
Universiteit
Leiden
The Netherlands

MC^2: Galaxy Imaging and Redshift Analysis of the Merging Cluster CIZA J2242.8+5301

Dawson, W.; Jee, M.; Stroe, A.; Ng, Y.; Golovich, N.; Wittman, D.; ... ; Weeren, R.

Citation
Dawson, W., Jee, M., Stroe, A., Ng, Y., Golovich, N., Wittman, D., ... Weeren, R. (2015). MC^2: Galaxy Imaging and Redshift Analysis of the Merging Cluster CIZA J2242.8+5301. *The Astrophysical Journal*, *805*(2), 143. doi:10.1088/0004-637X/805/2/143

Version:	Not Applicable (or Unknown)
License:	Leiden University Non-exclusive license
Downloaded from:	https://hdl.handle.net/1887/48591

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

The Astrophysical Journal, 805:143 (18pp), 2015 June 1 doi:10.1088/0004-637X/805/2/143
© 2015. The American Astronomical Society. All rights reserved.

MC^2: GALAXY IMAGING AND REDSHIFT ANALYSIS OF THE MERGING CLUSTER CIZA J2242.8+5301

William A. Dawson[1], M. James Jee[2], Andra Stroe[3], Y. Karen Ng[2], Nathan Golovich[2], David Wittman[2], David Sobral[3,4,5], M. Brüggen[6], H. J. A. Röttgering[3], and R. J. van Weeren[7]
[1] Lawrence Livermore National Lab, 7000 East Avenue, Livermore, CA 94550, USA; dawson29@llnl.gov
[2] University of California, One Shields Avenue, Davis, CA 95616, USA
[3] Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands
[4] Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, PT1349-018, Portugal
[5] Center for Astronomy and Astrophysics of the University of Lisbon, Tapada da Ajuda—Edificio Leste—2° Piso, 1349-018 Lisbon, Portugal
[6] Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany
[7] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
Received 2014 October 12; accepted 2015 February 26; published 2015 May 28

ABSTRACT

X-ray and radio observations of CIZA J2242.8+5301 suggest that it is a major cluster merger. Despite being well studied in the X-ray and radio, little has been presented on the cluster structure and dynamics inferred from its galaxy population. We carried out a deep ($i < 25$) broadband imaging survey of the system with Subaru SuprimeCam (g and i bands) and the Canada–France–Hawaii Telescope (r band), as well as a comprehensive spectroscopic survey of the cluster area (505 redshifts) using Keck DEep Imaging Multi-Object Spectrograph. We use these data to perform a comprehensive galaxy/redshift analysis of the system, which is the first step to a proper understanding of the geometry and dynamics of the merger, as well as using the merger to constrain self-interacting dark matter. We find that the system is dominated by two subclusters of comparable richness with a projected separation of $6\farcm9^{+0.7}_{-0.5}$ ($1.3^{+0.13}_{-0.10}$ Mpc). We find that the north and south subclusters have similar redshifts of $z \approx 0.188$ with a relative line-of-sight (LOS) velocity difference of 69 ± 190 km s^{-1}. We also find that north and south subclusters have velocity dispersions of 1160^{+100}_{-90} and 1080^{+100}_{-70} km s^{-1}, respectively. These correspond to masses of $16.1^{+4.6}_{-3.3} \times 10^{14}$ and $13.0^{+4.0}_{-2.5} \times 10^{14}$ $M_\odot$, respectively. While velocity dispersion measurements of merging clusters can be biased, we believe the bias in this system to be minor due to the large projected separation and nearly plane-of-sky merger configuration. We also find that the cDs of the north and south subclusters are very near their subcluster centers, in both projection (55 and 85 kpc, respectively) and normalized LOS velocity ($|\Delta v|/\sigma_v = 0.43 \pm 0.13$ and 0.21 ± 0.12 for the north and south, respectively). CIZA J2242.8+5301 is a relatively clean dissociative cluster merger with near 1:1 mass ratio, which makes it an ideal merger for studying merger-associated physical phenomena.

Key words: galaxies: clusters: individual (CISA) – galaxies: distances and redshifts

Supporting material: machine-readable table

1. INTRODUCTION

Under the hierarchical structure formation paradigm all clusters are formed from merging substructures. When the mergers involve two approximately equal mass subclusters, a *dissociative* merger can occur where the baryonic plasma of each subcluster collides, forms shocks, is slowed relative to the effectively collisionless galaxies, and becomes dissociated for a time post-merger (some examples include: the Bullet Cluster, Clowe et al. 2004; the Musket Ball Cluster, Dawson et al. 2012; and Pandora's Cluster, Merten et al. 2011). These plasma shocks can lead to sharp X-ray bow shock features (Markevitch et al. 2002, 2005) and, coupled with the intra-cluster magnetic fields, can lead to radio relics, which are diffuse synchrotron sources typically at the periphery of cluster mergers (see Feretti et al. 2012, for a review). It is still unclear exactly what effect these merger-related phenomena have on the constituent galaxies. There is observational evidence that cluster mergers trigger star formation (e.g., Miller & Owen 2003; Ferrari et al. 2005; Owen et al. 2005; Hwang & Lee 2009), quench it (Poggianti et al. 2004), or have no immediate effect (Chung et al. 2010). In addition to enabling the study of baryonic physical phenomena, merging clusters can be used to constrain the dark matter (DM) self-interaction cross section by comparing the location of the DM (measured through gravitational lensing) with the location of the collisonal gas and effectively collisionless galaxies (e.g., Randall et al. 2008; Dawson et al. 2012). There are seemingly conflicting results where galaxy-DM offsets have been observed in some systems (A520, Jee et al. 2012, 2014a; and the Musket Ball, Dawson 2013b) but not in others (the Bullet Cluster, Bradač et al. 2006; El Gordo, Jee et al. 2014b).

In an attempt to resolve some of these apparent discrepancies and properly infer the underlying micro-physics, we have formed the Merging Cluster Collaboration[8] (MC^2), which is undertaking a systematic X-ray, broad/narrowband optical, spectroscopic, and radio survey of an ensemble of merging clusters. In this paper we will present the global galactic properties of CIZA J2242.8+5301, the first merger of this systematic approach. Jee et al. (2014c) present the weak-lensing analysis of this system, Stroe et al. (2014b) present a narrowband Hα galaxy study of this system, and Sobral et al. (2015) present spectroscopic cluster galaxy evolution analyses (using the spectra discussed in this paper).

CIZA J2242.8+5301 (aka the Sausage) was first discovered by Kocevski et al. (2007) in the second Clusters in the Zone of Avoidance (CIZA) sample, which is a survey of clusters of

[8] http://www.mergingclustercollaboration.org.

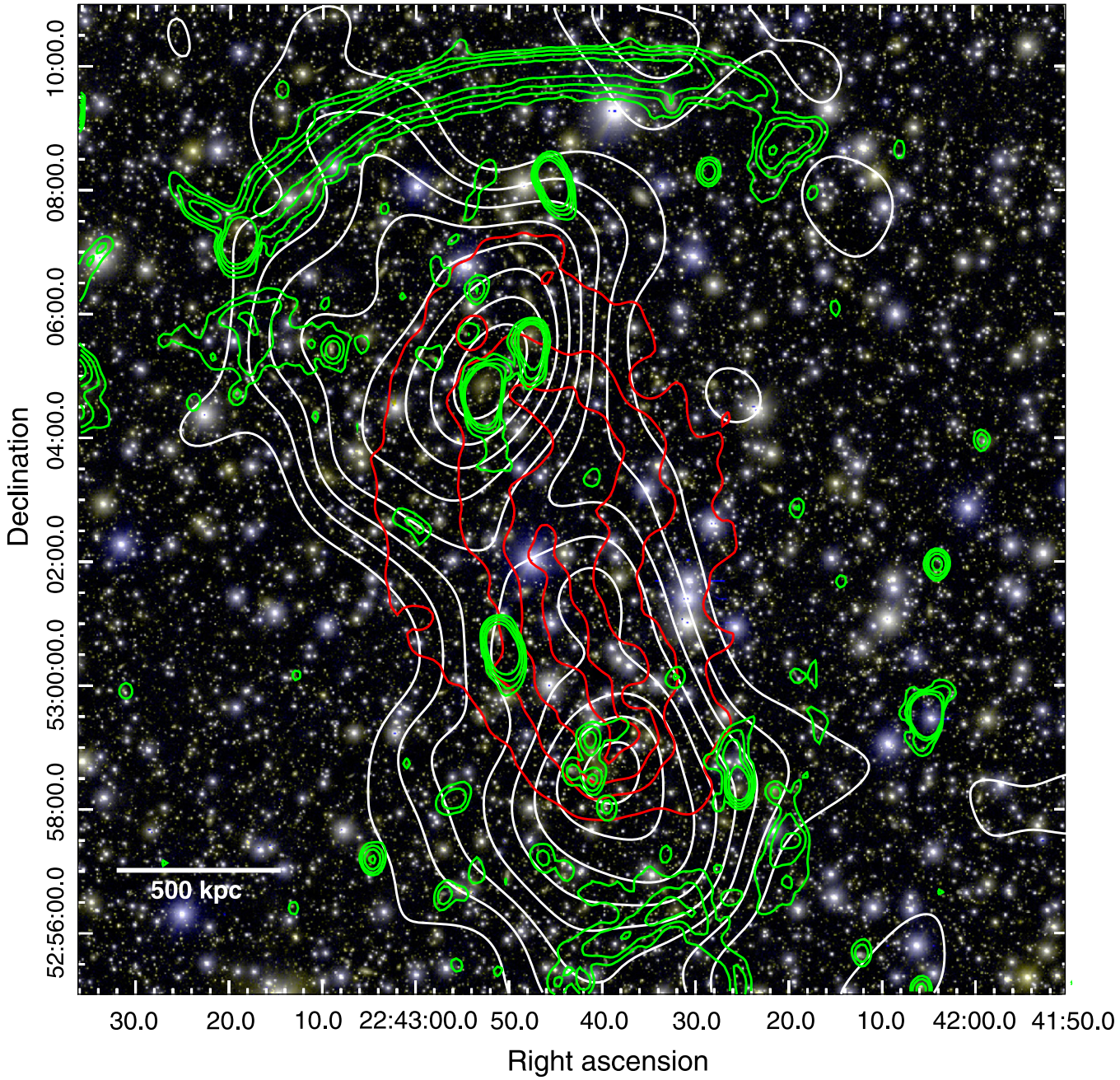

Figure 1. Subaru *gi* color composite image of CIZA J2242.8+5301. The red contours are a linear scale mapping of the *XMM-Newton* X-ray luminosity map. The green contours are a linear scale mapping of the WSRT radio emission, and the radio relics are the extended and diffuse emission near the periphery of the north and south subclusters. The cluster galaxy number density contours (white) based on our red sequence selection begin at 100 galaxies Mpc^{-2} and increase linearly with increments of 25 galaxies Mpc^{-2} (copied from Figure 10). CIZA J2242.8+5301 is an example of a dissociative radio relic merger, with two radio relics at the periphery and the bulk of the cluster gas dissociated between two subclusters.

galaxies behind the Milky Way. Its galactic coordinates are (104°11′20.″61, −05°06′15.″87), so it is very near the disk of the Galaxy but away from the bulge. This cluster is in a field with high Galactic dust extinction ($A_v = 1.382$; Schlafly & Finkbeiner 2011), which is likely the reason there have been limited optical studies of the system (with the exception of our recent work and Stroe et al. 2014c's Hα study).

Van Weeren et al. (2010) conducted the first comprehensive radio survey of the system (including Westerbork Synthesis Radio Telescope, Giant Meter-wave Radio Telescope, and Very Large Array observations). They observed two radio relics at the north and south periphery of the cluster (see green contours in Figure 1). These radio relics are elongated diffuse radio emission (approximately 10:1 length-to-width ratios). These radio relics are evidence of shock acceleration and spectral aging associated with the outward-moving shock (this was later confirmed with the follow-up study of Stroe et al. 2013). They also observed that the northern relic is strongly polarized at the 50%–60% level and used this to infer that the merger angle must be within ∼30° of the plane of the sky. They also used the spectral index to infer a Mach number of ∼4.6. Stroe et al. (2014a) performed a more detailed spectral age modeling of the radio relic and found a slightly lower Mach number of $2.9^{+0.10}_{-0.13}$.

Three detailed X-ray analyses of CIZA J2242.8+5301 have been conducted, one with *XMM-Newton* (Ogrean et al. 2013), one with *Suzaku* (Akamatsu & Kawahara 2013), and one with *Chandra* (Ogrean et al. 2014). The Ogrean et al. (2013) *XMM-Newton* analysis shows an extreme N-S elongation of the X-ray gas largely consistent with the merger axis suggested by the radio relics (see red contours of Figure 1). The *XMM-Newton* instrumental background levels prevent them from characterizing the surface brightness profile at the location of the northern radio relic (this is remedied by the Akamatsu & Kawahara 2013 *Suzaku* observations); however, near the southern radio relic they find evidence for a shock with Mach number ~1.2–1.3. Ogrean et al. (2013) also note two interesting features of the gas. The first feature is a "wall" of hot gas east of the cluster center, and while not associated with a radio relic, it does extend into the region behind the southern relic. They note that a simple binary merger is not expected to create such a feature and suggest that it may be indicative of a more complex merger scenario (e.g., a triple merger), or "a lack of understanding on our part of the complex structures formed during real cluster mergers." The second feature is a "smudge" of enhanced X-ray emission coincident with the eastern 1/5 of the northern radio relic. Akamatsu & Kawahara (2013) found evidence for a temperature jump at the location of the northern radio relic corresponding to a Mach number of 3.15 ± 0.52, consistent with the radio-estimated Mach number of Stroe et al. (2014a). Akamatsu & Kawahara (2013) did not see a jump in the surface brightness profile, but they claim that this is due to the large *Suzaku* point-spread function (PSF) (~380 kpc) being much larger than the width of the relic (~55 kpc). Ogrean et al. (2014) found evidence for two inner density discontinuities, trailing the northern and southern radio relics by ~0.5 Mpc. They argue that these discontinuities are not likely cold fronts given that their large distance from the cluster center (≈1.5 Mpc) would make them the most distant cold fronts ever detected. Additionally, the measured temperature of ~8–9 keV would make them the hottest of all known cold fronts. Instead, they argue that the inner density discontinuities could be caused by the violent relaxation of DM tidal tails that were generated at the far sides of the DM halos post-merger.

A number of simulations of the system have been performed. Van Weeren et al. (2011) conducted a suite of simulations studying potential analogs to the system and argue that CIZA J2242.8+5301 is undergoing a merger in the plane of the sky ($\lesssim$10° from edge-on), with a mass radio of about 2:1, an impact parameter of $\lesssim$400 kpc, and a core pass that happened about 1 Gyr ago. Interestingly, they suggest that the southern subcluster should be slightly less massive, given the relative size of the southern relic. Kang et al. (2012) conducted diffusive shock acceleration simulations of the Sausage and found that Mach numbers from 2 to 4.5 were supported depending on the amount of pre-existing cosmic-ray electrons. However, they question the ability of the merger event to produce such an elongated shock.

The only thorough optical analysis of CIZA J2242.8+5301 to date was an Hα survey conducted by Stroe et al. (2014c). They find an order-of-magnitude boost in the normalization of the Hα galaxy luminosity function in the vicinity of the relics, even greater than that of other known mergers at the same redshift. One important note is that they made assumptions about the contamination of their cluster Hα population. Stroe et al. (2014b) have used the redshifts presented in this paper to show that their original assumptions were overly conservative. Stroe et al. (2014b) find an even larger boost than Stroe et al. (2014c) based on updated contamination estimates.

In this paper we add to this picture with broadband optical and spectroscopic analyses of CIZA J2242.8+5301, which are key components to properly interpreting the merger. In Section 2 we discuss our Issac Newton Telescope Wide Field Camera, Subaru SuprimeCam, and Canada–France–Hawaii Telescope (CFHT) Megacam observations. In Section 3 we discuss our Keck DEep Imaging Multi-Object Spectrograph (DEIMOS; Faber et al. 2003) and William Hershel Telescope (WHT) AF2 spectroscopic observations. In Section 4 we discuss our spectroscopic and imaging selection of cluster members. In Section 5 we discuss our identification of the system's subclusters, and in Section 6 we present the galactic properties of those subclusters. In Section 7.1 we place the current work in context of the X-ray and radio analyses of the system and where necessary revise existing interpretations. Finally, in Section 8 we summarize our results. As we previously noted, companion papers present the weak-lensing analysis (Jee et al. 2014c), a narrowband Hα galaxy study (Stroe et al. 2014b), and a spectroscopic cluster galaxy evolution analysis of this system (using the imaging and spectra discussed in this paper; Sobral et al. 2015).

We assume a flat ΛCDM universe with $H_0 = 70\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. At the redshift of the cluster ($z = 0.188$), 1′ corresponds to 189 kpc. Magnitudes are in the AB system.

2. OBSERVATIONS: IMAGING

We first observed CIZA J2242.8+5301 in the optical using the Wide Field Camera (WFC) on the 2.5 m Issac Newton Telescope (INT) at the Roque de Los Muchachos Observatory on La Palma. We carried out the observations over two nights (2009 October 06–07), observing the system in the *B*, *V*, *R*, and *I* filters with total exposure times of 12,000, 9000, 9000, and 9000 s, respectively. The data reduction was carried out with IRAF and the *mscred* package (Valdes 1998). Standard bias and flat-field corrections were carried out, and the *R*- and *I*-band images were fringe corrected with *rmfringe*. As a final step the images were registered to the 2MASS WCS coordinate system and co-added, rejecting pixels above $3.0\sigma_{\mathrm{rms}}$. The seeing ranged from 1.″5 to 2″. This relatively large PSF, coupled with the high stellar densities in the low galactic latitude field, made it difficult to discriminate between stars and galaxies when we used the imaging for spectroscopic target selection, as discussed in Section 3.1.1.

CIZA J2242.8+5301 was observed with CFHT MegaCam during queue scheduling during 2013 July 3–12 in *r* (P.I. A. Stroe). The total integration is 24,000 s, consisting of 40 short (600 s) exposures. The median seeing is ~0.″74, and the 5σ limiting magnitude is 24.1 (~1.3 M^* at $z = 0.188$). We also observed CIZA J2242.8+5301 with Subaru SuprimeCam on 2013 July 13 in *g* and *i* (P.I. D. Wittman), for a total integration time of 700 s in *g*, consisting of four 180 s exposures, and a total integration time of 3060 s in *i*, consisting of eight 360 s and one 180 s exposures. We rotated the field between each exposure (30° for *g* and 15° for *i*) 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 detected galaxies. The median seeing for *g* and *i*

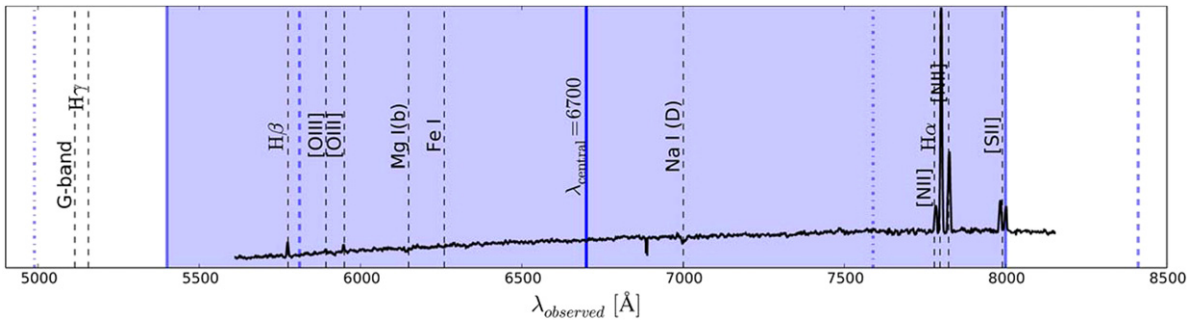

Figure 2. Spectral coverage of the Keck DEIMOS observations (shaded blue region), along with the redshifted location of common cluster emission and absorption features (black dashed lines). The blue dot-dashed pair and the blue dashed pair of lines show the variable range depending on where the slit was located along the width of the slit mask. The solid black line shows an example galaxy spectrum from our DEIMOS survey.

images is 0.″72 and 0.″65, respectively. The observed 5σ limiting magnitudes are 24.9 and 25.4 for the g and i filters. These limiting magnitudes correspond to M^*_g+4.9 and M^*_i+7.9 at $z = 0.188$, assuming $M^*_B(z = 0) = -20.5$ (Norberg et al. 2001), K-corrections $K_{g-B} = -0.6$ and $K_{i-B} = -2.3$ (Frei & Gunn 1994), and extinctions $A_g = 1.417$ and $A_i = 0.729$ (Schlafly & Finkbeiner 2011). The details of the CFHT and Subaru data reduction and photometric dust extinction correction are presented in Jee et al. (2014c) and Stroe et al. (2014b).

3. OBSERVATIONS: SPECTROSCOPIC

3.1. Keck DEIMOS Observations

We conducted a spectroscopic survey of CIZA J2242.8+5301 with the DEIMOS instrument on the Keck II 10 m telescope over two observing runs on 2013 July 14 and 2013 September 05. Observations during both runs were taken using 1″ wide slits with the 1200 line mm^{-1} grating, tilted to a central wavelength of 6700 Å, resulting in a pixel scale of 0.33 Å pixel^{-1}, a resolution of ∼1 Å (50 km s^{-1}), and typical wavelength coverage of 5400–8000 Å, shown in Figure 2. The actual wavelength coverage may be shifted by ∼±410 Å depending on where the slit is located along the width of the slit mask. For most cluster members this enabled us to observe Hβ, [O III] $\lambda\lambda$4960 and 5008, Mg I (b), Fe I, Na I (D), [O I], Hα, and the [N II] doublet (Figure 2). This spectral setup enables us to also study the star formation properties of the cluster galaxies; see related work by Sobral et al. (2015). The position angle (PA) of each slit was chosen to lie between ±5° and 30° of the slit mask PA to achieve optimal sky subtraction[9] during reduction with the DEEP2 version of the spec2d package (Newman et al. 2013). Within this range the slit PA was chosen to minimize the effects of chromatic dispersion by the atmosphere by aligning the slit, as much as possible, with the axis connecting the horizon, object, and zenith (see e.g., Filippenko 1982). We observed a total of four slit masks with approximately 120 slits per mask. For each mask we took three 900 s exposures.

Since the central wavelength of 6700 Å is bluer than typical DEIMOS setups, we found it necessary to modify the default DEIMOS arc lamp calibration procedure. We began by turning the Hg, Ne, Cd, Kr, Ar, and Zn lamps on at the same time, after 1 s we turned off the Hg and Ne lamps, after 7 s we turned off the Cd lamp, after 8 s we turned off the Kr lamp, and we stopped exposing after 16 s. We found this sequence necessary to prevent the brighter emission lines on the red side from saturating while exposing long enough to get lines of sufficient signal on the blue side.

3.1.1. Keck DEIMOS: Target Selection

Our primary objective for the spectroscopic survey was to maximize the number of cluster member spectroscopic redshifts. Since the SuprimeCam imaging was unavailable at the time of our spectroscopic survey planning, we used the WFC imaging to determine the approximate red sequence of the cluster and create a galaxy number density map. The DEIMOS 5′ × 16.′7 field of view is well suited to survey the elongated CIZA J2242.8+5301 system, ∼7′ × 15′, and we aligned the long axis of our slit masks with the long axis of the system. Cluster member target selection was challenging due to the low galactic latitude ($b = -5°$) with a stellar surface density approximately 2.7 times the galaxy surface density, as well as variable extinction ($\Delta E(B - V) \sim$ 0.4–0.6 mag) across the field (Schlafly & Finkbeiner 2011; Stroe et al. 2014c). The difficulty of star–galaxy separation is also compounded by the 1.″5–2″ seeing of the INT/WFC imaging, which results in many stars being blended (especially binary pairs), which results in many blended pairs of stars passing morphological cuts designed to eliminate point sources. We find that the majority of the stars in the field are bluer than the cluster galaxy population; thus, we did not target any object with $R - I < 0.9$. We found it difficult to clearly define the cluster red sequence due to variable extinction across the field plus the red star contamination. Thus, rather than exclude galaxies redder than the brightest cluster galaxy (BCG; $R - I = 1.2$), we linearly down-weighted the probability of selecting galaxies redder than the BCG as a function of their $R - I$ color. In addition to these weights, we weighted each galaxy's probability of being targeted by $10^{-(R-22)}$, thus preferentially selecting brighter galaxies likely to have higher signal-to-noise ratios (S/Ns). We then divided our potential targets into a bright sample (Sample 1; $R < 22.5$) and a faint sample (Sample 2; $22.5 < R < 23.5$). We first filled our mask with as many Sample 1 targets as possible and then filled in the remainder of the mask with Sample 2 targets.

We used the DSIMULATOR package[10] to design each slit mask. DSIMULATOR automatically selects targets by maximizing the sum total weights of target candidates, by first selecting as many objects from Sample 1 as possible and then

[9] http://astro.berkeley.edu/∼cooper/deep/spec2d/slitmask.html.

[10] http://www.ucolick.org/∼phillips/deimos_ref/masks.html.

Table 1
Keck DEIMOS Redshift Catalog of the CIZA J2242.8+5301 Field

R.A. (hh:mm:ss.sss)	Decl. (dd:mm:ss.ss)	z	σz	i (mag)	σi (mag)
22:42:43.719	+52:54:17.317	0.187259	0.000026	18.792	0.003
22:42:50.008	+52:54:17.651	0.184404	0.000057	18.949	0.003
22:42:51.259	+52:54:22.113	0.183752	0.000005	17.833	0.002
22:42:36.834	+52:54:48.770	0.454690	0.000044	19.617	0.006
22:43:00.240	+52:54:59.057	0.186943	0.000023	18.650	0.002

(This table is available in its entirety in machine-readable form.)

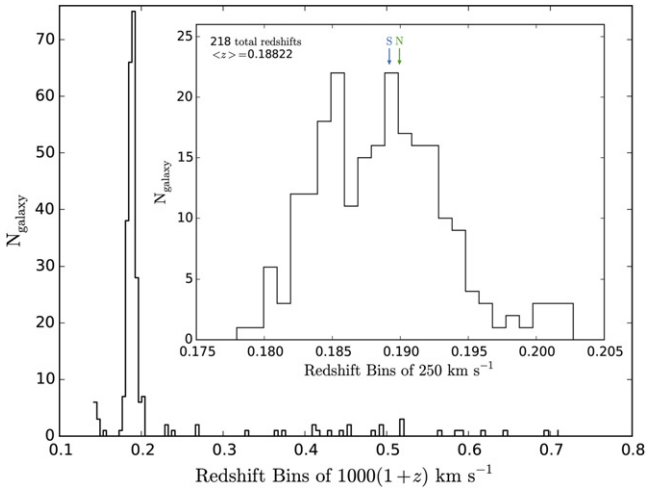

Figure 3. *Main:* redshift distribution of the Keck DEIMOS high-quality ($Q \geqslant 3$) galaxy spectroscopic redshifts. The overdensity near the cluster redshift $z = 0.188$ is clear, with 206 spectroscopic galaxies near the cluster redshift, 15 foreground galaxies, and 34 background galaxies. *Inset:* zoom-in of the spectroscopic histogram near the cluster redshift. The north brightest-cluster-galaxy (BCG) redshift is indicated by the green arrow, and the southern BCG redshift is indicated by the blue arrow.

filling in the remaining area of the slit mask with target candidates from Sample 2. We manually edited the automated target selection to increase the number of selected targets, e.g., by selecting another target between targets selected automatically by DSIMULATOR if it resulted in a small loss of sky coverage to their slits.

While we preferentially targeted likely red sequence cluster members, it was not always possible to fill the entire mask with these galaxies, in which case we would place a slit on other galaxies in the field. In our 2013 July 14 observations we serendipitously observed nine galaxies from the Stroe et al. (2014c) Hα catalog. In our 2013 September 05 observations we purposefully targeted 17 galaxies from that catalog.

3.1.2. Keck DEIMOS: Data Reduction

The exposures for each mask were combined using the DEEP2 versions of the *spec2d* and *spec1d* packages (Newman et al. 2013). This package combines the individual exposures of the slit mosaic and performs wavelength calibration, cosmic-ray removal, and sky subtraction on a slit-by-slit basis, generating a processed two-dimensional spectrum for each slit. The *spec2d* pipeline also generates a processed one-dimensional spectrum for each slit. This extraction creates a one-dimensional spectrum of the target, containing the summed flux at each wavelength in an optimized window. The *spec1d* pipeline then fits template spectral energy distributions (SEDs) to each one-dimensional spectrum and estimates a corresponding redshift. There are SED templates for various types of stars, galaxies, and active galactic nuclei (AGNs). We then visually inspect the fits using the *zspec* software package (Newman et al. 2013), assign quality rankings to each fit (following a convention closely related to Newman et al. 2013), and manually fit for redshifts where the automated pipeline failed to identify the correct fit. The highest-quality galaxy spectra ($Q = 4$; 229 total) have a mean S/N of 10.9 and standard deviation of 5.5, while the minimum-quality galaxy spectra used on our redshift analysis ($Q = 3$; 27 total) have a mean S/N of 4.8 with a standard deviation of 1.4. For emission-line galaxies it is possible to have a high-quality ranking yet low-S/N estimate, since the S/N estimate is dominated by the continuum of a spectroscopic trace (for example, in our data set the minimum S/N of all $Q = 4$ galaxies is 1.2). An example of one of the reduced spectra is shown in Figure 2, and more are shown in a related CIZA J2242.8+5301 galaxy evolution paper (Sobral et al. 2015).

3.2. WHT AF2 Observations

We also conducted a separate spectroscopic survey using WHT/AF2. This survey primarily targeted Hα cluster member candidates identified in our narrowband survey of the system (Stroe et al. 2014c, 2014b). In total 73 objects were targeted over an area roughly $30' \times 30'$. Specific details regarding the target selection and data reduction processes are presented in Sobral et al. (2015).

3.3. Spectroscopic Redshift Catalog

3.3.1. DEIMOS Spectroscopic Redshifts

We obtained 505 spectra with DEIMOS. Of these, we were able to obtain reliable redshifts for 447 objects (89%; see Table 1), leaving 58 spectra that were either too noisy or had ambiguous redshift solutions (e.g., those with a single emission line). Figure 3 shows the redshift distribution of the 255 (51%) high-quality ($Q \geqslant 3$; see Newman et al. 2013 for an explanation on the quality codes) DEIMOS galaxy spectra. Of the high-quality spectra, 206 (41%) fall within $0.176 \leqslant z \leqslant 0.2$, which is $z_{\rm cluster} \pm 3 \times \sigma$, where $z_{\rm cluster} = 0.188$ and σ is the approximate velocity dispersion (1000 km s^{-1}; see Section 6.2). Of the high-quality spectra, 15 (34), or 3% (7%), are

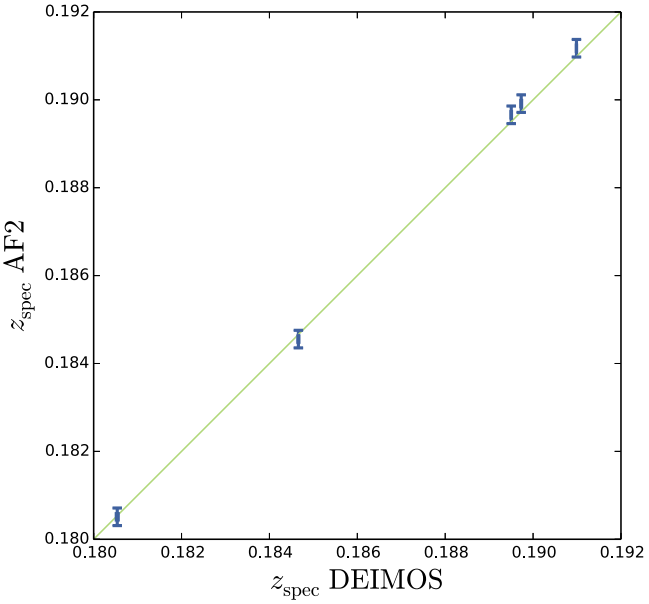

Figure 4. Spectroscopic redshift comparison of the five galaxies in our CIZA J2242.8+5301 survey that have overlapping DEIMOS and AF2 spectra and pass both DEIMOS and AF2 quality cuts (blue error bars). The green line shows the expected 1:1 ratio.

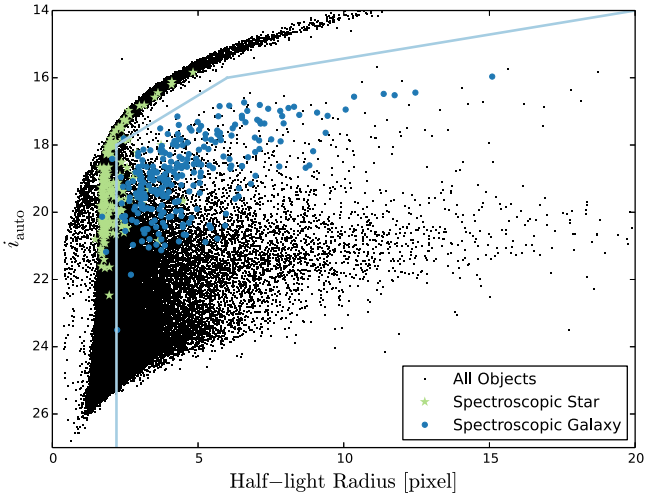

Figure 6. Size–magnitude diagram based on Subaru extinction-corrected i-band magnitude and half-light radius. Spectroscopically confirmed stars (green stars) and galaxies (blue circles) are overlaid. The stellar track is visible to the left and above the light blue lines, which designate our defined star/galaxy separation border. For $i > 18$ stars are defined to have half-light radii <2.2 pixels, at $i = 18$ the slope changes to −0.53, and at $i = 16$ it changes to −0.14 in order to track the changing stellar sequence slope due to saturation. A half light radius of 2.2 pixels is 0.″44 for SuprimeCam. Several spectroscopic stars have half-light radii greater than 2.2 pixels due to blending with neighboring objects.

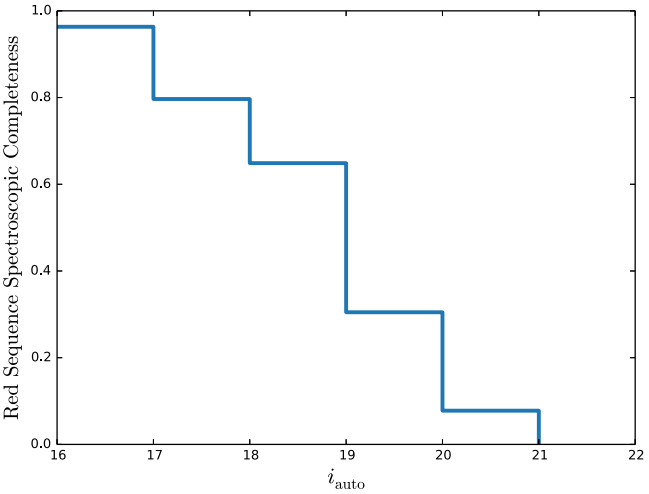

Figure 5. Estimate of the Keck DEIMOS spectroscopic completeness of the cluster red sequence galaxies as a function of extinction-corrected Subaru i-band magnitude in the Keck DEIMOS survey footprint area.

foreground (background) galaxies and 186 (37%) of the spectra are stars. Of these, 72 were serendipitous spectra, meaning that they were not the primary spectroscopic target. Many of these were paired with primary spectra that were also stars, owing to the fact that often binary stars appeared as a single elliptical in the 1.″5–2″ seeing INT/WFC images.

3.3.2. AF2 Spectroscopic Redshifts

We targeted 73 objects with our AF2 spectroscopic survey of cluster Hα candidates, 42 of which fall within the 15′ radius of the cluster center being analyzed for this article. Of those 42 objects, 11 (26%) are stars and 19 (45%) are high-quality galaxy spectra with reliable redshifts. Five of those galaxies also have high-quality DEIMOS spectra (Section 3.3.1). As can be seen from Figure 4, we find that the spectroscopic redshifts for the two surveys are consistent within the measurement errors. In the following analysis we use just the DEIMOS redshift values for these galaxies due to their smaller uncertainties. Of the 14 unique high-quality AF2 redshifts within a 15′ radius of the cluster center, 11 are new cluster members and 3 are higher-redshift galaxies.

4. GALAXY CLUSTER MEMBER SELECTION

To determine which galaxies are members of the CIZA J2242.8+5301 cluster, we utilize both spectroscopic and red sequence cluster member selection methods. The spectroscopic sample has the advantage of being a purer sample, and the precise redshifts are a necessity for many of the following analyses (see Sections 5.1 and 6). While the red sequence sample is not as pure, it is more complete and is not subject to the undersampling bias that affects the spectroscopic sample (Section 4.1); thus, it is advantageous for some analyses (see Section 5.2). In this section we define each sample and quantify the sample completeness and purity.

In what follows we limit our consideration to galaxies within 15′ (2.8 Mpc at $z = 0.188$ or $\sim 1.4\,R_{200}$ for an individual subcluster; Jee et al. 2014c) of the center of the Subaru field, R. A. $= 22^{\rm h}42^{\rm m}43\overset{\rm s}{.}762$, decl. $= 53^{\circ}02'06\overset{''}{.}3$. This radius corresponds to $\sim 1.4\,R_{200}$. Subaru SuprimeCam is strongly vignetted beyond this radius, with the corner pixels receiving approximately half the light as the center (von der Linden et al. 2014). Since the cluster fits well within this radius, there would be little gained by including galaxies outside this radius.

4.1. Spectroscopic Redshift Selection

All spectroscopic galaxies within the range $0.176 \leqslant z \leqslant 0.2$ are considered to be cluster members. This range is defined by $z_{\rm cluster} \pm 3 \times \sigma$, where $z_{\rm cluster} = 0.188$ and σ is the approximate velocity dispersion of each subcluster (1000 km s^{-1}; see

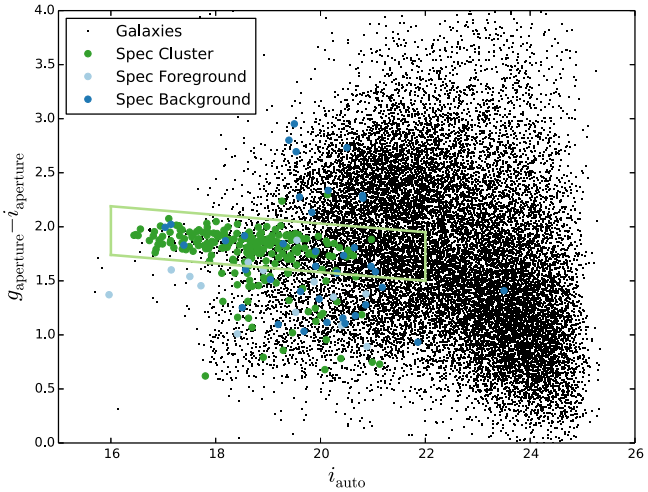

Figure 7. Color–magnitude diagram of galaxies within a 15′ radius of the system center, based on extinction-corrected Subaru g and i magnitudes. Spectroscopic cluster (green), foreground (light blue), and background (dark blue) galaxies are overlaid. Our red sequence selection region is outlined in light green.

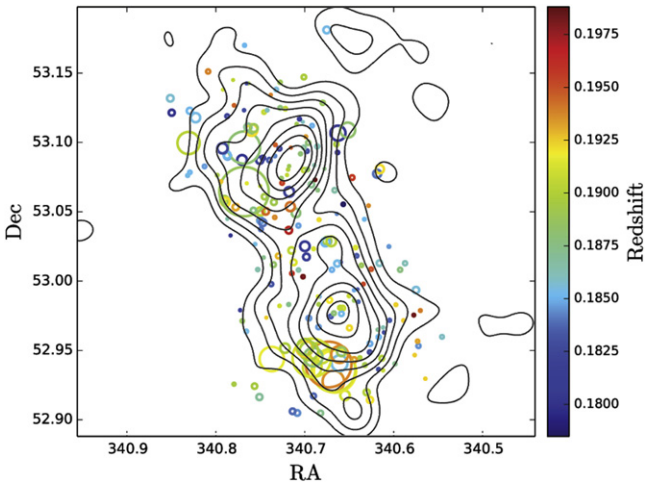

Figure 8. Projected locations of Keck DEIMOS spectroscopic cluster members color coded according to their redshift, with red sequence sample number density contours (see Figure 10 for detailed description). The diameter of each circle is proportional to 10^{δ}, where δ is the DS-δ value for each galaxy: the larger the circle, the more likely that galaxy belongs to a substructure with disparate velocity and/or velocity dispersion from that of the bulk system properties. There are 15 spectroscopic galaxies in the south that show signs of constituting a substructure (i.e., clustering of large circles) with $z \sim 0.191$. Note that the DS-test is not expected to identify the larger north and south subclusters because they have nearly identical velocity distributions.

Section 6.2). This is not exactly a 3σ selection cut, since the velocity dispersions of the northern and southern subclusters are 1160 ± 95 and 1080 ± 90 km s^{-1}, respectively, and they have a line-of-sight (LOS) relative velocity difference of -69 ± 190 km s^{-1} (see Section 6.2). This selection results in 206 Keck DEIMOS spectroscopic redshifts and 11 unique WHT AF2 spectroscopic redshifts, for a total of 217 spectroscopic cluster member redshifts.

Since our Keck DEIMOS spectroscopic survey targeted primarily cluster red sequence galaxies (see Section 3.1.1), it is an incomplete survey of the cluster blue cloud galaxies. While the WHT AF2 survey adds a number of blue-cloud galaxies, there are only 11 unique additional spectra in the 15′ radius surrounding the cluster. Also, since the blue cloud region of color–magnitude space has a large amount of stellar contamination, it is difficult to estimate our completeness of this population of cluster galaxies. However, we are able to use the cluster red sequence to estimate our spectroscopic completeness in this region of color–magnitude space. After correcting for the purity of our red sequence imaging selection (Section 4.2.2) and Keck DEIMOS survey area, we estimate the spectroscopic completeness for red sequence cluster galaxies as a function of i-band magnitude (Figure 5).

While our Keck DEIMOS spectroscopic survey is a $\gtrsim$70% complete sample of cluster red sequence galaxies with $i < 19$ (mass $\gtrsim 10^{10}$ $M_{\odot}$), it is important to note the undersampling bias that affects the densest parts of the subclusters. Since Keck DEIMOS utilizes slit masks and the reduction software (Section 3.1.2) is not designed for slits that overlap in the dispersion direction, we undersample the dense cores of each subcluster (see, e.g., the insets of Figure 10) relative to the less dense periphery of each subcluster. This bias will affect the southern subcluster more than the northern subcluster, due to its higher galaxy concentration.

4.2. Red Sequence Selection

Despite our spectroscopic survey being a $\gtrsim$70% complete sample of cluster red sequence galaxies with $i < 19$, we are able to obtain a more complete and less biased survey of cluster members through color–magnitude selection. In this subsection we first discuss our star/galaxy separation schema and then discuss our red sequence cluster membership selection schema, as well as the purity of this sample.

4.2.1. Star–Galaxy Separation

The excellent 0.″65 seeing of the Subaru i-band imaging facilitates star–galaxy separation via size (or half-light radius) cuts. We couple this with each objects' magnitude to perform a size–magnitude cut to distinguish between stars and galaxies (see Figure 6). In Figure 6 we overlay spectroscopically confirmed stars and galaxies, as well as our defined border between the star–galaxy phase space. For $i > 18$ stars are defined to have half-light radii <2.2 pixels (0.″44), at $i = 18$ the slope changes to −0.53, and at $i = 16$ it changes to −0.14 in order to track the changing stellar sequence slope due to saturation. Our star–galaxy separation schema errs more toward galaxy completeness than purity, since blending results in a large number of stars with measured half-light radii >2.2 pixels.

We also investigated whether a color–magnitude cut or PSF-based cut (Jee et al. 2014c) would increase our star–galaxy discriminating power. All reasonable color–magnitude cuts and PSF-based cuts resulted in a sample of stars that were already subsumed by the size–magnitude selected sample.

4.2.2. Red Sequence Properties

We find that after star–galaxy separation and dust extinction corrections there is a well-defined and relatively tight cluster red sequence (see Figure 7; see Appendix A for discussion of the dust extinction corrections). We further accentuate this by plotting spectroscopically confirmed cluster members in this color–magnitude space (green points in Figure 7). Across the 15′ radius field there are 2605 presumed galaxies that fall within our defined red sequence region. We estimate the purity

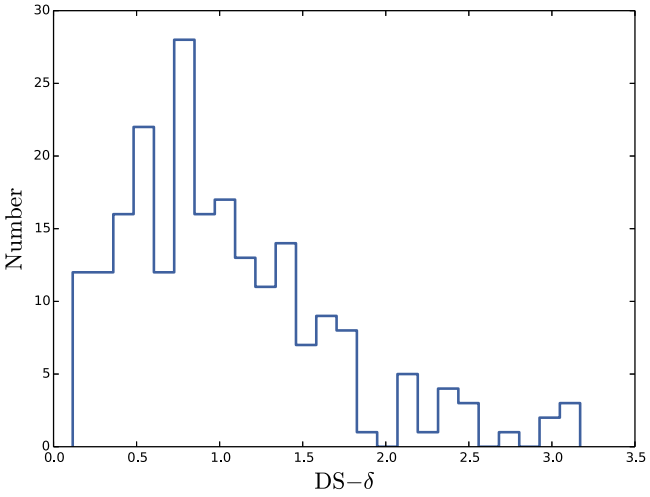

Figure 9. DS-δ distribution for Keck DEIMOS spectroscopic cluster members. All but three of the galaxies with $\delta > 2.0$ are compactly clustered in the south. We define the galaxies to make up the *interloper*.

of a cluster red sequence selected sample by studying the population of spectroscopic stars and galaxies within our red sequence selection region. Our red sequence selection region extends to $i = 22$ and our spectroscopic sample completeness falls to $<10\%$ beyond $i > 20$ (see Figure 5), so our purity calculations should be considered rough estimates. Within the red sequence selection region there are 234 spectroscopic objects with secure redshifts: 179 (77%) are cluster galaxy members, 4 (2%) are foreground galaxies, 14 (6%) are background galaxies, and 38 (16%) are stars. We also find no evidence for clustering of the contaminants and thus expect no resulting propagation of bias in our subcluster location estimates. We do not attempt to estimate the completeness of this cluster red sequence membership selection schema since our spectroscopic survey was not a magnitude-limited survey, instead targeting primarily red sequence galaxies.

5. SUBCLUSTER IDENTIFICATION

No one subcluster identification method is optimal for all types of subcluster configurations (Pinkney et al. 1996; Einasto et al. 2012), so we employ three separate complementary methods of subcluster identification based on galaxy cluster membership discussed in Section 4. The first is the Dressler & Shectman (1988) analysis, the second is a projected galaxy number/luminosity overdensity analysis, and the third is a Gaussian Mixture Model (GMM) clustering analysis. The DS-test has been shown to be one of the best at identifying substructure in clusters (Pinkney et al. 1996; Einasto et al. 2012); however, it has a notable weakness when attempting to identify substructures with very similar redshifts and velocity dispersions (e.g., subclusters of similar mass that are merging near the plane of the sky). While the projected galaxy number/luminosity overdensity method largely ignores redshift information (except in broad cluster membership selection), it is best at identifying substructure with large relative projected separations (e.g., subclusters that are merging near the plane of the sky; Pinkney et al. 1996). The GMM analysis, coupled with a Bayesian information criterion (BIC) analysis, enables an objective determination of the optimal number of subclusters and their member galaxies. We perform the GMM analysis on the one-dimensional redshift distributions, as well as the three-dimensional galaxy distribution (right ascension, declination, and redshift).

5.1. Dressler–Shectman Test

In an attempt to identify the main subclusters and minor substructures in the system, we perform a DS-test (Dressler & Shectman 1988) analysis where we calculate the DS-δ value of each spectroscopic cluster member (see Section 4.1). For each galaxy the DS-δ parameter is calculated as

$$\delta^2 = \frac{N_{\rm local}}{\sigma^2}\Big[(\bar{v}_{\rm local} - \bar{v})^2 + (\sigma_{\rm local} - \sigma)^2\Big], \quad (1)$$

where $N_{\rm local}$ is the number of nearest neighbors (including the galaxy itself) to include when calculating $\bar{v}_{\rm local}$, the average LOS velocity, and $\sigma_{\rm local}$, the local velocity dispersion. We let $N_{\rm local} = \lceil\sqrt{N_{\rm total}}\rceil$, where $N_{\rm total}$ is the total number of spectroscopic cluster members, following the best practice identified by Pinkney et al. (1996). Cluster substructures will have larger δ values.

In Figure 8 we plot the projected location of each spectroscopic cluster member and represent it as a circle with diameter proportional to 10^{δ}. We find that there is a concentration of galaxies with large δ values in the south, indicative of cluster substructure with a mean LOS velocity and/or velocity dispersion different from that of the system average. Looking at the distribution of δ values (Figure 9), we find that there is an apparent break near $\delta = 2.0$. Fourteen of the galaxies with $\delta > 2.0$ are compactly clustered in the south. These make up a small fraction of the total number of spectroscopic galaxies (206) and are considerably fainter than the more massive galaxies that define the southern subcluster peak. This leads us to define the galaxies as members of a substructure we call *Interloper*.

We investigate the significance of the interloper substructure by comparing the cumulative deviation,

$$\Delta = \sum_{i}^{N_{\rm total}} \delta_i, \quad (2)$$

of the observed system with that of 10,000 realizations where we maintain the projected galaxy locations but shuffle the redshifts. When we do this for all of the cluster redshifts, we find $\Delta = 221.6$, which is only a 0.55σ deviation from the mean of the distribution defined by the 10,000 resamplings. If we instead investigate the significance of the interloper by considering only redshifts within 625 kpc of the peak of the southern subcluster (Section 6.1), we find $\Delta = 93.0$, which is a 1.8σ deviation. Since there is only marginal evidence for the interloper being a distinct substructure, in Section 6, we consider both the cases where the interloper is a distinct substructure and where the interloper galaxies are just members of the southern subcluster.

With the exception of the interloper galaxies, the redshift distributions in the north and south regions of the system are similar (see Figure 8). The DS-test, designed primarily to identify velocity substructure, cannot separate structures with such similar radial velocity distributions. Thus, the results are not inconsistent with the previous findings of van Weeren et al. (2011), suggesting that the CIZA J2242.8+5301 system consists of two nearly equal mass subclusters.

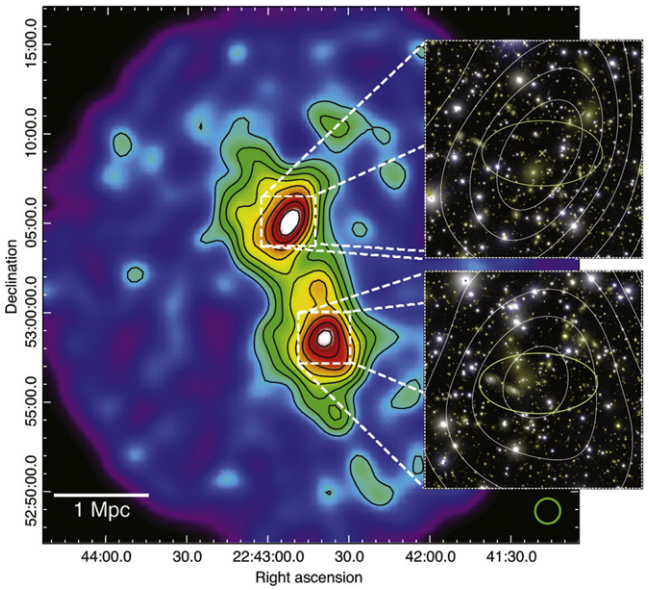

Figure 10. Smoothed galaxy luminosity density map of CIZA J2242.8+5301 based on cluster red sequence selection. The cluster galaxy number density contours (black) based on our red sequence selection begin at 100 galaxies Mpc^{-2} and increase linearly with increments of 25 galaxies Mpc^{-2}. Color composites based on the Subaru *g*- and *i*-band observations are shown for the peaks of the north and south subclusters. The light green ellipses show the 68% confidence regions for the locations of each subcluster based on 10,000 bootstrap resamplings of the cluster red sequence galaxies. The dark green circle in the bottom right of the map shows the scale of the KDE bandwidth used to create the map.

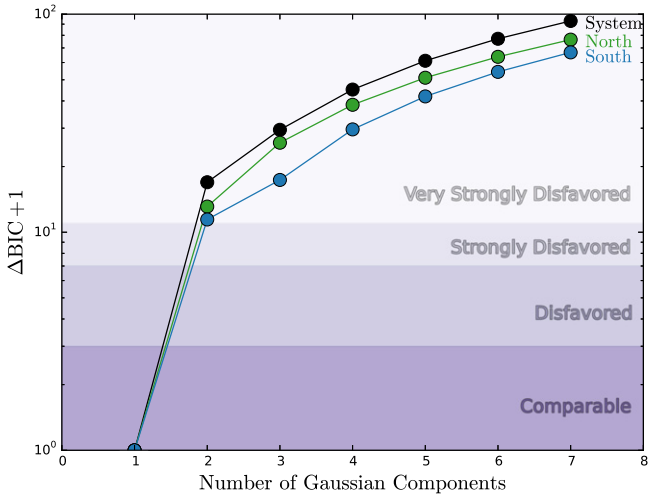

Figure 11. ΔBIC plot comparing GMM fits to the redshift distributions of the entire cluster system (black), northern subcluster (green), and southern subcluster (blue), with varying number of Gaussian components. Note that the *diag* and *full* covariance structures are indistinguishable for one-dimensional data. The purple shaded regions roughly denote how a given model compares with the model that has the lowest BIC score. All distributions are best fit by a single-component model.

5.2. Projected Densities

Given the similar redshift distributions (see Figures 8 and 14) and velocity dispersions in the north and south (as we will discuss in Section 6) and the known failure mechanism of the DS-test, we also look for substructures in projected space. We independently use both the spectroscopic cluster member sample, Section 4.1, and the red sequence cluster member sample, Section 4.2. For each of these samples we study both the projected galaxy number density distribution and the projected luminosity density distribution (essentially the same as the number density except that we weight each galaxy by its observed *i*-band luminosity, assuming that it is at the average redshift of the cluster).

We employed a kernel density estimation (KDE)[11] to model the structures of the galaxy data. We made use of KDE as the number density estimate,

$$\hat{f}(\boldsymbol{x}, \boldsymbol{h}) = \frac{1}{n \prod_{j=1}^{p} h_j} \sum_{i=1}^{n} \left[\prod_{j=1}^{p} K\left(\frac{\boldsymbol{x_i} - \boldsymbol{X_{ij}}}{\boldsymbol{h_j}} \right) \right], \quad (3)$$

where we use $p = 2$ as the number of spatial variables, n as the number of galaxies, $\boldsymbol{X_i} = (X_{i1}, X_{i2})$ as the two spatial values of each galaxy, $\boldsymbol{h} = (h_1, h_2)$ as the bandwidth for each dimension, and K as the bivariate Gaussian kernel function.

The most important aspect of performing a KDE is to pick suitable bandwidths $\boldsymbol{h}$ (smoothing length). The smaller the bandwidth, the greater the variance in the KDE; however, the greater the bandwidth the greater the bias. The consideration for choosing suitable bandwidths is what is known as the bias-variance trade-off. We picked our smoothing bandwidth by performing an exhaustive leave-one-out cross-validation (Stone 1984) in each dimension, while maximizing the likelihood of fit between our KDE and the data. The cross-validated score in each dimension (l) can be written as

$$CV_l(h_i) = \frac{1}{N} \sum_{i}^{n} \ln \hat{f}_{-i,\mathrm{kern}}(X_i), \quad (4)$$

where we constructed n data sets, with data of the ith galaxy being left out in each data set, and we performed a grid search of suitable h_i values to maximize the score. When we apply this procedure to the red sequence selected sample, we find that the most suitable bandwidths (i.e., those with the maximum *CV* score) for the R.A. and decl. are 62″ and 42″, respectively, and when we apply it to the spectroscopic cluster member sample, we find 90″ and 67″ for the R.A. and decl. dimensions, respectively. To avoid anamorphic distortions in the projected R.A.–decl. space, we use the smaller of the two bandwidths for each dimension, 42″ for the red sequence sample and 67″ for the spectroscopic cluster member sample. We choose the smaller of the two bandwidths in each case since this minimizes bias. While this choice will slightly increase the variance, we have verified that we are still able to maintain subcluster peak density S/Ns >9 by performing bootstrap error analyses of each map with 1000 resamplings of the respective galaxy populations. We find general agreement between each of the four resulting density maps. For the sake of simplicity, in what follows we will consider just the red sequence number density map; however, we present the four resulting density maps in Appendix B.

From the red sequence number density map presented in Figure 10 it is apparent there are two distinct subclusters (one in the north and one in the south) with similar size and density. We compare this galaxy density distribution with the X-ray and

[11] A more comprehensive discussion of KDE can be found in either Feigelson & Babu (2012) or Ivezić et al. (2014).

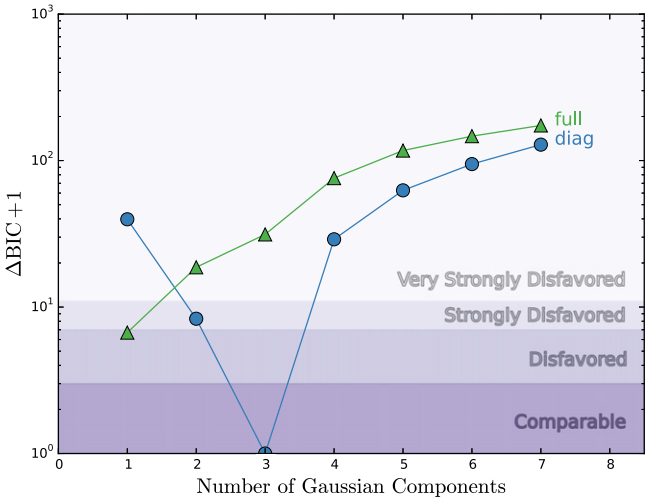

Figure 12. ΔBIC plot comparing GMM fits to the three-dimensional (right ascension, declination, and redshift) distribution of all the cluster member spectroscopic galaxies (see Section 4), with varying number of Gaussian components and covariance type. We plot the results for models with *diag* (blue triangles) and *full* (green squares) covariance types. The purple shaded regions roughly denote how a given model compares with the model that has the lowest BIC score. The best fit is a three-component model with *diag* covariance structure.

radio emission of the system in Figure 1. The two dominant subclusters that are aligned closely with the merger axis were inferred from the radio relics (van Weeren et al. 2010) and elongated X-ray gas distribution (Ogrean et al. 2013, 2014). Furthermore, the X-ray gas distribution is largely located between the two galaxy subclusters, as expected for a dissociative merger. We discuss the galaxy distribution in relation to the other cluster emission further in Section 7.1, and Jee et al. (2014c) discuss the galaxy distribution in relation the the weak-lensing mass distribution.

5.3. Gaussian Mixture Model

We implement a modified version[12] of *scikit-learn*'s (Pedregosa et al. 2011) GMM program and apply it to the one-dimensional redshift distributions, as well as the three-dimensional galaxy distribution (right ascension, declination, and redshift). We consider mixtures of one to seven multivariate Gaussian components and two types of covariance structures: *diag*, each Gaussian component has an uncorrelated covariance structure; and *full*, each Gaussian component can have a different unstructured covariance. Note that we do not consider a *tied* covariance type (where each Gaussian component has the same unstructured covariance) since there is no reason to expect that merging subclusters should have the same size and velocity dispersion; nor do we consider a *spherical* covariance type (where each Gaussian component has an equicorrelation covariance structure) since there is no reason to expect the scale of the cluster in projection to be tied to its velocity dispersion scale. For each number of components (n) and covariance structure (c) we calculate the BIC and use this to infer the optimal number of subclusters. We plot these results as

$$\Delta\mathrm{BIC}_{nc} = \mathrm{BIC}_{nc} - \min\left(\mathrm{BIC}_{nc} | n \in \mathbb{Z}_{1\ldots7}, c \in \mathbb{C}\right), \quad (5)$$

where $\mathbb{Z}_{1\ldots7}$ is the set of integers from 1 to 7, and $\mathbb{C}$ is the set of covariance structures {diag, full}. For convenience of interpretation we color-code regions of the ΔBIC plot according to the broad model comparison categories suggested by Kass & Raftery (1995).

5.3.1. One-dimensional GMM

To test for LOS substructure, we apply our GMM analysis to the redshift distribution for all cluster member spectroscopic galaxies (see Section 4), as well as the distribution of cluster member galaxies within a 625 kpc radius of the respective north and south red sequence number density locations (see Section 6.1). These apertures were chosen to be as large as possible while maintaining mutual exclusivity of the subcluster membership. In total we use 69 and 76 redshifts when analyzing the northern and southern subclusters, respectively. For each redshift distribution we find that it is optimally fit by a single Gaussian and that multiple Gaussian component fits are strongly disfavored (see Figure 11). Thus, there is no significant evidence of substructure in the LOS dimension, suggesting that the two major subclusters identified in Section 5.2 have a relative LOS velocity difference less than their respective velocity dispersions.

5.3.2. Three-dimensional GMM

We also apply our GMM analysis to the three-dimensional (right ascension, declination, and redshift) distribution of all the cluster member spectroscopic galaxies (see Section 4). We find that the data are best fit by a three-component Gaussian model with *diag* covariance structure (see Figure 12). In Figure 13 we plot the three-dimensional distribution of the spectroscopic cluster members and their most likely subcluster membership assignment for this best-fit model. For the projected one-dimensional distributions we plot the marginalized Gaussian components for the best-fit model (dashed lines). For the projected two-dimensional distributions we plot marginalized 68% confidence ellipses of the best-fit model Gaussian components. While a three-component model is preferred, the majority of the galaxies belong to two subcluster components (blue diamonds and green circles in Figure 13), corresponding to the ones we identified in the projected density analysis of Section 5.2. The third component (black triangles in Figure 13) consists almost entirely of AF2 spectroscopic galaxies and is an artifact of the AF2 Hα survey sparsely sampling a larger footprint relative to the DEIMOS survey. If we exclude the AF2 spectroscopic redshifts, we find a two-component optimal model fit, which is essentially just the north (green) and south (blue) components of Figure 13 (note that the GMM galaxy membership assignment of the north and south subclusters is similar to what we obtained with the projected apertures defined in Section 5.3.1). Thus, our three-dimensional GMM analysis confirms the results of Section 5.2 but finds no significant evidence for the potential interloper we identified in Section 5.1.

[12] https://github.com/wadawson/scikit-learn/commit/ea033dcc3c04957bad7f7737c6800b657ed29454.

6. SUBCLUSTER PROPERTIES

Having defined the northern, southern, and interloper subclusters in Section 5, we now present the macroscopic galaxy properties of each subcluster. Of particular interest are the subcluster locations, redshifts, and velocity dispersions (σ_v).

6.1. Subcluster Locations

Accurate subcluster locations are necessary for the dynamic analysis of the system. They are also necessary for some constraints on the DM self-interactions that rely on accurately constraining the offset between the effectively collisionless galaxies and DM. The dynamic analysis also depends on accurate estimates of each subcluster's redshift, or their relative LOS velocities.

To estimate the subcluster locations, we use the four KDE projected density maps discussed in Section 5.2. We measure the locations of the north and south subclusters as the peaks of density maps in the north and south regions, respectively. To estimate the uncertainty distribution on these peak locations, we generate 10,000 bootstrap samples from the respective cluster member sample and repeat the same smoothing and peak location process for each, limiting the search region to ~500 kpc × 500 kpc regions surrounding each peak in the original density map. We find consistent location estimates in each of the four maps for both the north and south subclusters (see Figure A2). We report here location estimates for the red sequence cluster member sample (Section 4.2) and KDE projected number density map (Section 5.2), since it is less affected by the spectroscopic undersampling bias (see discussion in Section 4.1) and since our bootstrap resampling for the luminosity-weighted maps is potentially biased due to resampling the galaxies rather than units of luminosity. We find that the north subcluster is located at (R.A. = $22^{\rm h}42^{\rm m}50^{\rm s}{}^{+50^{\rm s}}_{-50^{\rm s}}$, decl. = $53^\circ05'06''{}^{+32''}_{-23''}$) and the south subcluster is located at (R.A. = $22^{\rm h}42^{\rm m}39^{\rm s}{}^{+50^{\rm s}}_{-50^{\rm s}}$, decl. = $52^\circ58'35''{}^{+31''}_{-18''}$). These

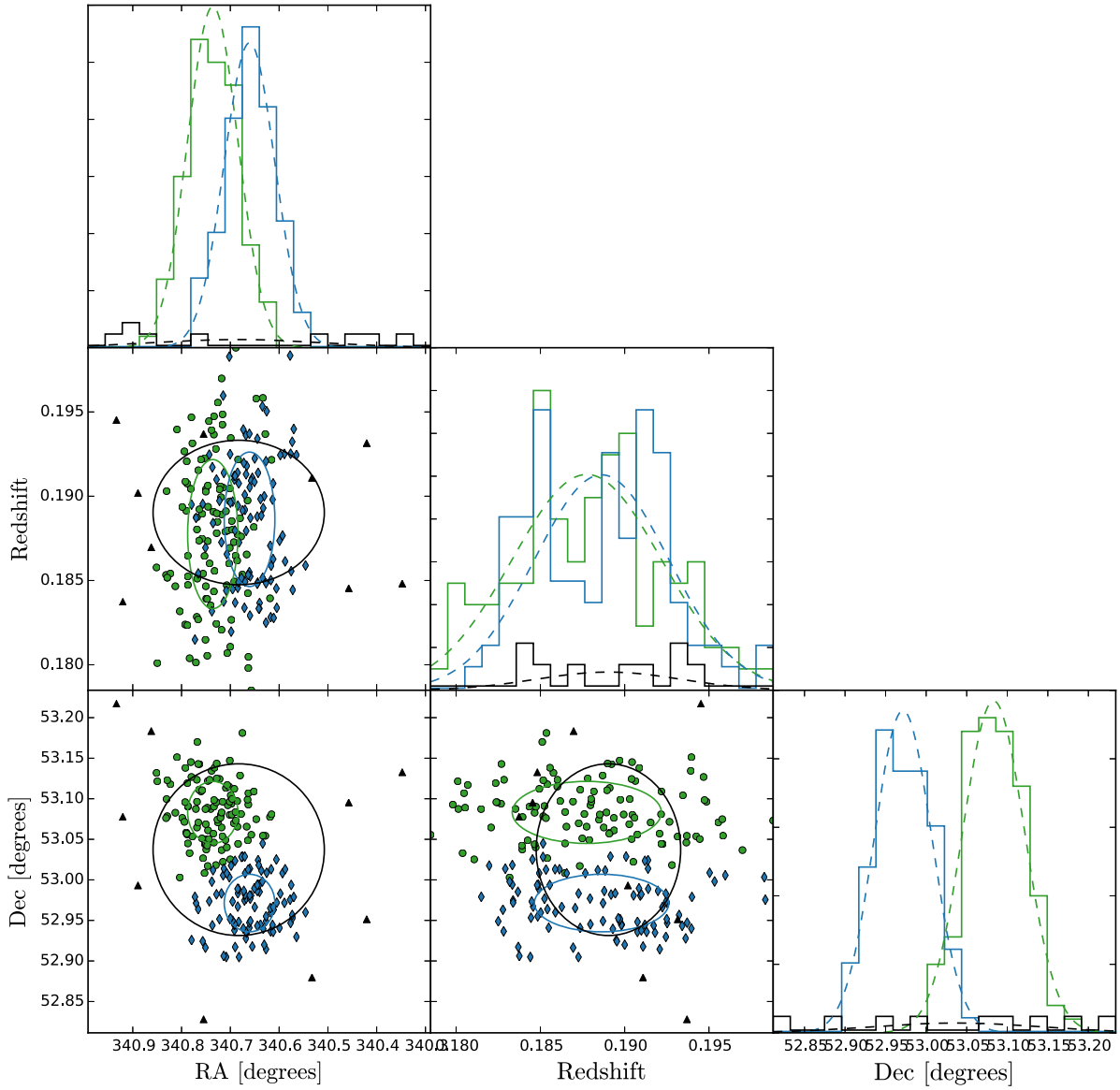

Figure 13. Three-dimensional distribution of the spectroscopic cluster members (Section 4) and their most likely subcluster membership assignment for the best-fit GMM (see Figure 12). For the projected one-dimensional distributions we plot the marginalized Gaussian components for the best-fit model (dashed lines). For the projected two-dimensional distributions we plot projected ellipses that encompass ~68% of the corresponding members in the best-fit model Gaussian components.

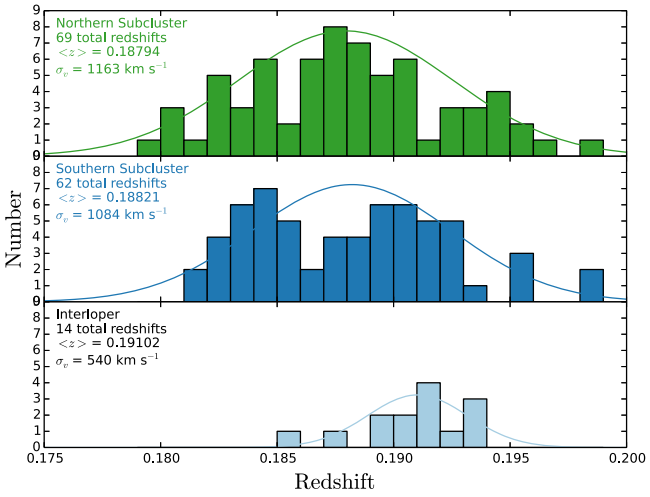

Figure 14. Redshift distributions of the northern subcluster (green), southern subcluster (dark blue), and the potential interloper (light blue). Redshift locations and velocity dispersions are listed in the upper left of each subpanel. The northern and southern subcluster histograms include spectroscopic members within a 625 kpc radius of the peak location of each subcluster (Section 6.1). Interloper galaxies were excluded from the southern subcluster distribution.

locations, as well as the 68% confidence regions, are shown in Figure 10.

6.1.1. Location Comparisons

Given the north and south subcluster locations, we estimate that the projected separation of the two subclusters is $6\farcm9^{+0.7}_{-0.5}$, which corresponds to $1.3\,\mathrm{Mpc}^{+0.13}_{-0.10}$. We estimate the projected separation PDF by selecting 10,000 random samples from the aforementioned north and south subcluster location bootstrap samples and calculate the spherical trigonometric separation of the two in each case.

As can be seen from the north and south zoomed insets of Figure 10, the BCGs of the north and south subclusters are near the subcluster red sequence number density peak locations, 55 and 85 kpc, respectively, and within the 68% confidence uncertainty of each (~90 kpc). In the case of the northern subcluster the BCG is the closest central galaxy (for galaxies with $i < 20$). In the case of the southern subcluster the second-brightest subcluster galaxy is the closest central galaxy (offset 43 kpc from the redshift number density peak; for galaxies with $i < 20$), although both it and the southern BCG are consistent with being closest to the various subcluster peak locations. In Section 6.2 we compare the redshifts of the subclusters and their respective BCGs, and in Section 7 we attempt to interpret these projected and redshift offsets.

We find that the northern and southern subclusters of galaxies are trailing their respective radio relics by $4\farcm7 \pm 0\farcm6$ and $2\farcm4 \pm 0\farcm6$ (0.85 ± 0.11 and 0.45 ± 0.11 Mpc at $z = 0.188$), respectively. Given that the plasma shock waves, suspected of sourcing the radio relics, are gravitationally decoupled from the system, it is expected that they will lead the subclusters, which are gravitationally coupled. Thus, the subcluster-relic offset is expected to increase with time. In a follow-up dynamics analysis of CIZA J2242.8+5301 (W. A. Dawson et al. 2015, in preparation) we will use the offsets as a prior to the Dawson (2013a) method to constrain the dynamic properties of the merger, in a similar manner to the Ng et al. (2014) dynamics analysis of the El Gordo merger.

The north and south subclusters are leading the peak of the smoothed *XMM-Newton* X-ray luminosity (R. A. $= 22^{\mathrm{h}}42^{\mathrm{m}}43\fs8$, decl. $= 53°00'55''$; Gaussian smoothing kernel with $\sigma = 0\farcm5n$) by $4\farcm5 \pm 0\farcm7$ and $2\farcm4 \pm 0\farcm5$, respectively. These offsets correspond to 0.89 ± 0.13 Mpc and 0.45 ± 0.09 Mpc, respectively. These offsets are comparable to those observed in other dissociative mergers (see, e.g., Bradač et al. 2006, 2008; Mahdavi et al. 2007; Dawson et al. 2012). Note that the respective X-ray peak-subcluster offsets and radio relic-subcluster offsets of the north and south subclusters are nearly identical. We will further explore this finding in a follow-up dynamics analysis of CIZA J2242.8+5301 (W. A. Dawson et al. 2015, in preparation).

We perform a detailed galaxy versus weak-lensing location comparison in Jee et al. (2014c), where we find $\sim 1'$ offsets between the galaxy–weak-lensing locations, although we find that these offsets are not highly significant given the measurement uncertainties. The combined probability that the galaxies and mass have different locations is 83% (combining the individual offset p-values in Jee et al. 2014c, with Fischer's formalism).

6.2. Subcluster Redshifts and Velocity Dispersions

To investigate the redshift and velocity dispersions of each subcluster, we consider all spectroscopic cluster member galaxies within a 625 kpc radius of the respective red sequence number density location (see Section 6.1). These apertures were chosen to be as large as possible while maintaining mutual exclusivity of the subcluster membership. In total we use 69 and 76 redshifts when analyzing the northern and southern subclusters, respectively. We also consider the possibility that the interloper is an independent structure and exclude the 14 associated galaxies (Section 5.1) from the southern subcluster membership and estimate their redshift and velocity dispersion separately. The redshift distributions of each of these selections are shown in Figure 14. While the southern subcluster redshift distribution appears bimodal, there is no sign of corresponding clustering in projected space, as discussed in Section 5.1.

We estimate each subcluster's redshift and velocity dispersion using the biweight-statistic and bias-corrected 68% confidence limit (Beers et al. 1990) applied to 100,000 bootstrap samples of each subcluster's spectroscopic redshifts. We summarize these results in Table 2. We find very similar redshifts for the northern and southern subclusters, $0.18794^{+0.00054}_{-0.00054}$ and $0.18900^{+0.00050}_{-0.00049}$, respectively. These translate to a relative LOS velocity difference in the frame of the cluster of $v_{\mathrm{north}} - v_{\mathrm{south}} = -73 \pm 190\ \mathrm{km\ s^{-1}}$. This suggests that either they are both nearly in the plane of the sky, they have slowed as they near the apocenter of the merger, or a combination of the two. Van Weeren et al. (2010) argue that the merger is occurring close to the plane of the sky. As we will show in a more detailed dynamics analysis (W. A. Dawson et al. 2015, in preparation), it is likely a combination of the two effects. Comparing the relative redshift of each subcluster's BCG with respect to the median subcluster redshift, we find relative LOS velocity differences of $v_{\mathrm{north}} - v_{\mathrm{north\ BCG}} = -500 \pm 140$ and $v_{\mathrm{south}} - v_{\mathrm{south\ BCG}} = -240 \pm 130\ \mathrm{km\ s^{-1}}$. In Section 7 we attempt to interpret these redshift offsets in conjunction with the projected offsets (see Section 6.1.1).

Table 2
Observed Subcluster Properties

Subcluster	R.A.[a]	Decl.[a]	Redshift	σ_v (km s^{-1})	σ_v M_{200} (10^{14} $M_\odot$)	WL M_{200}[b] (10^{14} $M_\odot$)
North	$22^h42^m50^s$	53°05′06″	$0.18794^{+0.00054}_{-0.00054}$	1160^{+100}_{-90}	$16.1^{+4.6}_{-3.3}$	$11.0^{+3.7}_{-3.2}$
South with Interloper	$22^h42^m39^s$	52°58′35″	$0.18900^{+0.00050}_{-0.00049}$	1120^{+100}_{-80}	$14.5^{+4.1}_{-2.8}$	$9.8^{+3.8}_{-2.5}$
South without Interloper	...	...	$0.18821^{+0.00054}_{-0.00052}$	1080^{+100}_{-70}	$13.0^{+4.0}_{-2.5}$	...
Interloper	$22^h42^m43^s$	52°56′38″	$0.19102^{+0.00055}_{-0.00065}$	540^{+190}_{-110}	$1.6^{+2.4}_{-0.8}$	...

[a] The uncertainties on these locations are discussed in Section 6.1.
[b] As presented in Jee et al. (2014c) weak-lensing analysis of CIZA J2242.8+5301.

While velocity dispersion mass estimates have been shown to be biased measures in disturbed systems (Pinkney et al. 1996), we estimate them here since they provide an independent mass estimate to compare with the less systematic prone weak-lensing mass estimates (Jee et al. 2014c). Furthermore, given the relatively large offset of the subclusters and their relatively small LOS velocity difference, we expect that the velocity dispersion bias is not significantly larger than the statistical uncertainty. We find similar velocity dispersions for the northern and southern subclusters, 1160^{+100}_{-90} and 1130^{+100}_{-80} km s^{-1}, respectively. We perform a Kolmogorov–Smirnov test to compare the redshift distribution of the north and south subclusters relative to the normal distribution defined by the calculated bi-weight location (redshift) and scale (velocity dispersion) of each subcluster. We find p-values of 0.95 and 0.07 for the north and south subclusters, respectively. Thus, both redshift distributions are consistent with being normally distributed, although there is some tension for the southern subcluster (see, e.g., Figure 14), and its velocity dispersion estimate may be more biased than the northern subcluster's. The velocity dispersion estimates are consistent with the picture of subclusters of similar richness seen in the galaxy density maps (see, e.g., Figure 10). Converting these velocity dispersions into M_{200} mass estimates using the Evrard et al. (2008) scaling relation, we estimate masses of $16.1^{+4.6}_{-3.3} \times 10^{14}$ and $14.5^{+4.1}_{-2.8} \times 10^{14}$ $M_\odot$ for the northern and southern subclusters, respectively. These mass estimates are consistent with the Jee et al. (2014c) weak-lensing mass estimates, shown here in Table 2, although slightly larger.

When we consider the possibility of the interloper being an independent structure, by excluding the 14 associated galaxies (Section 5.1) from the southern subcluster membership, we find that the redshift of the southern subcluster changes slightly to $z = 0.188821^{+0.00054}_{-0.00052}$ and the interloper has a redshift of $z = 0.19102^{+0.00055}_{-0.00065}$. Comparing the relative LOS velocities of the southern subcluster (without the interloper galaxies) and the interloper, we find $v_{\rm south} - v_{\rm interloper} = -710 \pm 200$ km s^{-1}; while larger than the northern and southern subcluster relative LOS velocity, it is still smaller than the velocity dispersion of the southern subcluster (see Figure 14). For the interloper we estimate a velocity dispersion of 540^{+190}_{-110} km s^{-1}, which translates to $M_{200} = 1.6^{+2.4}_{-0.8} \times 10^{14}\,M_\odot$. However, as we discussed in Section 5.1, the evidence for the interloper is not highly significant according to the DS-test. Even if it is a valid substructure, we caution that its velocity dispersion estimate is only based on 14 redshifts and the quoted statistical uncertainties are likely underestimated. The southern velocity dispersion changes by an insignificant amount when the interloper galaxies are excluded (Table 2).

7. DISCUSSION

Our findings largely support the general interpretation that CIZA J2242.8+5301 is a major cluster merger being observed sometime after the first pericentric collision. We have found that CIZA J2242.8+5301 is dominated by two subclusters of similar scale, density, and mass. Furthermore, we are able to accurately locate the subclusters at the leading edges of the elongated X-ray emission, with the bulk of the X-ray-emitting gas located between the two subclusters, making this a textbook dissociative merger. In this section we discuss how our findings add to the multi-wavelength picture of CIZA J2242.8+5301 and place our observed BCG-subcluster offsets in the context of existing works.

7.1. Multi-wavelength Merger Picture

Van Weeren et al. (2010) suggest that the observed radio relic polarization of 50%–60% indicates that the merger angle must be within ~ degrees of the plane of the sky. Our observed relative velocity of the subclusters, $v_{\rm north} - v_{\rm south} = -69 \pm 190$ km s^{-1}, is consistent with a merger occurring nearly in the plane of the sky; however, without further analysis we cannot rule out the possibility that the merger has a larger inclination angle and is just being observed near the merger apocenter where the subclusters have slowed just before or after turnaround. We will address this in a future paper where we present our detailed geometric and dynamic analysis of the system (W. A. Dawson et al. 2015, in preparation).

Van Weeren et al. (2011) conducted simulations of the system and argue that they suggest a bimodal merger with a 2:1 mass ratio between the north and south subclusters and an impact parameter less than 400 kpc. While our velocity-dispersion-based mass estimates prefer closer to a 1:1 mass ratio, our mass estimates are consistent with a 2:1 mass ratio (the same statement can be made of the Jee et al. 2014c weak-lensing mass estimates). It is difficult for us to place tight constraints on the impact parameter; however, we find excellent agreement between the merger axis inferred from the radio relics and elongated gas distribution (see Figure 1), suggesting that the impact parameter is not significantly larger than our uncertainty on the subcluster locations $\lesssim$200 kpc.

We find that the north and south subcluster-gas projected offsets (0.89 ± 0.13 and 0.45 ± 0.09 Mpc, respectively) and subcluster-relic projected offsets (0.85 ± 0.11 and 0.45 ± 0.11 Mpc, respectively) are asymmetric (see Section 6.1.1). The asymmetry between the north and south offsets is suggestive of a non-equal-mass merger and is consistent with a 2:1 mass ratio. As an aside, we note that the magnitudes of the subcluster-gas and subcluster-relic offsets

are nearly identical for their respective subclusters. While we expect the offsets to be correlated, since the relic shock wave propagation velocity is related to the collision velocity of the subclusters, it is likely just a coincidence that they are nearly equal. At some other merger time they should be unequal since the subcluster galaxies are gravitating while the radio relic shock is not.

Ogrean et al. (2013) find a "wall" of hot gas east of the cluster center. They note that a simple binary merger is not expected to create such a feature and suggest that it may be indicative of a more complex merger scenario (e.g., a triple merger), or "a lack of understanding on our part of the complex structures formed during real cluster mergers." Other than an insignificant substructure we find no evidence for a complex multiple merger scenario, and instead our analysis favors a relatively clean bimodal merger. This then suggests that our understanding of how complex X-ray structures are formed during cluster mergers is incomplete.

As discussed in Section 1, Ogrean et al. (2014) found evidence for two inner density X-ray discontinuities, trailing the northern and southern radio relics by ∼0.5 Mpc. They argue that these discontinuities are not likely cold fronts given that their large distance from the cluster center (≈1.5 Mpc) would make them the most distant cold fronts ever detected. However, the location of these potential cold fronts should be measured with respect to the two subclusters and not the cluster system center. In this case the potential cold fronts are offset less than 500 kpc from the centers of their respective subclusters. However, this does not help explain the anomalously high temperatures of ∼8–9 keV. Our observations may also lend support to the possibility that the inner density discontinuities are the result of violent relaxation of the DM halos since the galaxies suggest that their are two DM halos near the discontinuities.

Stroe et al. (2014c) find an order-of-magnitude boost in the normalization of the galaxy luminosity function in the vicinity of the relics. Our analysis finds no evidence that this boost in star formation is simply due to coincidental infalling groups or other substructure, suggesting that the enhanced star formation may be the result of cluster members responding to changes in their environment due to the merger.

Our finding of two dominant subclusters, with similar mass based on velocity dispersions, is consistent with our weak-lensing analysis of the system (Jee et al. 2014c). In that paper we perform a detailed analysis comparing the relative locations of the galaxies and mass.

7.2. BCG-subcluster Offsets

In Section 6.1.1 we noted that the BCGs of the north and south subclusters are near the subcluster peak locations, 55 and 85 kpc, respectively, and within the 68% confidence uncertainty of each (∼90 kpc). In Section 6.2 we calculated the relative LOS velocity differences of each subcluster's BCG with respect to the median subcluster redshift and found $v_{\rm north} - v_{\rm north\ BCG} = -500 \pm 140$ and $v_{\rm south} - v_{\rm south\ BCG} = -240 \pm 130$ km s^{-1}. Normalizing these velocities by the respective velocity dispersion of each subcluster ($|\Delta v|/\sigma_v$) results in 0.43 ± 0.13 and 0.21 ± 0.12 for the north and south, respectively. We compare the projected and LOS velocity offsets of the BCGs relative to their subclusters with the empirical PDFs of Einasto et al. (2012) for single-component (i.e., relaxed) and multi-component (i.e., disturbed) clusters,

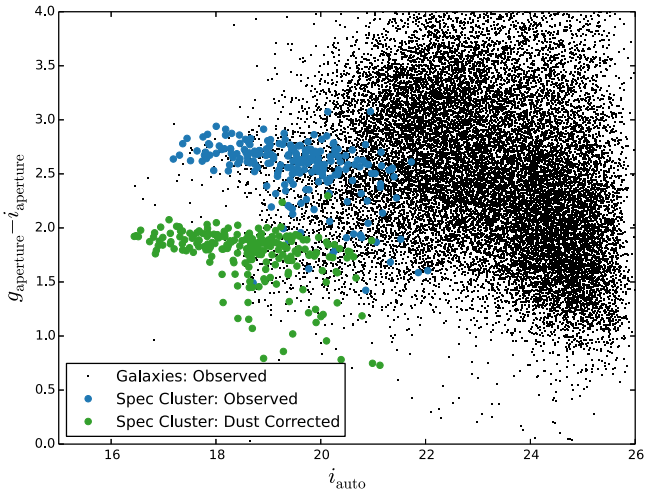

Figure A1. Color–magnitude diagram of galaxies within a 15′ radius of the system center, based on non-dust-corrected Subaru g and i magnitudes (*observed*). Spectroscopic cluster members using non-dust-corrected magnitudes are shown in blue. For comparison the spectroscopic cluster members using dust-corrected magnitudes are copied from Figure 7 and shown in green. The red sequence is noticeably tighter in color space after applying the dust corrections, which is reflective of the highly varying attenuation across the field.

and we find that the CIZA J2242.8+5301 BCG offsets are consistent with offsets observed in both single- and multi-component clusters, although with a larger joint likelihood for the single-component cluster PDFs.

If a cluster potential remains relatively static (i.e., does not undergo major mergers), then the most massive galaxy (i.e., BCG) should migrate toward the potential minimum at the center of the cluster largely due to dynamical friction (Lecar 1975; Ostriker & Tremaine 1975; Merritt 1983; Malumuth 1992). Given the age of the universe at the CIZA J2242.8+5301 cluster redshift (∼11 Gyr), and that the subclusters likely collided ∼0.8 Gyr ago (Stroe et al. 2014a; W. A. Dawson et al. 2015, in preparation; H. Akamatsu et al. 2015 in preparation), there was sufficient time for each subcluster to have formed a BCG and for that BCG to have migrated to the center of the cluster prior to the merger (Malumuth 1992). This assumes that there were no major mergers in the ∼2 Gyr prior to the current merger (the approximate timescale it takes the BCGs to settle to the center of their subcluster via dynamical friction after being offset by a past merger; Merritt 1984; Martel et al. 2014). This assumption is consistent with the fact that we find no significant evidence for major substructure in either the north or south subclusters. Given these assumptions, we argue that the respective projected BCG-subcluster center offset and the BCG-subcluster LOS velocity difference prior to the merger were likely small. If this is the case, then the CIZA J2242.8+5301 major merger appears to have caused BCG offsets of $\lesssim$100 kpc in projection and $\lesssim$0.5 in $|\Delta v|/\sigma_v$. This appears to be in tension with the commonly proposed argument (Martel et al. 2014 and references therein) that major cluster mergers are responsible for the observed large BCG offsets (∼1 Mpc) and large relative BCG velocity differences ($|\Delta v|/\sigma_v \sim 1.5$; see, e.g., Einasto et al. 2012; Lauer et al. 2014).

Martel et al. (2014) provide an example from their simulations where a BCG-cluster center offset of 2.2 Mpc is experienced in the early stages of a major merger (see Figure 9

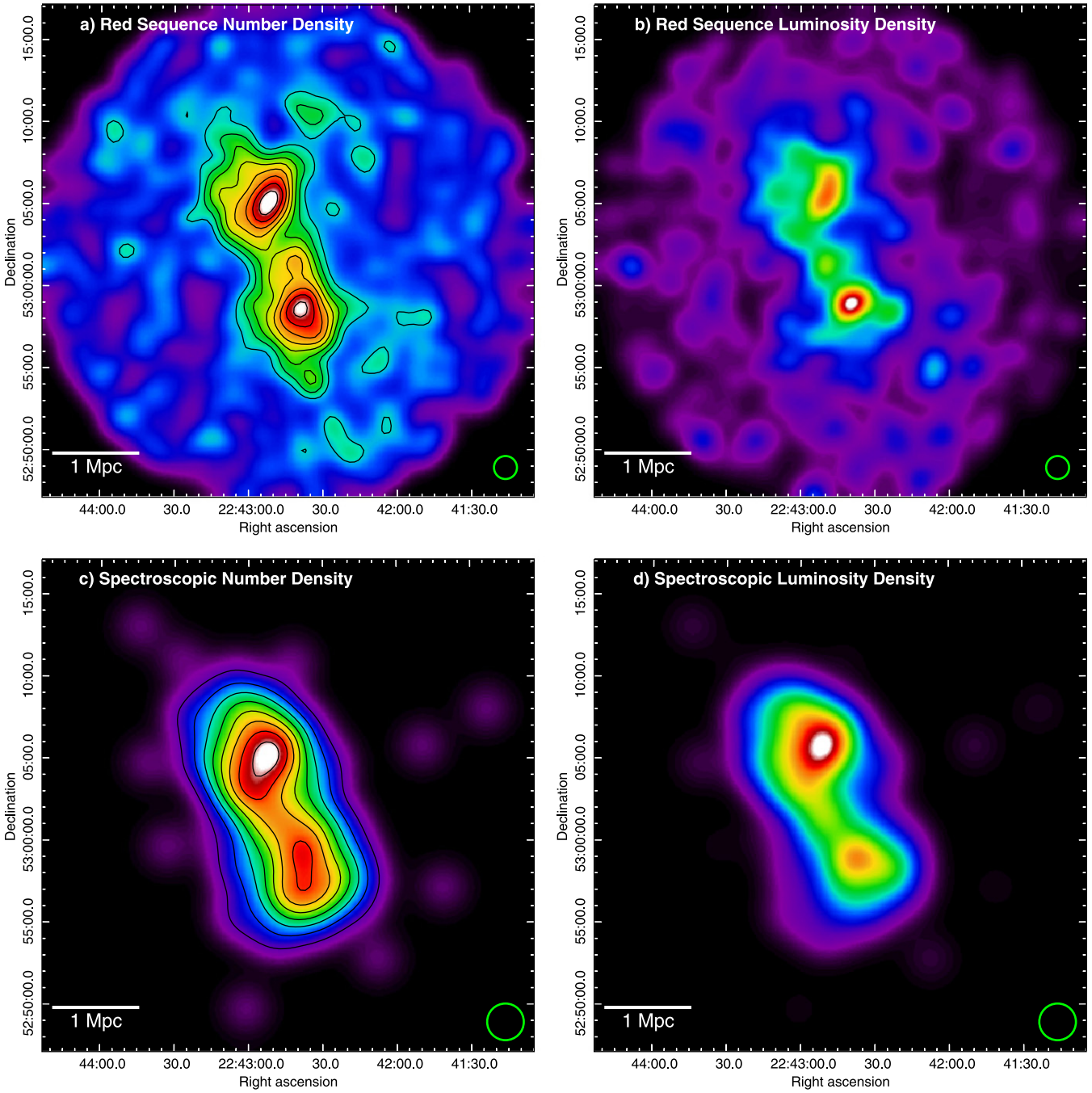

Figure A2. The four galaxy density maps discussed in Section 5.2. Each map has a linear scaling with black being less and white being more galaxies Mpc^{-2}. The KDE bandwidth (i.e., smoothing scale) is shown by the green circle in the lower right of each map. The white bar in the lower left gives the physical scale at the cluster redshift $z = 0.188$. (a) The cluster red sequence sample number density map. The cluster galaxy number density contours (black) begin at 100 galaxies Mpc^{-2} and increase linearly with increments of 25 galaxies Mpc^{-2}. The scale has been corrected for contamination (see discussion in Section 4.2). (b) The cluster red sequence sample luminosity density map. (c) The cluster spectroscopic sample number density map. The cluster spectroscopic number density contours (black) begin at 10 galaxies Mpc^{-2} and increase linearly with increments of 10 galaxies Mpc^{-2}. (d) The cluster spectroscopic luminosity density map.

and the related discussion in their paper). It is clear from the $z = 0.53$ frame of their Figure 9 that such a large offset is an artifact of their friends-of-friends numerical algorithm linking three clearly distinct subclsuters. Much in the same way, we would find a BCG-cluster center offset of ∼0.9 Mpc for CIZA J2242.8+5301 if we were not to distinguish between the northern and southern subclusters and treat the galaxy population as a single cluster or similarly if we were to measure the BCG offset with respect to the peak of the X-ray distribution (e.g., as done in Lauer et al. 2014). This is consistent with the finding of Beers & Geller (1983) that D or cD galaxies are located near $(180 \pm 50\,\mathrm{kpc}\,h_0^{-1})$ local galaxy density peaks and the finding of Einasto et al. (2012) that the BCG offsets are smaller with respect to GMM component centers (i.e., subcluster centers) than the cluster system center. This highlights the importance of carefully defining the cluster center, since different definitions can result in offset estimates that differ by ∼1 dex. With regard to studying the evolution of

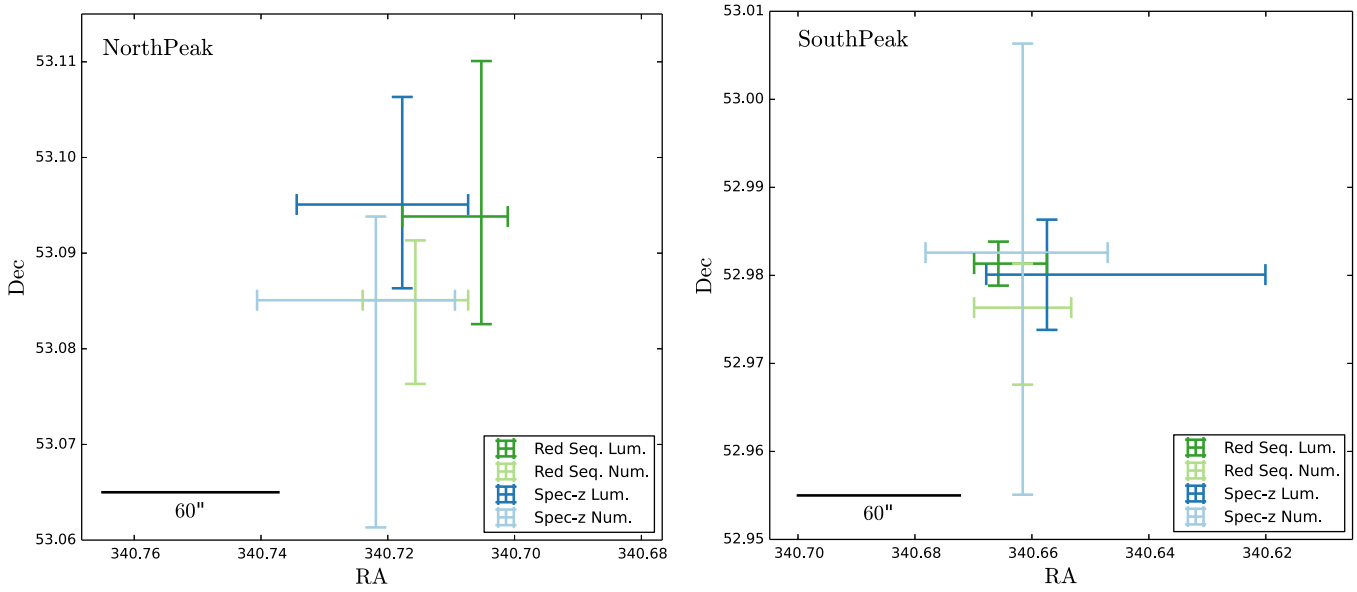

Figure A3. Comparison of the peak 65% confidence intervals for the north (left) and south (right) subclusters, as measured for the red sequence number (light green) and luminosity (dark green) projected density maps, and spectroscopic cluster sample number (light blue) and luminosity (dark blue) projected density maps. The dimensions of each figure are chosen to prevent anamorphic distortions. We find general agreement between each location measure.

BCGs, it seems more appropriate to measure their offsets with respect to any significant local galaxy density peak (i.e., subcluster center), since the effectively collisionless galaxies are expected to be better tracers of the local gravitational potential than the cluster gas.

8. CONCLUSIONS

We have presented a comprehensive broadband imaging (Section 2) and spectroscopic survey (Section 3) of the CIZA J2242.8+5301 cluster's galaxy population and used this information to provide a new perspective on the cluster's global properties. We confirm that the cluster is a textbook dissociative major merger, with the bulk of the gas being offset between two galaxy subclusters. We also find excellent agreement between the merger axis inferred from the two radio relics, elongated X-ray gas, and bimodal galaxy distribution (see Figure 1).

We find that the system is dominated by two subclusters of comparable richness (Section 5) and accurately measure their locations (Section 6.1), which imply a projected separation of $6\farcm9^{+0.7}_{-0.5}$ ($1.3\,\mathrm{Mpc}^{+0.13}_{-0.10}$). We find that the north and south subclusters have redshifts of $0.18794^{+0.00054}_{-0.00054}$ and $0.18821^{+0.00054}_{-0.00052}$, respectively, corresponding to a relative LOS velocity of 69 ± 190 km s^{-1}. This is consistent with previous suggestions that the merger is occurring close to the plane of the sky; however, without a more detailed dynamics analysis we cannot rule out the possibility that the merger has a larger inclination angle and is just being observed near the merger apocenter.

We also find that north and south subclusters have velocity dispersions of 1160^{+100}_{-90} and 1080^{+100}_{-70} km s^{-1}, respectively (Section 6.2). These correspond to masses of $16.1^{+4.6}_{-3.3} \times 10^{14}$ and $13.0^{+4.0}_{-2.5} \times 10^{14}$ $M_\odot$, respectively. While velocity dispersion measurements of merging clusters can be biased, we believe the bias in this system to be minor due to the large projected separation and nearly plane-of-sky merger configuration. In this regard we find the velocity dispersion inferred masses to be consistent with our weak-lensing inferred masses (Jee et al. 2014c).

We also find that the BCGs of the north and south subclusters are very near their subcluster centers, in both projection (55 and 85 kpc respectively) and normalized LOS velocity ($|\Delta v|/\sigma_v = 0.43 \pm 0.13$ and 0.21 ± 0.12 for the north and south, respectively), suggesting that the major merger had little impact on the location of the BCGs with respect to their parent subclusters. CIZA J2242.8+5301 is a relatively clean dissociative cluster merger, potentially occurring near the plane of the sky, with near 1:1 mass ratio, which makes it an ideal merger for modeling merger-associated physical phenomena.

We would like to thank the broader membership of the MC^2 for their continual development of the science motivating this work, and which has been instrumental in the acquisition of the data used in this paper. We would like to thank Anja von der Linden for the initial recommendation to rotate the Subaru SuprimeCam instrument 90° between exposures to better probe instrument systematics. Slight modifications to this strategy enabled us to reduce the effects of stellar bleeds and nearly double the number of detected objects. We would like to thank Cristbal Sifón Andalaft and the referee for suggestions regarding additional substructure tests. We would also like to thank the referee for suggesting added discussion of the subcluster-BCG offsets. We also would like to thank Michael Schneider for valuable feedback regarding the presentation of the current work. M.J.J., D.W., and W.D. acknowledge support from *HST*-GO-13343.01-A. A.S. acknowledges financial support from NWO. D.S. acknowledges financial support from the Netherlands Organisation for Scientific research (NWO) through a Veni fellowship, from FCT through an FCT Investigator Starting Grant and Start-up Grant (IF/01154/2012/CP0189/CT0010), and from FCT grant PEst-OE/FIS/UI2751/2014. R.W. is supported by NASA through the Einstein Postdoctoral grant number PF2-130104 awarded by the *Chandra* X-ray Center, which is operated by the

Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060. M.B. acknowledges support by the research group FOR 1254 funded by the Deutsche Forschungsgemeinschaft. M.B. acknowledges funding from the Deutsche Forschungsgemeinschaft under SFB 676. Part of this work was performed under the auspices of the U.S. DOE by LLNL under Contract DE-AC52-07NA27344. This research has made use of NASA's Astrophysics Data System. The William Herschel Telescope and Isaac Newton Telescope are operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. Funding for the DEEP2/DEIMOS pipelines has been provided by NSF grant AST-0071048. The DEIMOS spectrograph was funded by grants from CARA (Keck Observatory) and UCO/Lick Observatory, an NSF Facilities and Infrastructure grant (ARI92-14621), the Center for Particle Astrophysics, and gifts from Sun Microsystems and the Quantum Corporation.

Facilities: Keck:II (DEIMOS), Subaru (Suprime-Cam), CFHT (MegaCam), ING:Herschel (AF2), ING:Newton (WFC), XMM, WSRT.

APPENDIX A
EFFECT OF GALACTIC EXTINCTION ON THE COLOR–MAGNITUDE RELATION

As discussed in Stroe et al. (2014c) and Jee et al. (2014c), the Galactic dust extinction in the CIZA J2242.8+5301 field results in high attenuation (A_v=1.382), which varies significantly across the field ($0.35 < E(B-V) < 0.52$; Schlafly & Finkbeiner 2011). We use extinction values from Schlafly & Finkbeiner (2011) to recover reddened magnitudes. Given the spatial resolution of $\sim 4'$, we interpolate between the extinction pierce points using cubic interpolation to predict 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).

In Figure A1 we plot the color–magnitude diagram using the non-dust-corrected magnitudes (*observed*). For comparison we copy the spectroscopic cluster members using dust-corrected magnitudes from Figure 7 and shown in green. The red sequence is noticeably tighter in color space after applying the dust corrections, which is reflective of the highly varying attenuation across the field. Additionally, the translation of the red sequence in the $i_{\rm auto}$ direction illustrates the large average attenuation ($A_v \sim 0.75$) across the field.

APPENDIX B
GALAXY DENSITY COMPARISONS

Here we compare the galaxy density maps for the red sequence selection (Section 4.2) and spectroscopic cluster member selection (Section 4.1) samples, with and without galaxy luminosity weighting. This is an extension of our projected galaxy density discussion in Section 5.2.

We find that each of these four representations has the same general features (see Figure A2): two dominant subclusters (one in the north and one in the south) that are aligned closely with the merger axis inferred from the radio relics (van Weeren et al. 2010) and elongated X-ray gas distribution (Ogrean et al. 2013); see, for example, the red sequence sample number density map in Figure 10 (all four maps are shown in Figure A2 of the Appendix). We also find close agreement between the subcluster peak locations for each representation. There are some notable differences between the maps. The number density maps result in slightly less concentrated distributions compared to the luminosity density maps. The red sequence sample luminosity density map shows that the southern subcluster is more densely concentrated relative to the northern subcluster (see Figure 10), whereas the spectroscopic sample luminosity density map shows them to be about the same. This can be explained by spectroscopic undersampling bias (see Section 4.1 for discussion). In the zoom of the southern subcluster region in Figure 10 it is apparent that there are a number of very densely packed bright cluster galaxies (compare with the less densely packed northern subcluster peak). Thus, we were able to obtain a more complete spectroscopic survey of the northern subcluster.

REFERENCES

Akamatsu, H., & Kawahara, H. 2013, PASJ, 65, 16
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Beers, T. C., & Geller, M. J. 1983, ApJ, 274, 491
Bradač, M., Allen, S. W., Treu, T., et al. 2008, ApJ, 687, 959
Bradač, M., Clowe, D., Gonzalez, A. H., et al. 2006, ApJ, 652, 937
Chung, S. M., Gonzalez, A. H., Clowe, D., Markevitch, M., & Zaritsky, D. 2010, ApJ, 725, 1536
Clowe, D., Gonzalez, A., & Markevitch, M. 2004, ApJ, 604, 596
Dawson, W. A. 2013a, ApJ, 772, 131
Dawson, W. A. 2013b, PhD thesis, UMI Dissertations Publishing
Dawson, W. A., Wittman, D., Jee, M. J., et al. 2012, ApJL, 747, L42
Dressler, A., & Shectman, S. A. 1988, AJ, 95, 985
Einasto, M., Vennik, J., Nurmi, P., et al. 2012, A&A, 540, A123
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
Feigelson, E., & Babu, G. 2012, Modern Statistical Methods for Astronomy: With R Applications (Cambridge: Cambridge Univ. Press)
Feretti, L., Giovannini, G., Govoni, F., & Murgia, M. 2012, A&ARv, 20, 54
Ferrari, C., Benoist, C., Maurogordato, S., Cappi, A., & Slezak, E. 2005, A&A, 430, 19
Filippenko, A. V. 1982, PASP, 94, 715
Frie, Z., & Gunn, J. E. 1994, AJ, 108, 1476
Hwang, H. S., & Lee, M. G. 2009, MNRAS, 397, 2111
Ivezić, Ž, Connolly, A., Vanderplas, J., & Gray, A. 2014, Statistics, Data Mining and Machine Learning in Astronomy (Princeton, NJ: Princeton Univ. Press)
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., Stroe, A., Dawson, W., et al. 2014c, arXiv:1410.2898
Kang, H., Ryu, D., & Jones, T. W. 2012, ApJ, 756, 97
Kass, R. E., & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773
Kocevski, D. D., Ebeling, H., Mullis, C. R., & Tully, R. B. 2007, ApJ, 662, 224
Lauer, T. R., Postman, M., Strauss, M. A., Graves, G. J., & Chisari, N. E. 2014, ApJ, 797, 82
Lecar, M. 1975, in IAU Symp. 69, Dynamics of the Solar Systems, ed. A. Hayli (Cambridge: Cambridge Univ. Press), 161
Mahdavi, A., Hoekstra, H., Babul, A., Balam, D. D., & Capak, P. L. 2007, ApJ, 668, 806
Malumuth, E. M. 1992, ApJ, 386, 420
Markevitch, M., Gonzalez, A. H., David, L., et al. 2002, ApJL, 567, L27
Markevitch, M., Govoni, F., Brunetti, G., & Jerius, D. 2005, ApJ, 627, 733

Martel, H., Robichaud, F., & Barai, P. 2014, ApJ, 786, 79
Merritt, D. 1983, ApJ, 264, 24
Merritt, D. 1984, ApJ, 276, 26
Merten, J., Coe, D., Dupke, R., et al. 2011, MNRAS, 417, 333
Miller, N. A., & Owen, F. N. 2003, AJ, 125, 2427
Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, ApJS, 208, 5
Ng, K. Y., Dawson, W. A., Wittman, D., et al. 2014, arXiv:1412.1826
Norberg, P., Baugh, C. M., Hawkins, E., et al. 2001, MNRAS, 328, 64
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
Ostriker, J. P., & Tremaine, S. D. 1975, ApJL, 202, L113
Owen, F. N., Ledlow, M. J., Keel, W. C., Wang, Q. D., & Morrison, G. E. 2005, AJ, 129, 31
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825
Pinkney, J., Roettiger, K., Burns, J. O., & Bird, C. M. 1996, ApJS, 104, 1
Poggianti, B. M., Bridges, T. J., Komiyama, Y., et al. 2004, ApJ, 601, 197
Randall, S. W., Markevitch, M., Clowe, D., Gonzalez, A. H., & Bradač, M. 2008, ApJ, 679, 1173
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Sobral, D., Stroe, A., Dawson, W. A., et al. 2015, MNRAS, in press (arXiv:1503.0207)
Stone, C. J. 1984, AnSta, 12, 1285
Stroe, A., Harwood, J. J., Hardcastle, M. J., & Röttgering, H. J. A. 2014a, MNRAS, 445, 1213
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
Valdes, F. G. 1998, in ASP Conf. Ser. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco, CA: ASP), 53
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
von der Linden, A., Allen, M. T., Applegate, D. E., et al. 2014, MNRAS, 439, 2