\langle \beta \rangle$ alone to represent a broad distribution biases the result. A first order correction is given when we update the reduced shear g with the following equation (Seitz & Schneider 1997):

$$g' = [1 + (\langle \beta^2 \rangle / \langle \beta \rangle^2 - 1)\kappa]g. \qquad (9)$$

Since we base our selection on the two-filter broadband photometry as described in Section 3.8, we estimate the source redshift indirectly utilizing a publicly available reference catalog. We choose the Cosmic Evolution Survey (COSMOS) photometric redshift catalog[14] (Ilbert et al. 2009) as our reference because (1) the magnitude-limit $I < 25$ is well-suited to the current analysis, (2) the result is obtained based on 30 filters, and (3) the survey area is reasonably large (2 sq. degs) so that the result does not greatly suffer from the cosmic variance. In addition, the Ilbert et al. (2009) photometry is also derived from the

---

[14] Available at http://irsa.ipac.caltech.edu.
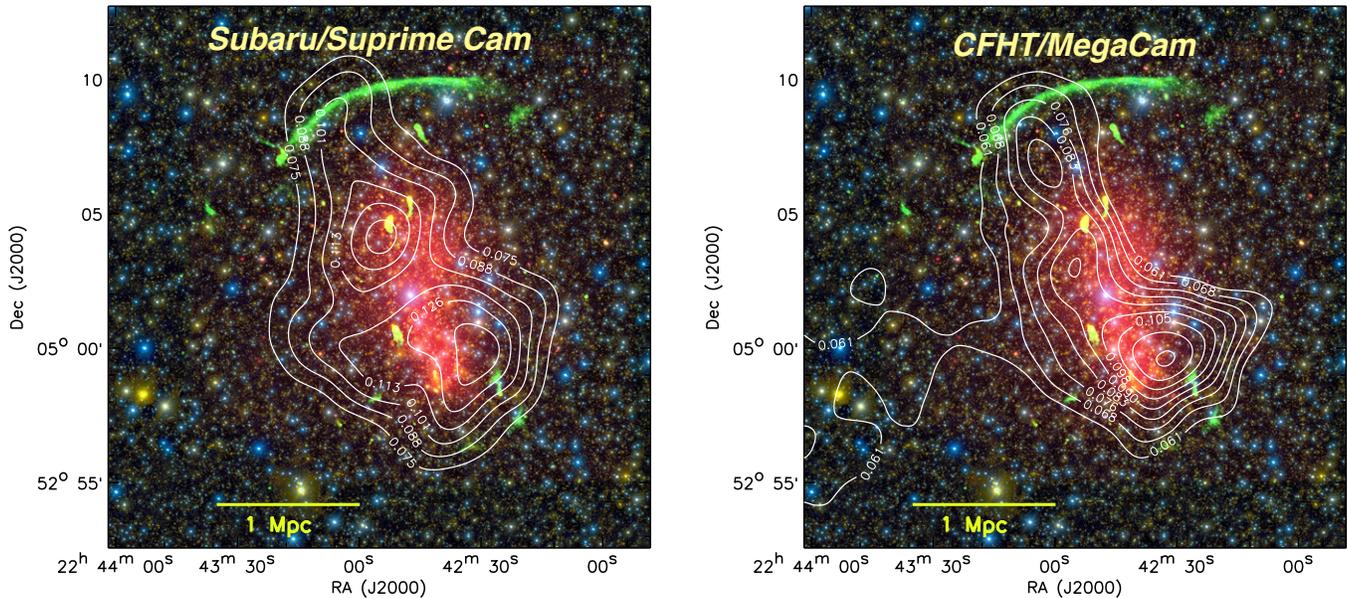
Figure 4. Maximum-entropy mass reconstructions from Subaru (left) and CFHT (right). The white contours represent the projected mass density (κ), which is subject to the mass-sheet degeneracy ($\kappa \rightarrow \lambda\kappa + 1 - \lambda$). We arbitrarily scale the mass map in such a way that the average κ value becomes approximately zero near the field boundary. The Subaru and CFHT results are consistent with other each. The consistency between independent telescopes serves as verification that the observed substructures in both maps are not due to residual systematics. Readers are reminded that the PSF patterns are very different between the two observations (Figure 2). It is also important to note that the source number density from the CFHT data is a factor of two lower.

Subaru Telescope, which obviates the need for any transformation between the current CIZA J2242.8+5301 and the COSMOS photometric systems.

We apply the source selection criteria described in Section 3.8 to the COSMOS galaxy catalog. We obtain $\langle \beta \rangle = 0.656$ and the corresponding effective redshift of this subset is $z_{\mathrm{eff}} = 0.626$. Because the depth of our CIZA J2242.8+5301 image is shallower than the COSMOS images, it is necessary to correct for the difference. We compute this correction factor by constructing magnitude histograms for the two catalogs and weight galaxies according to the ratio of our source number density per magnitude bin to the COSMOS one. The stellar contamination discussed in Section 3.7 is also taken into account in this step. Both $\langle \beta \rangle$ and z_{eff} decrease to 0.616 and 0.549, respectively. We obtain $\langle \beta^2 \rangle = 0.438$, which is needed to account for the width of the redshift distribution (Equation (9)).

5. TWO-DIMENSIONAL MASS RECONSTRUCTION

The reduced shear defined in Equations (4) and (5) is directly observable (up to shear calibration) by averaging object ellipticities. However, in order to obtain the surface mass density κ, we need to know shears γ, which are related to reduced shears via

$$g = \frac{\gamma}{1 - \kappa}. \qquad (10)$$

Often, $g = \gamma$ is assumed in the very weak-lensing regime where $\kappa \ll 1$. Under this assumption, we can obtain the two-dimensional convergence field by performing the following integral:

$$\kappa(\mathbf{x}) = \frac{1}{\pi} \int D^*(\mathbf{x} - \mathbf{x}')\gamma(\mathbf{x}')d^2\mathbf{x}, \qquad (11)$$

where D^* is the convolution kernel defined as $D(\mathbf{x}) = -1/(x_1 - ix_2)^2$ and $\mathbf{x}$ is the coordinate. This direct inversion first used in Kaiser & Squires (1993) is still a popular method to reconstruct a two-dimensional mass distribution.

However, near cluster centers, κ is non-negligible, and thus we need to include the nonlinear relation between g and γ. In the current paper, we used the maximum-entropy inversion code of Jee et al. (2007b), which utilizes the entropy of the mass pixels as prior to regularize the result.

We show our mass reconstruction results in Figure 4 from both Subaru and CFHT. Clearly, the results demonstrate that the mass distribution is elongated along the merger axis inferred from the radio relics. Note that the substructures seen in both instruments are consistent. We regard this consistency as verification that the observed substructures are not due to any residual systematics in weak-lensing measurements. As shown in Figure 2, the PSF ellipticity patterns of Subaru and CFHT are different. Thus, if our PSF correction error were significant, we would not observe this level of consistency between the two instruments. Also, remember that the source number density in the CFHT weak-lensing catalog is a factor of two lower. This implies that the CFHT result can be regarded as one of the bootstrapping realizations of the Subaru result (in the absence of systematic errors). Most of the sources used for the CFHT weak-lensing analysis are present in the Subaru source catalog. Therefore, little difference from the Subaru result is observed when we combine the two source catalogs. Nevertheless, we note that our weak-lensing analysis hereafter is based on the union of both Subaru and CFHT shape catalogs.

Figure 5 compares the mass reconstruction from the combined catalog (CFHT+Subaru) with the cluster galaxies. The cluster member selection is described in Section 3.9. We smooth the galaxy distribution using a Gaussian kernel with a FWHM $\simeq 3\farcm4$. Both number and luminosity maps indicate that the cluster galaxies in CIZA J2242.8+5301 have a bimodal distribution, and our mass reconstruction reveals two dominant mass clumps that can be associated with the two peaks in both luminosity and number density maps. It is apparent that the mass centroids do not perfectly align with the cluster galaxies. The mean offset between the mass and luminosity/number density peaks is $\sim 1'$.
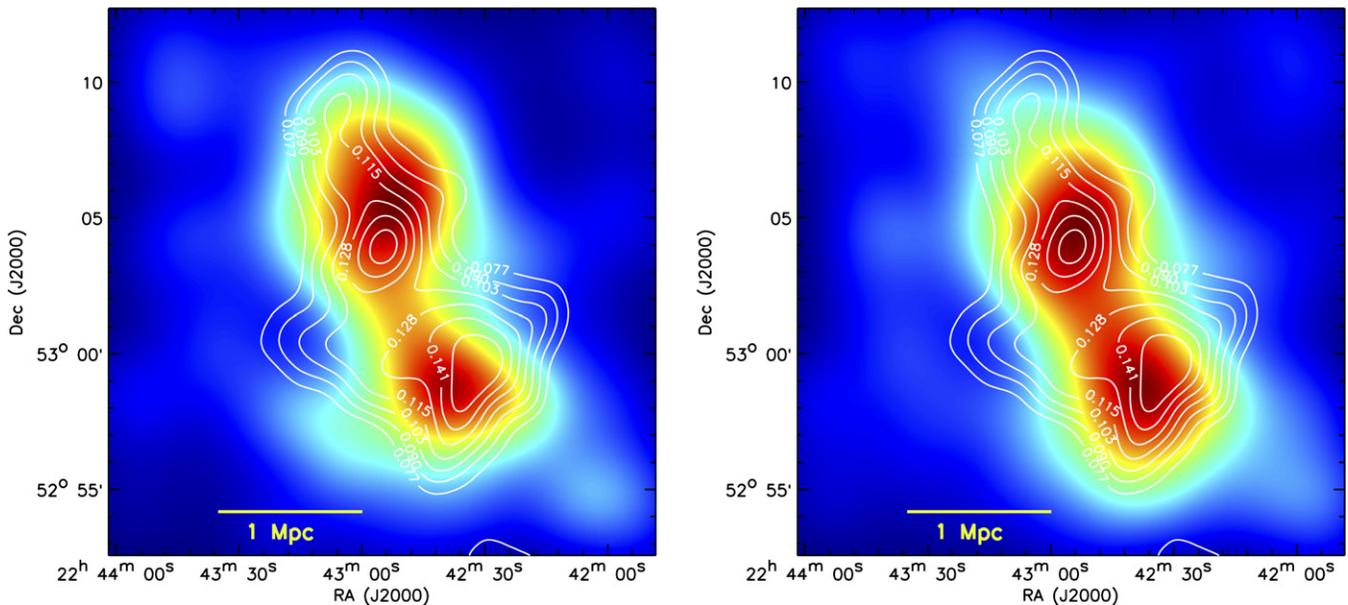
Figure 5. Mass distribution overlaid on galaxy distribution. The white contours represent the mass reconstruction when we combine shapes from CFHT and Subaru. The background is color-coded with smoothed galaxy luminosity (left) and number density (right). Smoothing is done with a Gaussian with a FWHM $\sim 3'.4$.

We present our analysis on the statistical significance of the mass–galaxy offset in Section 7.3.

6. HOW MASSIVE IS CIZA J2242.8+5301?

Together with the radio relics, galaxy distribution, and X-ray morphology, our weak-lensing mass reconstruction shows that CIZA J2242.8+5301 is comprised of at least two massive halos separated by ~ 1 Mpc. This is analogous to the situation where we need to perform accurate photometry for two blended objects that are of comparable luminosity. Ignoring the presence of the companion will lead to significant overestimation of each object's luminosity. Of course, this companion bias is avoided when we fit the two objects simultaneously and estimate the luminosity based on the resulting best-fit parameters. Similarly, if we assume a single halo and attempt to estimate the mass using the azimuthally averaged shear profile, the resulting mass will become non-negligibly biased. Hence, for the mass estimation of CIZA J2242.8+5301 , we fit the two halos simultaneously. One caveat in this approach is that we still have to assume that both halos can be represented by a Navarro–Frenk–White (NFW; Navarro et al. 1997) profile. In general, the difference between our model and actual cluster profiles is an important topic for cluster cosmology, where one desires to know the mass with bias smaller than 1% accuracy. In this study, we do not suggest a solution to this cluster model bias. Instead, we perform a separate mass estimation using aperture mass densitometry (Fahlman et al. 1994; Clowe et al. 2000), which provide non-parametric masses within an aperture, and compare these projected masses with the results derived from our parameterization. We demonstrate that the discrepancy is negligible within the current statistical error.

Although we emphasize above that the traditional method of fitting an analytic profile to one-dimensional (1D) tangential shear profile causes bias, we still present the results from this approach in Figure 6 because the results give rough estimates and provide opportunity to examine a possible presence of systematics (so-called B-mode test). The B-mode signal is consistent with zero, which shows that no significant systematic

Table 1
Mass Estimates of CIZA J2242.8+5301 Based on 1D Fit

Subclusters	σ_v (km s^{-1})	c_{200c}	R_{200c} (h_{70}^{-1} Mpc)	M_{200c} ($h_{70}^{-1} 10^{14} M_\odot$)
North	1046 ± 65	3.19 ± 0.06	2.0 ± 0.1	11.1 ± 2.2
South	1182 ± 58	3.10 ± 0.04	2.3 ± 0.1	15.7 ± 2.5

Note. The σ_v value is derived from the SIS fit whereas the rest are from the NFW fit.

errors are present in our analysis. The singular isothermal sphere (SIS) model is used to predict velocity dispersions whereas the NFW model is assumed for mass estimation. The vertical dotted line in Figure 6 is to denote the minimal radius inside which we exclude the data for fitting. This need for excluding the tangential shear at small radii is shared by the weak-lensing community because of the following reasons. First, crowding makes reliable shape measurements difficult. Second, the cluster galaxy contamination rate into the source catalog is higher. Third, mis-centering distorts the shear profile. Fourth, weak-lensing assumption becomes invalid when the cluster possesses strong-lensing features. We summarize the fitting results in Table 1, which are to be compared to the results from our simultaneous fitting explained below.

The complicated shape of the observed tangential shear profiles indicate that CIZA J2242.8+5301 cannot be described by a single halo without significant bias in cluster mass estimation. At relatively small radii ($r \lesssim 10'$), it seems that the substructures and halo mis-centering suppress the signal amplitude whereas at $r \gtrsim 10'$ the tangential shear profile rises because of the contribution from both halos. In Figure 6 the two-component models (blue) described below possess these features and better-fit the observed reduced shear profile.
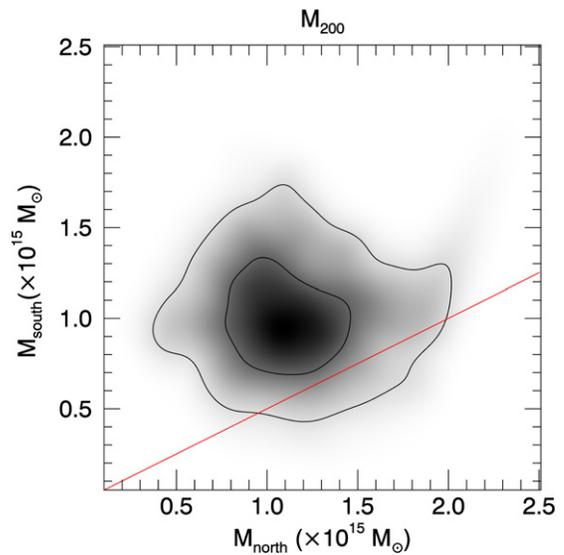
Because we desire to determine parameter uncertainties and examine their correlations robustly, we perform Markov Chain Monte Carlo (MCMC) analysis. In fact, this approach was also used in the case of ACT-CL J0102−4915 (Jee et al. 2014b), which closely resembles CIZA J2242.8+5301 in terms
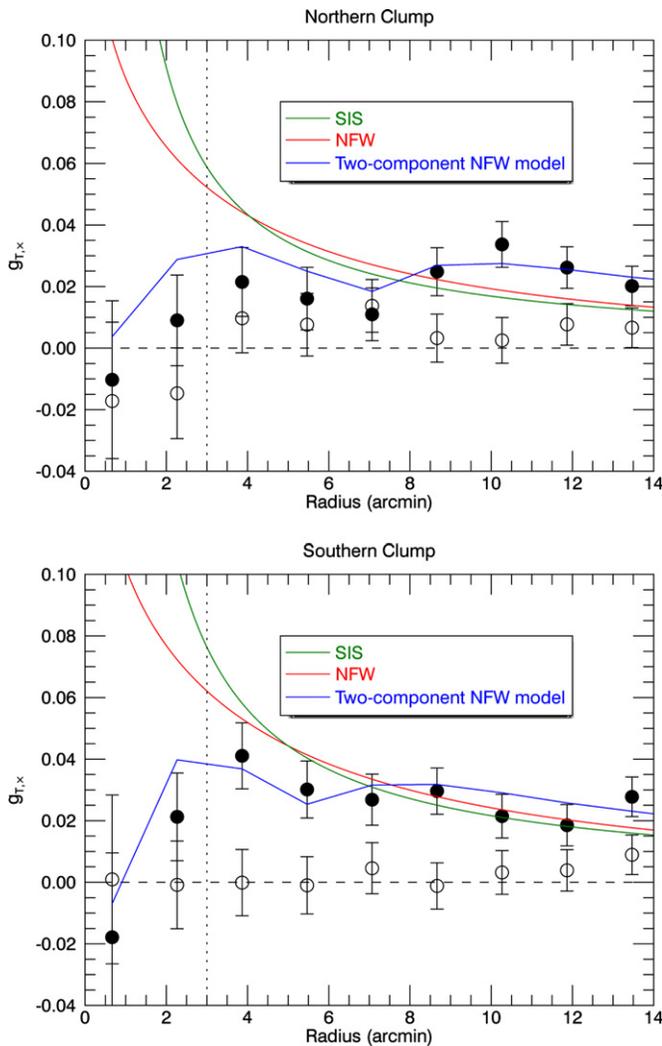
Figure 6. Tangential shear profile of CIZA J2242.8+5301. The top and bottom panels show the tangential shear profile for the northern and southern components, respectively. We choose the luminosity centers (Figure 5) as references for both. The filled circles represent the weak-lensing signal whereas the open circles show so-called *B*-mode signal, which should be consistent with zero as observed if the lensing signal comes from pure E-mode. We refer to Table 1 for the profile fitting results. The vertical dotted line is to denote the minimal radius inside which we exclude the data for fitting.

Figure 7. Mass constraints on CIZA J2242.8+5301 from simultaneous 2D fitting. The inner and outer contours correspond to 1σ and 2σ confidence. The solid red line depicts the 2:1 mass ratio assumed in the simulation of van Weeren et al. (2011).

Table 2
Mass Estimates of CIZA J2242.8+5301 Based on Simultaneous 2D Fit

| Subclusters	σ_v (km s^{-1})	c_{200c}	R_{200c} (h_{70}^{-1} Mpc)	M_{200c} ($h_{70}^{-1} 10^{14} M_\odot$)
North	967^{+113}_{-128}	$3.20^{+0.09}_{-0.08}$	$2.0^{+0.2}_{-0.2}$	$11.0^{+3.7}_{-3.2}$
South	1137^{+93}_{-101}	$3.23^{+0.08}_{-0.08}$	$1.9^{+0.1}_{-0.2}$	$9.8^{+3.8}_{-2.5}$

Note. The σ_v value is derived from the SIS fit whereas the rest are from the NFW fit.

of their structure and mass. ACT-CL J0102−4915, nicknamed "El Gordo" (Menanteau et al. 2012) meaning "the Fat One" in Spanish, is a merging system with two comparably massive ($\sim 10^{15} M_\odot$ each) clumps. Similarly to the case of ACT-CL J0102−4915, we place two NFW halos over the northern and southern sub-clusters. Throughout the paper, the Duffy et al. (2008) mass–concentration relation is assumed. We choose the galaxy luminosity center as an initial centroid and allow it to vary as we run our MCMC. We use a flat prior with a bounding box $200'' \times 200''$ centered on the luminosity peak. For the concentration parameter, we use a flat prior ranging from $c = 2$ to 4, which corresponds to the range $M_{200} = 7 \times 10^{13}$–$3 \times 10^{17} M_\odot$ according to the Duffy et al. (2008) relation. The 1D fitting results above are used as initial values. We exclude source galaxies when they are located within $r_{cut} = 200''$ circle. The mass varies $\sim 11\%$ when r_{cut} varies between $150'' < r_{cut} < 250''$.

The resulting parameter constraints are displayed in Figure 7. The most probable masses are $M_{200c} \sim 10^{15} M_\odot$ for both

clusters. We summarize the fitting results in Table 2. The mass of the southern halo based on the 1D analysis is higher by $\sim 60\%$, although the results are in good agreement for the northern clump. The large difference in the southern subcluster indicates that cluster mass estimation with weak-lensing requires care when non-negligible substructures are present.

The aperture mass densitometry can be performed by evaluating the following integral.

$$\zeta_c(r_1, r_2, r_{\max}) = \bar{\kappa}(r \leqslant r_1) - \bar{\kappa}(r_2 < r \leqslant r_{\max})$$

$$= 2 \int_{r_1}^{r_2} \frac{\langle \gamma_T \rangle}{r} dr + \frac{2}{1 - r_2^2/r_{\max}^2} \int_{r_2}^{r_{\max}} \frac{\langle \gamma_T \rangle}{r} dr, \quad (12)$$

where $\langle \gamma_T \rangle$ is the azimuthally averaged tangential shear, r_1 is the aperture radius, and r_2 and $r_{\max}$ are the inner- and the outer-radii of the annulus. $\zeta_c(r_1, r_2, r_{\max})$ provides a density contrast of the region inside $r < r_1$ with respect to the control annulus ($r_2, r_{\max}$). We choose $r_2 = 600''$ (~ 1.9 Mpc) and $r_{\max} = 800''$ (~ 2.5 Mpc) for the control annulus and estimate the density within this region by projecting our NFW fitting results (Table 2). The input to the equation of the densitometry is a shear, not a reduced shear. Therefore, we determine the aperture mass iteratively by updating κ. The result is displayed in Figure 8. Also, displayed in Figure 8 is the aperture mass estimated with the NFW fitting results above. In order to obtain this estimation, we first projected each NFW profile along the line-of-sight direction and sum the two cluster results. The projected masses from aperture mass densitometry and NFW fitting are in good agreement at large radii, which indicates that
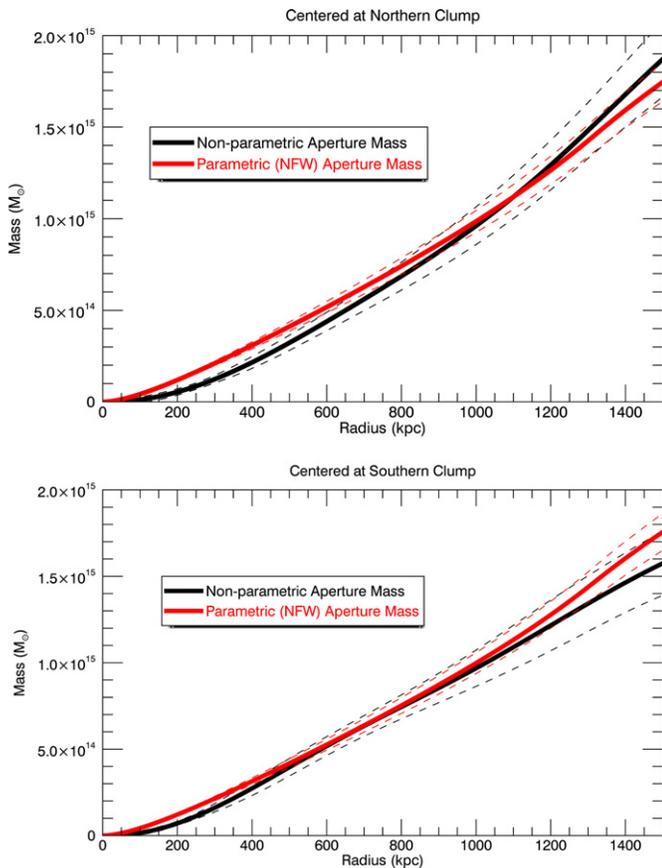
Figure 8. Comparison of projected masses between aperture mass densitometry and NFW fitting. The aperture mass densitometry is performed using the luminosity center as a origin. We set up a control annulus at $r = 1.9$–2.5 Mpc. The density within this region is estimated utilizing the two-dimensional NFW fitting results. For computation of the aperture mass based on NFW profile fitting, we stack two projected NFW profiles created with the parameters shown in Table 1. The results from the two different methods are in excellent agreement, indicating that despite the apparent violent merger, parameterizing the halos with NFW profile is still a good approximation in CIZA J2242.8+5301. The dashed lines depict 1σ (lower and upper) limits.

despite the apparent violent merger, parameterizing the halos with NFW profiles still provides a reasonable global mass in CIZA J2242.8+5301. At smaller radii ($r \lesssim 1$ Mpc), we observe some disagreements, which may arise not only from a genuine difference at the core, but also from centroid mis-alignments and/or data binning artifacts.

Given that CIZA J2242.8+5301 is comprised of two halos of approximately equal mass, we can estimate the total mass of the system in the following way. We adopt the mean position of the two halos as the cluster center. Namely, the coordinate of this location is $(\alpha, \delta) \simeq (22^{\mathrm{h}}42^{\mathrm{m}}47\overset{s}{.}3, 53°01'56'')$. We set up a three-dimensional box (3D) with a side of 6 Mpc containing the two halos of CIZA J2242.8+5301. A density is assigned to each volume element based on the parameters in Table 2. Then, it is straightforward to compute the total spherical mass as a function of radius. We estimate that $M_{200c} = (2.51 \pm 0.53) \times 10^{15} M_\odot$ is reached at $r_{200c} = 2.63$ Mpc. This mass is comparable to the value $M_{200c} = (2.76 \pm 0.51) \times 10^{15} M_\odot$ of "El Gordo" similarly comprised of two $\sim 10^{15} M_\odot$ halos. We remind readers that the total mass becomes greater than a mere sum of the two halos' masses because r_{200c} should increase. Although this total mass estimation assumes that the two subclusters are at the same distance from us, the result does not vary significantly with

viewing angles (see Jee et al. 2014b for the case of El Gordo) as long as the angle between the merger axis and the plane of the sky is less than $\sim70°$. The polarization of the radio relic indicates that the angle is less than $\sim30°$ (van Weeren et al. 2010), which is consistent with the result from our detailed dynamical analysis (W. Dawson et al., in preparation).

7. DISCUSSION

7.1. Comparison with Dynamical and X-Ray Studies

To date, we have 217 spectroscopically confirmed members of CIZA J2242.8+5301. A comprehensive dynamical analysis will be presented in a separate publication (W. Dawson et al., in preparation). Here we will present basic comparisons relevant to mass properties.

It is impossible to distinguish the membership of galaxies individually between the two halos because the difference in the line-of-sight velocity between the two halos whose virial radii enclose each other's center is an order of magnitude smaller than the velocity dispersion of each halo. Thus, we adopt a common operational definition based on the (projected) proximity to each clump. Within the $r = 200''$ circle (~625 kpc), we find that the number of galaxies belonging to the northern and southern components are 69 and 62, respectively. The redshift of each component is determined with bi-weight estimator, giving $\langle z \rangle = 0.18794 \pm 0.00053$ and 0.18821 ± 0.00052 for the northern and southern clumps. The line-of-sight velocity difference between the two components is $\Delta z = 0.00027$ (~69 km s^{-1}), which indicates that the merger is happening nearly in the plane of the sky.

The directly measured velocity dispersions of the northern and southern clumps are $\sigma_v = 1163^{+103}_{-86}$ km s^{-1} and 1084^{+102}_{-74} km s^{-1}, respectively. These estimates can be compared with our lensing results by assuming that the cluster is composed of two SISs. Our simultaneous two-dimensional fitting (Table 2) gives $\sigma_v = 967^{+113}_{-128}$ km s^{-1} and 1137^{+93}_{-101} km s^{-1} for the northern and southern components, respectively. The predicted velocity dispersion from our lensing analysis for both subclusters are consistent with the direct measurement.

One may choose to convert the directly measured velocity dispersion to a mass and compare the result to weak-lensing mass. Using the $M_{200c} - \sigma_{DM}$ relation of Evrard et al. (2008), we convert $\sigma_v = 1163^{+103}_{-86}$ km s^{-1} and 1084^{+102}_{-74} km s^{-1} to $M_{200c} = 1.61^{+0.46}_{-0.33} \times 10^{15} M_\odot$ and $1.30^{+0.40}_{-0.25} \times 10^{15} M_\odot$, respectively, where the errors on the mass include only the statistical noise (not the scatter of the relation). These values are consistent with our lensing results (Table 2).

The first gas temperature measurement of CIZA J2242.8+5301 is presented in Ogrean et al. (2013). The result is based on deep *XMM-Newton* data. The global temperature 7.7 ± 0.3 keV indicates that the cluster is massive. Assuming that the cluster follows an isothermal beta model with $\beta_X = 0.7$ and $r_c = 100$ kpc, we obtain $M_{200c} \sim 1.3 \times 10^{15} M_\odot$. This value is a factor of two smaller than our weak-lensing result, which is not surprising because the entire system CIZA J2242.8+5301 cannot be modeled with a single isothermal beta halo.

7.2. Implication for Merging Scenario

The giant radio relic of CIZA J2242.8+5301 provides direct evidence for the on-going merger, which is also supported by the gas and galaxy distribution. The numerical simulation of van Weeren et al. (2011) successfully reproduces the observed
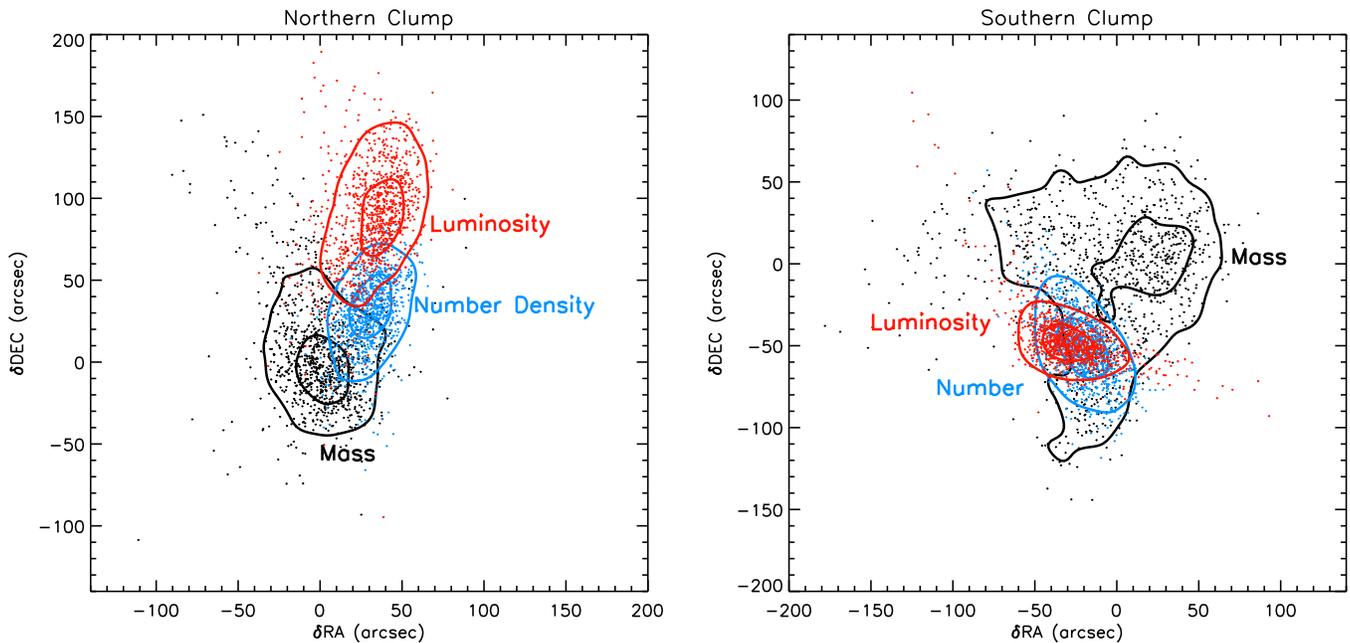
Figure 9. Centroid uncertainty estimation using bootstrap resampling. The coordinate (0, 0) represents the approximate median of the mass centroid distribution. Each dot represents a measurement from a single realization (out of the 1000 bootstrapping runs). The inner and outer contours correspond to 1σ and 2σ levels, respectively. The centroid distribution for each mass clump is similar to the shape in mass distribution. Accordingly, the northern mass clump, which is more peaked than the southern clump, has a smaller centroid uncertainty. On the other hand, the size of the galaxy centroid distribution is similar between the northern and southern clusters. Our bootstrapping analysis shows that both northern and southern mass peaks are offset from the corresponding galaxy peaks at least at the $\sim 2\sigma$ level; the 1σ contours of the mass and galaxy centroids do not overlap.

features of the radio relics in CIZA J2242.8+5301. Here we discuss the implication of the current weak-lensing results on the initial condition by noting the following two apparent differences between our findings and the assumptions made for the simulation of van Weeren et al. (2011).

First, the most striking difference is the total mass of the system. Our weak-lensing analysis finds that the total mass of CIZA J2242.8+5301 should exceed $M_{200c} \sim 2 \times 10^{15}\,M_\odot$ (Section 6). On the other hand, in van Weeren et al. (2011) the total mass was assumed to be $5.5 \times 10^{14}\,M_\odot$, which is a factor of four lower.

Second, the mass ratio is different. van Weeren et al. (2011) found that the mass ratio of the northern cluster to the southern cluster should be close to 2:1 in order to reproduce the observed asymmetry of the two relics. However, this 2:1 mass ratio is not favored by our weak-lensing analysis, which shows that the best-fit masses of the two sub-clusters are in fact of nearly equal mass.

The first difference (i.e., total mass) comes from the fact that van Weeren et al. (2011) assume that the total mass of CIZA J2242.8+5301 is $5.5 \times 10^{14}\,M_\odot$, which is based on the luminosity-mass ($L_X - M$) relation of Pratt et al. (2009). One of the most important merger parameters that will be affected by updating the mass would be the impact velocity between the two halos. Using the simplified assumption made in Sarazin (2002), we can estimate the initial separation d_0 using Kepler's 3rd law

$$d_0 \sim 4.5 \left(\frac{M_1 + M_2}{10^{15}\,M_\odot} \right)^{1/3} \left(\frac{t_{\mathrm{impact}}}{10^{10}\,\mathrm{yr}} \right)^{2/3}\,\mathrm{Mpc}, \quad (13)$$

where t_{impact} is the time elapsed at the impact since the unknown epoch t_0 when the two halos detached from the Hubble flow and started to accelerate toward each other. It is important to remember that the impact velocity that we discuss below is not sensitive to the exact value of d_0 (t_0).

It is impractical to set up a numerical simulation to begin at $t = t_0$ and run it until $t = t_{\mathrm{impact}}$. Therefore, most N-body simulations choose a reasonable initial distance d as a starting point. The relative velocity between the two halos when they are separated by d is given by (Sarazin 2002)

$$v \sim 2930 \left(\frac{M_1 + M_2}{10^{15}\,M_\odot} \right)^{1/2} \left(\frac{1 - d/d_0}{1 - (b/d_0)^2} \right)^{1/2}$$
$$\times \left(\frac{d}{1\,\mathrm{Mpc}} \right)^{-1/2}\,\mathrm{km\,s^{-1}}, \quad (14)$$

where b is the impact parameter. When we assume $t_{\mathrm{impact}} = t_{\mathrm{age}}$ (~ 11 Gyr), $b = 0$, and $M_1 + M_2 = 5.5 \times 10^{14}$, the relative velocity of the two halos is $\sim 1000\,\mathrm{km\,s^{-1}}$ at $d = 2\,\mathrm{Mpc}$. In fact, this relative velocity is used as an initial condition in the simulation of van Weeren et al. (2011). Instead, if we substitute our weak-lensing masses into $M_1 + M_2$, the relative velocity increases to $\sim 2500\,\mathrm{km\,s^{-1}}$ at the same separation $d = 2\,\mathrm{Mpc}$. This difference is non-negligible and will greatly amplify at the core pass-through. The estimation of the exact relative speed at the core impact requires N-body simulations, and we will defer the estimation to our future studies.

The second difference may appear puzzling because the simulation clearly shows that a mass ratio of 2:1 (the northern clump should be more massive than the southern clump by a factor of two) is required to reproduce the observed asymmetry in the two radio relics. Nevertheless, given the size of the mass uncertainties in the current study, our result does not rule out the case of the 2:1 mass ratio. As shown in Figure 7, the tension is only slightly greater than 1σ. In addition, even in the case that the global mass ratio is 1:1, the asymmetry in the relics of CIZA J2242.8+5301 can be accommodated in the scenario, wherein the two subcluster's concentration values are different. Since a halo with a higher concentration packs more mass within the
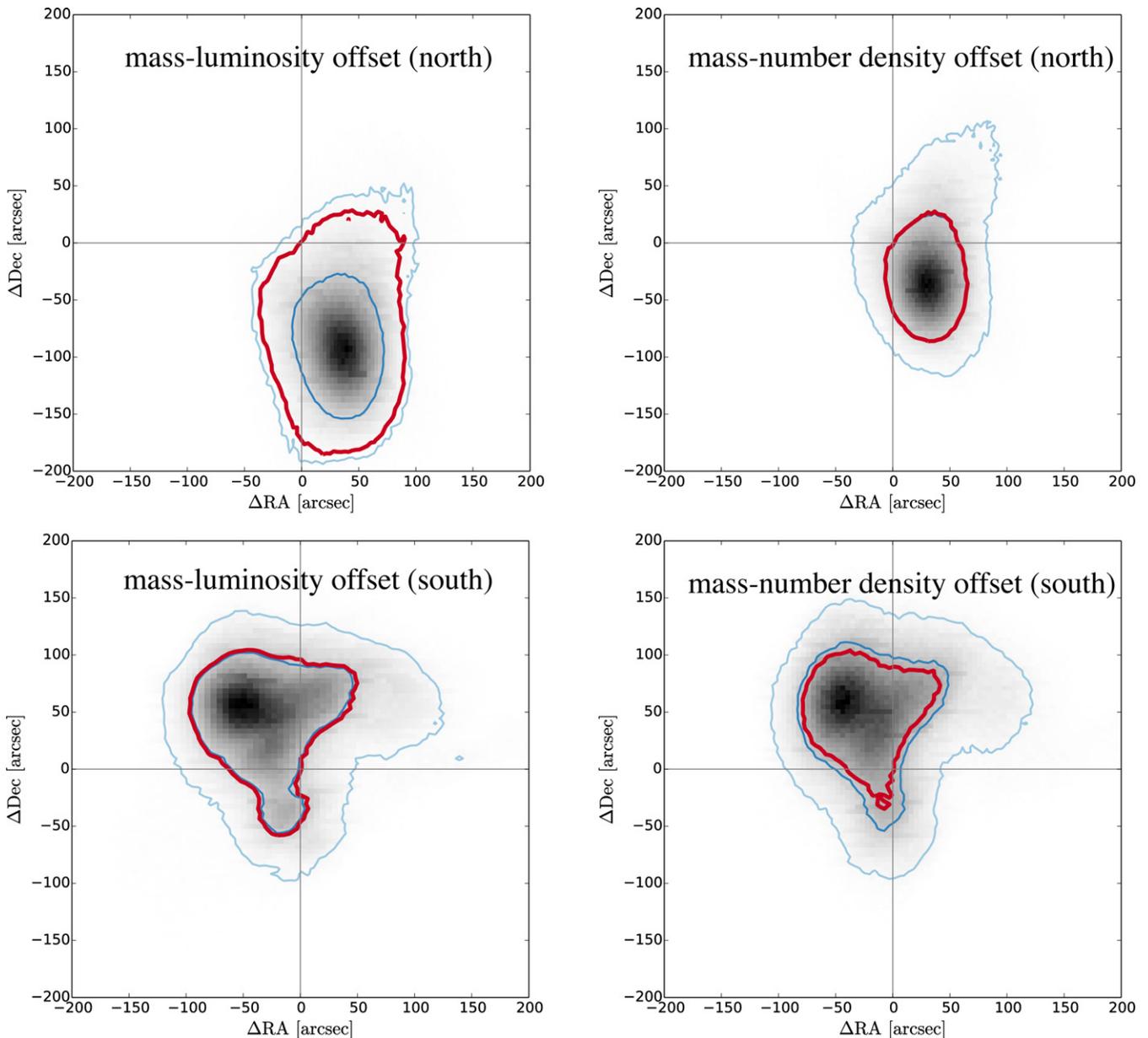
Figure 10. Significance of mass–galaxy offset. We examine the spatial distribution of the mass–galaxy offsets extracted from the bootstrap resampling. The blue contours show the 68% and 95% confidence limits. The origin in each plot corresponds to the null hypothesis of no offset, and the red contour shows the confidence limit that just touches this point in parameter space. The corresponding *p*-values are listed in Table 3.

same volume (at a small radius), it may be possible to cause the observed disparity in radio relics if the northern cluster has a denser core.

7.3. Mass–Galaxy Offset

The comparison of the mass and galaxy centroids in CIZA J2242.8+5301 shows that on average they are offset by $\sim1'$. The offset may be regarded as interesting in the context of a SIDM theory with a non-negligible collisional cross-section, which predicts that in general mass and galaxy distributions will differ after collision in merging clusters. In CIZA J2242.8+5301 the northern galaxy centroid appears to lead the corresponding mass centroid whereas the mass–galaxy offset is rather lateral for the southern clump. Also, the axis defined by the line connecting the two mass centroids is rotated by $\sim10°$ with respect to the merger axis defined by the cluster mass, X-ray gas, and radio relics.

In order to investigate the significance of the mass centroid, we re-sample source galaxies while allowing duplication. Similarly, we bootstrap the cluster galaxies as well to estimate the uncertainty of their centroid in both luminosity and number density. The resulting centroid distributions are displayed in Figure 9. We note that the luminosity centroid uncertainty shown here should be considered an upper limit because we assign equal probabilities of elimination/selection to all galaxies regardless of their luminosity; a tighter centroid distribution would be obtained if brighter galaxies were made harder to drop in the re-sampled catalog.

Before we discuss the mass–galaxy offset, it is instructive to examine individual contours in detail. First, we note that the mass centroid contours closely resemble the shapes of the mass peaks (Figure 5). Second, the northern mass peak, which is more peaked, has a tighter centroid distribution. Third, the galaxy centroid distributions also resemble the corresponding

Table 3
Significance of the Mass–Galaxy Offsets in CIZA J2242.8+5301

Subclusters	Mass–Luminosity	Mass–Number
North	0.082	0.310
South	0.288	0.431

Note. We quote p-values for the null hypothesis that the mass and galaxy centroids coincide.

smoothed galaxy (both luminosity and number density) distribution. Fourth, the size of the galaxy centroid uncertainties are similar between the northern and southern clumps.

The significance of the mass–galaxy offset is not easily deduced from Figure 9. One question is whether number density or luminosity is a better tracer of the galaxy population. Luminosity is a better tracer of the mass in stars in principle, but its location can be more highly disturbed by interloping foreground galaxies. We therefore tabulate both types of offset in both north and south subclusters, yielding four potential offsets. In Figure 10, we plot the spatial distribution of these four types of offsets extracted from the bootstrap resampling. The blue contours show the 68% and 95% confidence limits. The origin in each plot corresponds to the null hypothesis of no offset, and the red contour shows the confidence limit that just touches this point in parameter space. The corresponding p-values are listed in Table 3. The lowest p-value is seen in the mass–luminosity offset in the north, meaning that the probability of detecting the observed offset by a random chance is $\sim 8\%$.

Can we use the current mass–galaxy (luminosity) offset for the argument of SIDM? We believe that it is still premature to conclude that our observation supports SIDM because of the following caveats. First, we do not know the ICM mass exactly. As seen in Figure 1, the X-ray emission is more centrally concentrated than the cluster galaxies, which may imply that a significant fraction of ICM may be present between the two halos of CIZA J2242.8+5301. Hence, one can suspect that our mass centroids are somewhat pulled inward (toward the cluster center) if the fraction of the mass locked up in the ICM is significant. Dawson (2013) showed that for the Musket Ball Cluster (Dawson et al. 2012), this systematic offset can be of the same order of magnitude as the weak-lensing and galaxy centroid uncertainties. Second, we cannot yet exclude the possibility that any line-of-sight interlopers are present. This issue may be resolved in the future when we reach more completeness in our spectroscopic survey of the field. Currently, our existing spectroscopic catalog does not hint at the possibility. Third, our weak-lensing source catalog may be "asymmetrically" contaminated by the cluster galaxies. With our dependence on the color–magnitude relation in source selection, we expect that our source catalog contains some blue cluster galaxies inevitably. Typically, we assume that the contamination is azimuthally symmetric in isolated clusters. However, in clusters with elongated morphology like CIZA J2242.8+5301, the assumption may turn out to be incorrect, especially, when the merger triggers star-formation activity preferentially along the merger axis in the outskirts (Stroe et al. 2014c). Fourth, interpretation of galaxy ellipticity is ambiguous in the strong lensing regime. That is, we do not know whether a galaxy ellipticity e becomes g or $1/g^*$ a priori. The two cases can be reliably distinguished only when strong lensing data with known redshifts are used. Because CIZA J2242.8+5301 is a merging cluster, a probable shape of the critical curve is a long

ellipse elongated in the north–south direction similar to the one modeled by Zitrin et al. (2013) for "El Gordo." Given the weak-lensing mass of CIZA J2242.8+5301, the Einstein radius of each halo is likely to be $r = 0\rlap{.}{''}5 \sim 1'$, which is comparable to the size of the observed mass–galaxy offset. Therefore, it is worth investigating in the future how weak-lensing mass centroids are affected in a similar mass configuration.

8. CONCLUSIONS

We have presented a detailed weak gravitational lensing analysis of CIZA J2242.8+5301, which is a rare target boasting a textbook example of a giant radio relic. Our mass reconstruction shows that the cluster consists of two halos of nearly equal mass, consistent with the cluster galaxy distribution. In order to quantify the mass while minimizing the contamination bias, we perform MCMC by simultaneously fitting the two halos with NFW profiles. Although we adopt the luminosity centers as our initial values, the mass centroids are not fixed, but allowed to move during our MCMC run. We obtain $M_{200c} = 11.0^{+3.7}_{-3.2} \times 10^{14}\ M_\odot$ and $9.8^{+3.8}_{-2.5} \times 10^{14}\ M_\odot$ for the northern and southern halos, respectively. The total mass of the system is $M_{200c} = (2.51 \pm 0.53) \times 10^{15}\ M_\odot$ at $r_{200c} = 2.63$ Mpc, which makes CIZA J2242.8+5301 among the most massive clusters detected so far. Because of the low redshift of the cluster, this mass alone does not cause any tension with the ΛCDM paradigm. This large cluster mass suggests that the previous N-body simulation assuming the total mass $\sim 5.5 \times 10^{14} M_\odot$ should be revised.

We detect somewhat large offsets ($\sim 1'$) between galaxy and mass centroids. To quantify the significance, we run bootstrapping for both mass and galaxy distributions and find that the offsets are significant at least at the 2σ level for the northern clump. Although the direction of the offset is in favor of SIDM at face value, we identify some caveats that may lead to the current observation. They are (1) the gas mass that can pull the total mass (galaxy+gas+dark matter) centroids toward the cluster center with respect to the galaxies, (2) a possible line-of-sight structure that is uncorrelated with the cluster, but may bias our mass centroids, (3) the cluster galaxy contamination in our source catalog, whose contamination rate may differ azimuthally and thus shift the two weak-lensing mass centroids, and (4) the ambiguity of shear interpretation between strong and weak lensing regimes. We believe that all these four issues can be addressed in the future as more concerted efforts are made to enhance our understanding of the system.

REFERENCES

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543

Akamatsu, H., & Kawahara, H. 2013, PASJ, 65, 16

Bartelmann, M., & Schneider, P. 2001, PhR, 340, 291

Bertin, E. 2006, in ASP Conf. Ser. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique (San Francisco, CA: ASP), 112

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in ASP Conf. Ser. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco, CA: ASP), 228

Clowe, D., Bradač, M., Gonzalez, A. H., et al. 2006, ApJL, 648, L109

Clowe, D., Luppino, G. A., Kaiser, N., & Gioia, I. M. 2000, ApJ, 539, 540

Dawson, W. A. 2013, PhD thesis University of California, Davis

Dawson, W. A., Jee, M. J., Stroe, A., et al. 2014, arXiv:1410.2893

Dawson, W. A., Wittman, D., Jee, M. J., et al. 2012, ApJL, 747, L42

Duffy, A. R., Schaye, J., Kay, S. T., & Dalla Vecchia, C. 2008, MNRAS, 390, L64

Ensslin, T. A., Biermann, P. L., Klein, U., & Kohle, S. 1998, A&A, 332, 395

Evrard, A. E., Bialek, J., Busha, M., et al. 2008, ApJ, 672, 122

Faber, S. M., Phillips, A. C., Kibrick, R. I., et al. 2003, Proc. SPIE, 4841, 1657

Fahlman, G., Kaiser, N., Squires, G., & Woods, D. 1994, ApJ, 437, 56

Hoekstra, H. 2013, arXiv:1312.5981

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Jee, M. J., Blakeslee, J. P., Sirianni, M., et al. 2007a, PASP, 119, 1403

Jee, M. J., Ford, H. C., Illingworth, G. D., et al. 2007b, ApJ, 661, 728

Jee, M. J., Hoekstra, H., Mahdavi, A., & Babul, A. 2014a, ApJ, 783, 78

Jee, M. J., Hughes, J. P., Menanteau, F., et al. 2014b, ApJ, 785, 20

Jee, M. J., Mahdavi, A., Hoekstra, H., et al. 2012, ApJ, 747, 96

Jee, M. J., & Tyson, J. A. 2011, PASP, 123, 596

Jee, M. J., Tyson, J. A., Schneider, M. D., et al. 2013, ApJ, 765, 74

Kaiser, N., & Squires, G. 1993, ApJ, 404, 441

Kahlhoefer, F., Schmidt-Hoberg, K., Frandsen, M. T., & Sarkar, S. 2014, MNRAS, 437, 2865

Katgert, P., Katgert-Merkelijn, J. K., Le Poole, R. S., & van der Laan, H. 1973, A&A, 23, 171

Kocevski, D. D., Ebeling, H., Mullis, C. R., & Tully, R. B. 2007, ApJ, 662, 224

Mahdavi, A., Hoekstra, H., Babul, A., Balam, D. D., & Capak, P. L. 2007, ApJ, 668, 806

Mandelbaum, R., Rowe, B., Armstrong, R., et al. 2014, arXiv:1412.1825

Menanteau, F., Hughes, J. P., Sifón, C., et al. 2012, ApJ, 748, 7

Merten, J., Coe, D., Dupke, R., et al. 2011, MNRAS, 417, 333

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Ogrean, G. A., Brüggen, M., Röttgering, H., et al. 2013, MNRAS, 429, 2617

Ogrean, G. A., Brüggen, M., van Weeren, R., et al. 2014, MNRAS, 440, 3416

Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004, ApJ, 611, 660

Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, A&A, 498, 361

Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523

Rocha, M., Peter, A. H. G., Bullock, J. S., et al. 2013, MNRAS, 430, 81

Sarazin, C. L. 2002, in Merging Processes in Galaxy Clusters, ed. L. Feretti, I. M. Gioia, & G. Giovannini (Astrophysics and Space Science Library, Vol. 272; Dordrecht: Kluwer), 1

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103

Seitz, C., & Schneider, P. 1997, A&A, 318, 687

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Stroe, A., Rumsey, C., Harwood, J. J., et al. 2014a, MNRAS, 441, L41

Stroe, A., Sobral, D., Dawson, W., et al. 2014b, arXiv:1410.2891

Stroe, A., Sobral, D., Röttgering, H. J. A., & van Weeren, R. J. 2014c, MNRAS, 438, 1377

Stroe, A., van Weeren, R. J., Intema, H. T., et al. 2013, A&A, 555, A110

Swarup, G., Ananthakrishnan, S., Kapahi, V. K., et al. 1991, CSci, 60, 95

van Weeren, R. J., Brüggen, M., Röttgering, H. J. A., & Hoeft, M. 2011, MNRAS, 418, 230

van Weeren, R. J., Röttgering, H. J. A., Brüggen, M., & Hoeft, M. 2010, Sci, 330, 347

Yagi, M., Kashikawa, N., Sekiguchi, M., et al. 2002, AJ, 123, 66

Zitrin, A., Menanteau, F., Hughes, J. P., et al. 2013, ApJL, 770, 15

Zwart, J. T. L., Barker, R. W., Biddulph, P., et al. 2008, MNRAS, 391, 1545