Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **484,** 1946–1971 (2019) doi:10.1093/mnras/stz088
Advance Access publication 2019 January 10

X-ray scaling relations from a complete sample of the richest maxBCG clusters

Chong Ge,[1]⋆ Ming Sun,[1]⋆ Eduardo Rozo,[2] Neelima Sehgal,[3] Alexey Vikhlinin,[4] William Forman,[4] Christine Jones[4] and Daisuke Nagai[5,6,7]

[1] *Department of Physics and Astronomy, University of Alabama in Huntsville, Huntsville, AL 35899, USA*
[2] *Department of Physics, University of Arizona, Tucson, AZ 85721, USA*
[3] *Physics and Astronomy Department, Stony Brook University, Stony Brook, NY 11794, USA*
[4] *Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[5] *Department of Physics, Yale University, New Haven, CT 06520, USA*
[6] *Department of Astronomy, Yale University, New Haven, CT 06520, USA*
[7] *Yale Center for Astronomy & Astrophysics, Yale University, New Haven, CT 06520, USA*

Accepted 2019 January 5. Received 2019 January 5; in original form 2018 March 13

ABSTRACT

We use a complete sample of 38 richest maxBCG clusters to study the ICM–galaxy scaling relations and the halo mass selection properties of the maxBCG algorithm, based on X-ray and optical observations. The clusters are selected from the two largest bins of optical richness in the *Planck* stacking work with the maxBCG richness $N_{200} \geq 78$. We analyse their *Chandra* and *XMM–Newton* data to derive the X-ray properties of the ICM. We then use the distribution of $P(X|N)$, $X = T_X, L_X, Y_X$, to study the mass selection $P(M|N)$ of maxBCG. Compared with previous works based on the whole richness sample, a significant fraction of blended systems with boosted richness is skewed into this richest sample. Parts of the blended haloes are picked apart by the redMaPPer, an updated red-sequence cluster finding algorithm with lower mass scatter. Moreover, all the optical blended haloes are resolved as individual X-ray haloes, following the established L_X–T_X and L_X–Y_X relations. We further discuss that the discrepancy between ICM–galaxy scaling relations, especially for future blind stacking, can come from several factors, including miscentring, projection, contamination of low-mass systems, mass bias, and covariance bias. We also evaluate the fractions of relaxed and cool core clusters in our sample. Both are smaller than those from SZ or X-ray selected samples. Moreover, disturbed clusters show a higher level of mass bias than relaxed clusters.

Key words: galaxies: clusters: general – galaxies: clusters: intracluster medium – X-rays: galaxies: clusters.

1 INTRODUCTION

The spatial distribution and number density of massive dark matter haloes are very sensitive to the underlying cosmological parameters. These haloes host clusters of galaxies, and they have been used to study cosmology for over 20 yr (e.g. Henry & Arnaud 1991; Eke et al. 1998; Reiprich & Böhringer 2002; Vikhlinin et al. 2009b; Rozo et al. 2010; Allen, Evrard & Mantz 2011; Sehgal et al. 2011; Weinberg et al. 2013). The key questions to study cosmology with clusters are how to find them and how to measure their masses, which can be measured by the velocity dispersion of cluster galaxies and gravitational lensing, both of them are observationally expensive. They are not efficient methods for finding clusters, though they are important tools for direct mass calibration. These measurements find that most cluster mass is in dark matter, which cannot be detected directly, thus more conveniently the cluster finding and mass measuring are tied to baryons in clusters.

A small fraction of cluster baryons cooled and formed stars. Clusters can be selected optically by looking for overdense regions of galaxies. The optical identifications have resulted in large samples of clusters, benefiting from large optical surveys like *SDSS* (e.g. Koester et al. 2007; Rykoff et al. 2014). Mass proxies in optical are generally the total number or luminosity of member galaxies but such scaling relations suffer from large fractional scatter (e.g. $\sigma_{m|n} = 0.45$; Rozo et al. 2009a).

Most of the cluster baryons are in the hot intracluster medium (ICM). The ICM emits X-rays via thermal bremsstrahlung and line emission, which are proportional to the density square of the gas

---

⋆ E-mail: chong.ge@uah.edu (CG); ming.sun@uah.edu (MS)

Published by Oxford University Press on behalf of the Royal Astronomical Society

that traces the deep potential well of the cluster. Thus, the X-ray detection is very effective to find clusters and relatively free from projection contamination. Meanwhile, the cluster mass can be derived from the distribution of their X-ray emitting gas under the assumption of hydrostatic equilibrium (HE). A more economical way is using X-ray mass proxies like T_X, L_X, and Y_X through scaling relations (e.g. Kaiser 1986; Pratt et al. 2009; Vikhlinin et al. 2009a; Sun et al. 2009; Mantz et al. 2010; Sun et al. 2011; Sun 2012). These X-ray mass proxies, usually considered robust and low scattered, can also be calibrated with the robust mass indicators like weak-lensing mass (e.g. $\sigma_{m|t} = 0.26$; Mantz et al. 2016). The ICM scaling relations are also important to understand the key baryon physics that governs galaxy formation, e.g. cooling, efficiency of star formation, and stellar or AGN feedback (e.g. Kravtsov, Nagai & Vikhlinin 2005; Croton et al. 2006; Bower, McCarthy & Benson 2008; Le Brun et al. 2014; Hahn et al. 2017; Truong et al. 2018).

The X-ray emitting hot electrons also boost microwave background radiation (CMB) photons to higher energy through inverse Compton scattering (Sunyaev–Zel'dovich – SZ effect; Carlstrom, Holder & Reese 2002). The SZ signal, which is proportional to gas density and temperature, is a promising way to find clusters and complementary to X-ray measurements. Tremendous progress has been made in the past 10 yr with the SZ surveys from *Planck*, the Atacama Cosmology Telescope (*ACT*), and the South Pole Telescope (*SPT*). Hundreds of new clusters have been discovered (e.g. Planck Collaboration XXVII 2016). Scaling relations between the SZ observables and the other observables in X-ray and optical have been established (e.g. Planck Collaboration XI 2011). The total SZ flux has been considered a robust mass proxy, as it measures the total thermal energy of the ICM. However, the new SZ data have also raised new puzzles. One that has received much attention is the mismatch between the stacked *Planck* SZ fluxes and the model expectations for the optically selected clusters. All the model SZ fluxes from optically selected clusters overpredict the observed SZ fluxes from *Planck* (Planck Collaboration XII 2011), *ACT* (Sehgal et al. 2013), and *SPT* (Saro et al. 2017). Possible solutions include (1) miscentring because of the offset between the optical centre and the ICM peak; (2) contamination of optical richness estimates from line-of-sight (LOS) projections; (3) contamination of the SZ flux from radio galaxies, infrared galaxies, Galactic emission, and the SZ background from unresolved clusters, groups, and the intergalactic medium; (4) underestimate of the true mass by the X-ray HE mass because of additional non-thermal pressure from e.g. gas bulk motion and turbulence; (5) property covariance and X-ray selection bias (e.g. Rykoff et al. 2008; Planck Collaboration XII 2011; Biesiadzinski et al. 2012; Sehgal et al. 2013; Evrard et al. 2014; Rozo et al. 2014b; Saro et al. 2017). Among them, the hydrostatic mass bias seems the dominate source for the discrepancy with a typical value $1 - b \sim 0.75$ (e.g. Henson et al. 2017; Penna-Lima et al. 2017; Miyatake et al. 2018).

While the previous studies addressing this puzzle require the calibration of the true mass and assume pressure templates, we bypass this intermediate step to directly compare the X-ray properties with the optical richness, which further allows a direct comparison of the pressure content from X-rays and SZ in the future. We target the most massive maxBCG clusters as the discrepancy exists in all richness bins. These clusters should have the strongest X-ray and SZ signals. It is natural to study them directly to examine the reasons for the discrepancy. There are 38 clusters in the two richest bins (fig. 1 in Planck Collaboration XII 2011 and Table 1). When we first started this project, 31 of them had existing *Chandra* or *XMM* data, and were mostly X-ray selected. The remaining seven clusters are generally X-ray faint from the *ROSAT* All Sky Survey (RASS). We proposed new *XMM* observations (proposal id 74251 and 76159, PI: M. Sun) on these seven clusters (shown in Table 2) to have full X-ray coverage of this sample. This sample also provides an extremal study of halo mass selection properties of the maxBCG algorithm. What is being indirectly tested here (ignoring z-dependence) is the conditional probability $P(M_{500}|N_{200})$, which is the probability of a mass M_{500} halo with given N_{200} galaxies, where N_{200} is the maxBCG richness. The probability is typically assumed to be a lognormal distribution that characterized by a mean relation and an rms scatter. What is actually measured is $P(X|N_{200})$, with $X = \{T_X, L_X, Y_X\}$. In addition to the maxBCG catalogue, we also use the redMaPPer catalogue (Rykoff et al. 2014), which uses a similar but different richness measure with different halo selection properties from maxBCG.

We organize the paper as follows: Section 2 includes the cluster sample and catalogues; Section 3 details our X-ray data reduction and analysis; Section 4 presents results based on X-ray data and optical catalogues; Section 5 is the discussion; and Section 6 is our conclusions. We use the standard cold dark matter cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

2 CLUSTER SAMPLE

Table 1 lists the 38 richest maxBCG clusters with optical and SZ properties assembled from maxBCG, redMaPPer, and PSZ2 catalogues, which are introduced briefly below. We use N as a general representative of optical richness, particularly N_{200} for maxBCG and λ for redMaPPer.

2.1 maxBCG catalogue

The maxBCG cluster catalogue (Koester et al. 2007) contains 13 823 clusters selected from *SDSS* Data Release 5 (DR5), with photometric redshifts of 0.1–0.3 over approximately volume-limited 7500 deg^2 of sky. Each cluster is selected as a spatial overdensity of red-sequence galaxies that form a tight E/S0 ridgeline in the colour–magnitude diagram. The cluster centre is on the brightest cluster galaxy (BCG) with the highest likelihood. The tight relation between the ridgeline colour and redshift also provides an accurate photometric redshift estimate for each cluster ($\Delta_z \simeq 0.01$). The cluster richness N_{200} is defined as the number of red-sequence galaxies with rest-frame i-band luminosity $L_i > 0.4L_*$ within a projected scaled radius R_{200}, interior to which the density of galaxies with $-24 < M_r < -16$ is 200 times the mean density of such kind of galaxies. Applying the cluster selection algorithm to mock catalogues suggests that the catalogue is ~90 per cent pure and ~85 per cent complete for clusters with masses $\geq 10^{14}\,\mathrm{M}_\odot$. The maxBCG cluster mass has been calibrated (Johnston et al. 2007; Rozo et al. 2009a) with the M_{500c}–N_{200} relation.

2.2 redMaPPer catalogue

The redMaPPer cluster catalogue (v6.3; Rykoff et al. 2014) includes 26 111 clusters identified from *SDSS* DR8, with photometric redshifts of 0.08–0.60 (volume-limited in $z \leq 0.35$), covering nearly 10 000 deg^2 of sky. Each cluster is identified as an overdensity of red-sequence galaxies, which relies on iteratively self-training a model of the red-sequence as a function of redshift based on the red galaxies with known spectroscopic redshifts. Then this model is used on photometric data to assign membership probabilities to galaxies with luminosities $\geq 0.2L_*$ in the cluster vicinity. The

Table 1. Optical and SZ properties of the richest maxBCG clusters.

Cluster	RA[a]	Dec.[a]	z_{photo}[a]	N_{200}[a]	redMaPPer[b] RM	λ[b]	offset[b] arcmin	*Planck*[c] PSZ2	SNR[c]	offset[c] arcmin
A2142	239.58334	27.23342	0.103	188	J155820.0+271400.3	169.8 ± 4.2	0.00	G044.20+48.66	28.4	1.56
J150	150.68705	20.49151	0.297	164	J100214.1+203216.6	151.4 ± 5.8	7.72	G213.30+50.99	4.5	3.91
					J100311.3+203230.2	54.7 ± 4.4	6.88			
A1689	197.87292	– 1.34110	0.189	156	J131129.5–012028.0	164.7 ± 4.2	0.00	G313.33+61.13	16.7	0.81
A1443	180.31985	23.10909	0.262	139	J120112.2+230557.3	131.3 ± 5.9	1.20	G229.74+77.96	12.0	0.87
A781	140.10742	30.49406	0.292	126	J092052.5+302740.3	128.6 ± 5.7	6.09	G195.60+44.06	7.6	0.72
					J092030.0+302946.8	38.7 ± 5.2	0.92			
A1986	223.37531	21.87434	0.116	120	J145308.4+215339.3	68.8 ± 3.2	5.18	–	–	–
A1882	213.78496	– 0.49325	0.135	115	J141424.1–002239.5	105.1 ± 3.8	13.07	–	–	–
A1758	203.16008	50.55992	0.284	114	J133238.4+503336.0	152.9 ± 5.7	0.00	G107.10+65.32	16.5	2.59
					J133224.5+502425.8	51.9 ± 4.2	9.43			
A1760	203.53616	20.24806	0.176	113	J133431.5+201217.7	73.4 ± 3.7	5.94	G000.13+78.04	9.3	1.35
					J133359.1+201801.3	26.5 ± 3.7	3.85			
A1622	192.38517	49.89549	0.275	107	J124922.1+494742.1	129.3 ± 5.6	6.25	G123.66+67.25	5.7	2.02
A750	137.21745	11.02649	0.173	107	J090912.2+105824.9	174.7 ± 4.9	5.84	G218.81+35.51	8.4	4.84
A1682	196.69040	46.55854	0.235	106	J130650.0+463333.4	122.4 ± 5.3	0.74	G114.99+70.36	10.0	2.52
A1246	171.01653	21.49114	0.186	103	J112358.8+212849.7	113.6 ± 4.1	1.36	G224.00+69.33	8.8	0.67
A1961	221.07852	31.28601	0.240	101	J144431.8+311336.0	114.8 ± 4.6	4.51	G049.18+65.05	6.1	3.79
A2034	227.54883	33.48646	0.122	99	J151011.7+332911.3	99.8 ± 3.7	0.00	G053.53+59.52	15.9	0.42
A655	126.37104	47.13348	0.135	99	J082529.1+470800.9	131.6 ± 4.8	0.01	G172.63+35.15	7.4	1.19
A1914	216.48612	37.81645	0.167	98	J142556.7+374859.2	103.3 ± 4.0	0.00	G067.17+67.46	17.8	0.81
Z5247	188.57278	9.76624	0.243	98	J123424.1+094715.5	93.0 ± 4.2	2.08	G289.13+72.19	8.8	2.71
A657	125.83033	15.96272	0.157	94	J082319.3+155745.8	82.1 ± 3.9	0.00	–	–	–
J229	229.35068	– 0.73816	0.119	93	J151721.9–004256.3	58.1 ± 3.0	1.47	–	–	–
A1423	179.32219	33.61093	0.219	91	J115715.2+333638.5	86.0 ± 4.2	0.45	G180.60+76.65	11.0	0.56
A801	142.01889	20.52921	0.208	91	J092804.5+203145.1	93.0 ± 4.5	0.00	G209.53+43.32	4.8	2.32
A773	139.47261	51.72704	0.224	90	J091753.4+514337.5	153.6 ± 4.7	0.00	G166.09+43.38	12.4	0.51
A1576	189.24684	63.18658	0.294	89	J123658.6+631114.1	102.6 ± 4.8	0.08	G125.71+53.86	11.4	0.60
A2631	354.41554	0.27138	0.286	88	J233739.7+001616.9	126.3 ± 4.9	0.00	G087.03–57.37	11.2	1.00
A1703	198.77182	51.81738	0.286	87	J131505.2+514902.8	152.5 ± 4.9	0.00	G114.31+64.89	11.1	0.96
A2219	250.08254	46.71148	0.235	85	J164019.8+464241.5	199.5 ± 5.3	0.00	G072.62+41.46	27.4	0.48
A1319	174.05437	40.04327	0.292	85	J113613.0+400235.8	155.1 ± 5.9	0.00	G168.33+69.73	5.6	3.60
A1560	188.47087	15.19464	0.275	84	J123416.1+151508.4	136.9 ± 6.1	6.57	G283.26+77.37	6.0	1.53
J175	175.57289	10.29192	0.127	84	J114207.4+100857.6	38.4 ± 2.6	8.91	–	–	–
					J114215.3+102649.4	23.5 ± 2.4	9.32			
J249	249.89850	47.07569	0.227	84	J163938.0+470310.4	61.8 ± 3.4	1.43	–	–	–
A1201	168.22709	13.43584	0.176	83	J111250.1+132830.4	78.3 ± 3.5	2.59	G238.69+63.26	7.3	3.01
A2009	225.08133	21.36941	0.165	82	J150019.5+212209.9	91.8 ± 3.6	0.00	G028.89+60.13	7.5	0.79
A2111	234.91872	34.42425	0.243	81	J153940.5+342527.3	151.5 ± 5.2	0.00	G054.99+53.41	8.4	0.94
A815	143.03590	29.06782	0.281	80	J093208.6+290404.1	92.2 ± 5.9	0.00	–	–	–
					J093251.2+290318.0	33.3 ± 3.3	9.33			
Z1450	223.05481	37.87803	0.284	79	J145213.1+375251.9	101.9 ± 4.5	0.18	–	–	–
A1765	204.20577	10.44002	0.159	78	J133649.4+102624.1	82.7 ± 3.7	0.00	–	–	–
A1902	215.41700	37.29131	0.165	78	J142140.1+371728.7	82.5 ± 3.8	0.00	G066.68+68.44	7.1	0.90

[a] From maxBCG catalogue (Koester et al. 2007); RA and Dec. mark the cluster centre defined as the BCG location; z_{photo} is the photometric redshift; N_{200} is the maxBCG richness given by the number of E/S0 ridgeline galaxies more luminous than $0.4L_*$ within a scaled radius R_{200}.

[b] From redMaPPer catalogue (Rykoff et al. 2014); λ is the redMaPPer richness defined as the sum of the membership probabilities over all galaxies; offset relatives to the maxBCG centre.

[c] From the second *Planck* catalogue of SZ sources (Planck Collaboration XXVII 2016); the significance of the *Planck* detection and the offset to maxBCG centre.

cluster richness λ is the sum of the membership probabilities of the galaxies within a radius R_λ, which scales with richness as $R_\lambda = (\lambda/100)^{0.2}\,h^{-1}$ Mpc. The cluster photometric redshift is evaluated by simultaneously fitting all high probability cluster members with a single red-sequence model, with an accurate of 0.005–0.01. The cluster centre is on the central galaxy chosen with a probabilistic approach that weights not just galaxy luminosity, but also local galaxy density, as well as the consistency to cluster redshift. The completeness is $\gtrsim$99 per cent at $\lambda > 30$ and $z < 0.3$, and the purity is >95 per cent at all richness and redshift. The redMaPPer cluster mass has been calibrated by Simet et al. (2017) with the $M_{200m}-\lambda$ relation. We convert M_{200m} to M_{500c} assuming a cluster mass distribution of Navarro-Frenk-White (NFW; Navarro, Frenk & White 1997) profile with a typical concentration of 6 (e.g. Buote et al. 2007) and a median $z = 0.23$ for the sample. The ratio of M_{200m}/M_{500c} is 1.67.

Compared with maxBCG, some main improvements of redMaPPer include (i) using multicolour filter rather than single colour; (ii) the aperture used to estimate cluster richness is scaled with richness rather than the fixed scaled radius; (iii) the cluster centre is

Table 2. X-ray observations of the richest maxBCG clusters.

Cluster	$N_{\rm H}$ 10^{20} cm^{-2}	*Chandra* id	*Chandra* time ks	*XMM–Newton* id	*XMM–Newton* time ks
A2142	4.36	(1196 1228 5005 7692) (15186 16564 16565)	(231.2)	(0111870101 0111870401) 0674560201	51.5, 52.9, 41.1/108.7
J150	2.59	–	–	0742510101* 0761590101*	7.9, 7.8, 5.7/43.6
A1689	1.98	540 1663 5004 6930 7289 7701	196.7/199.8	0093030101	34.4, 35.3, 26.3/39.8
A1443	2.41	11762 16279	25.5/26.1	–	–
A781	1.75	(534 15128)	(45.6)	0150620201 0401170101 0401170201	73.1, 74.8, 58.2/105.2
A1986	3.18	17474	5.0/5.1	–	–
A1882	3.76	12904 12905 12906 12907 12908 12909 12910 12911 12912 17149 17150 17151 17671	392.2/397.9	0145480101 0762870501	30.2, 30.9, 20.0/45.6
A1758	1.06	(2213) 7710 13997 15538 15540	154.7/215.8	(0142860201)	(57.2)
A1760	1.87	17159	7.0/7.1	–	–
A1622	1.19	11763 (17154)	12.9/27.1	–	–
A750	3.66	924 7699	34.5/34.9	0605000901 0673850901	21.0, 21.2, 9.7/45.7
A1682	1.07	3244 11725	29.4/30.0	–	–
A1246	1.66	11770	5.0/5.1	–	–
A1961	1.21	11764	6.9/7.0	–	–
A2034	1.62	(2204 7695 12885 12886 13192 13193)	(259.1)	0149880101 0303930101 0303930201	18.8, 19.5, 11.1/75.4
A655	4.39	15159	7.9/8.1	–	–
A1914	1.10	(542) 3593 12197 12892 12893 12894	38.6/47.2	0112230201	19.1, 20.3, 8.8/25.8
Z5247	1.73	539 11727	29.3/30.5	(0673851101 0673852101)	(46.1)
A657	3.38	–	–	0742510401*	14.9, 15.7, 8.4/18.0
J229	6.09	–	–	0201902001 0201902101 0761590301*	76.0, 76.6, 58.4/83.8
A1423	1.93	538 11724	35.5/36.0	–	–
A801	4.37	11767	6.6/6.7	–	–
A773	1.34	533 3588 5006 13592 13594 13593 13591	60.2/ 61.3	0084230601	12.5, 14.3, 14.2/25.2
A1576	1.12	7938 15127	43.4/45.0	(0402250101 0402250601)	(28.8)
A2631	3.96	3248 11728	26.0/26.3	(0042341301)	(14.0)
A1703	1.39	15123 (16126)	29.1/78.8	(0653530101 0653530201 0653530301) (0653530401 0653530501 0653530601) (0653530701 0653530801 0653530901)	(135.5)
A2219	1.87	(896) 7892 13988 14431 14451 14355 14356	152.5/197.4	(0112231801 0112231901) 0605000501	12.0, 12.8, 6.8/55.1
A1319	2.20	11760 17153	23.8/24.1	–	–
A1560	2.63	11761 (17155)	12.1/20.8	(0404120101)	(31.9)
J175	4.51	–	–	0655380101	4.2, 4.3, 3.1/13.9
J249	1.94	–	–	0761590401*	20.2, 20.7, 13.7/23.0
A1201	1.66	(4216 7697) 9616	47.4/93.2	(0500760101)	(51.8)
A2009	3.89	10438	19.9/20.1	(0673851001 0673852201) 0693011001 (0693011301)	10.6, 12.3, 6.1/60.2
A2111	2.00	544 11726	31.2/31.6	(0673850601 0673852301)	(40.5)
A815	1.98	–	–	0742510701*	16.1, 16.6, 12.5/20.0
Z1450	1.19	–	–	0761590601*	19.6, 19.9, 6.3/23.0
A1765	2.33	–	–	0761590701*	21.7, 21.4, 15.6/ 24.0
A1902	1.02	16151	5.0/5.1	–	–

Note. The total Galactic column density of Hydrogen is from Willingale et al. (2013). The obsids and their total exposure time in brackets are not included in the data analysis of this work. The exposure time is the clean/total time for *Chandra* or clean MOS1, MOS2, pn/total time for *XMM–Newton*. The obsids with * are clusters from our *XMM* project (proposal id 74251 and 76159, PI: M. Sun).

on a weighted position rather than simply the BCG. These updates reduce the scatter at fixed richness (e.g. Rozo et al. 2009b).

2.3 PSZ2 catalogue

The second *Planck* Catalogue of Sunyaev–Zel'dovich Sources (PSZ2; Planck Collaboration XXVII 2016) exploits the 29 month full-mission data. It contains 1653 candidate clusters with a signal-to-noise ratio (SNR) above 4.5 and distributes across 83.6 per cent of the sky. Among them, more than 1203 are confirmed to be clusters with identified counterparts in external optical or X-ray samples or by dedicated follow-ups, and with a purity larger than 83 per cent. The median redshift is $z \sim 0.2$ and the farthest clusters are at $z \lesssim 1.0$.

3 DATA ANALYSIS

Table 2 provides the detail of observations from *Chandra* and *XMM–Newton* for this sample. We present here the data reduction procedure, the derivation of X-ray properties, and how we deal with complicated multiple cluster systems.

3.1 *Chandra*

We process the *Chandra* ACIS data with the *Chandra* Interactive Analysis of Observation (CIAO, version 4.9) and calibration data base (CALDB, version 4.7.3), following the procedures in Morandi et al. (2015). We reprocess the level-1 event files using `acis_process_events` tool to check for the presence of cosmic ray background events, correct for spatial gain variations due to charge transfer inefficiency (CTI), and recompute the event grades. We then filter the data to include the standard events grades 0, 2, 3, 4, and 6 only. Most observations were taken in very faint (VFAINT) mode, and in this case we applied VFAINT cleaning to both the cluster and stowed background observations. The light curve is then created with `dmextract` and filtered with `deflare` to exclude intervals of deviating more than 3σ of the mean value. Then we filter the ACIS event files in 0.3–12 keV to obtain a level-2 event file. Point sources and extended substructures are detected and removed using `wavedetect`, which provides candidate point sources, and the result is then checked through visual inspection. We produce the X-ray images from the level-2 event file and then create an exposure-corrected image from a set of observations using `merge_obs` (Fig. C2). We then measure the surface brightness radial profile S_X from the exposure-corrected images. We applied a direct subtraction of the cosmic X-ray (CXB) + particle + readout artefact backgrounds. For the particle background modelling, we use the stowed background scaled with the count rate in the 9.5–12 keV band, where the *Chandra* effective area is negligible and the flux is dominated by the particle background. In order to measure the CXB, we considered the region where the CXB is more dominant than the cluster emission, which can be determined from the flattened portion at the outer radial profile.

The spectra and response files of ACIS (0.7–7 keV) are extracted using SPECEXTRACT and fitted with XSPEC package (Arnaud 1996, version 12.9.1). We adopt the APEC emissivity model (Foster et al. 2012) to fit the on-cluster emission and the ATOMDB (version 3.0.8) of atomic data, the solar abundance tables are adopted from Asplund et al. (2009) and fixed to $0.3\,Z_\odot$, the redshift is fixed to the optical spectroscopy redshift of BCG. We apply the Tuebingen–Boulder absorption model (`TBABS`) for X-ray absorption by the interstellar medium (ISM), with fixed hydrogen column density N_H to the Galactic value from the NHtot tool (Willingale et al. 2013). The off-cluster background spectra are extracted from cluster emission insignificant regions ($\gtrsim R_{100}$) of the same exposure.

3.2 *XMM–Newton*

We reduce the *XMM–Newton* MOS and pn data using the Extended Source Analysis Software (ESAS; Kuntz & Snowden 2008; Snowden et al. 2008), as integrated into the *XMM–Newton* Science Analysis System (SAS, version 15.0.0) with the associated Current Calibration Files (CCF), following the procedures in Ge et al. (2016). We reproduce the raw event files from MOS and pn CCDs using tasks `emchain` and `epchain`, respectively. The solar soft proton (SP) flares are filtered out with `mos-filter` and `pn-filter` through the light-curve screening to obtain the clean event files. The MOS CCDs that are damaged or in the anomalous state are excluded in downstream processing. The point sources are detected by `cheese` and checked with visual inspection and then excluded. We use `mos-spectra` and `pn-spectra` to produce event images and exposure maps, as well as to extract spectra and response files. The instrumental background images and spectra are modelled with `mos_back` and `pn_back`. We combine the event images, background images, and exposure maps from MOS and pn with `comb`, and combine the images from multiple observations with `merge_comp_xmm`. We use `adapt` to produce the final background subtracted, exposure-corrected, and smoothed images (Fig. C2). The surface brightness profiles are extracted from exposure-corrected images of combined MOS1/MOS2 and pn separately. The derived X-ray properties are evaluated with error weight means of MOS and pn.

The spectra of MOS (0.3–11.0 keV) and pn (0.4–11.0 keV) are fitted jointly. The on-cluster spectra are also fitted with APEC model. The same solar abundance table (also fixed to $0.3\,Z_\odot$), ATOMDB version and the Galactic absorption model as in the *Chandra* analysis are used. The background spectra consist of mainly four components: CXB, quiescent particle background (QPB), residual SP, and solar wind charge exchange (SWCX). They are fit simultaneously with the on-cluster emission. The CXB is modelled with three components: an unabsorbed thermal emission ($E \sim 0.1$ keV) from the local hot bubble or heliosphere; an absorbed thermal emission ($E \sim 0.25$ keV) from the Galactic halo and/or intergalactic medium; and an absorbed power-law emission ($\Gamma \sim 1.46$) from an unresolved background of cosmological sources. An off-cluster RASS spectrum is also extracted from a 1 deg to 2 deg annulus surrounding the cluster and joint fit with other spectra to constrain the contribution of CXB. The QPB originates from the interaction of cosmic rays with detectors. The QPB continuum is subtracted as background spectra in XSPEC, while its bright instrumental fluorescent lines vary from observation to observation. Thus, they are not included in the QPB model spectra and are individually fit by Gaussian models. The residual SP may still exist after the light-curve screening. As they are not X-ray photons and not folded through the instrumental effective area, SP is modelled by a power law with diagonal response matrices supplied with the ESAS calibration files. The SWCX process may produce additional emission lines in the observed spectra, and they are modelled with the Gaussian components.

3.3 T_X, L_X, and Y_X

The spectroscopic X-ray temperature T_X is measured in 0.15–0.75 R_{500}. The inner boundary of $0.15\,R_{500}$ is chosen to exclude the central cool core (CC) and the outer boundary of $0.75\,R_{500}$ is limited by the quality of the spectroscopic data. The X-ray peak or centroid of a cluster is the position where the derivatives of the surface brightness variation along two orthogonal (e.g. X and Y) directions become zero. R_{500} is estimated from the $M-T_X$ relation (Sun et al. 2009) iteratively. In fact, the R_{500} is dependent on temperature with a power-law index of 0.55 and the R_{500} typically converges in three iterations.

The bolometric X-ray luminosity L_X is derived within R_{500}. We sum the count rates from surface brightness profiles in the 0.7–2 keV for *Chandra* (maximize the SNR of cluster emission and minimize the dependence of the cooling function on the temperature and metallicity) and 0.7–1.3 keV for *XMM–Newton* (maximize the SNR of cluster emission and minimize the contamination of SWCX lines below 0.7 keV and instrumental lines above 1.3 keV). Regions masked for point sources, chip gaps and bad pixels are added back in this process. We then use the best-fitting spectral model to convert the count rate to bolometric luminosity. For some clusters with shallow data, surface brightness profiles cannot be robustly constrained to R_{500}. We extrapolate the surface brightness profile using a power law measured from the profile at large radii. The typical slope is about -3, which corresponds to $\beta = 2/3$

for a β model and is typical for cluster density profiles around R_{500} (Morandi et al. 2015). The extrapolated correction factors ([data + extrapolation]/data) are always smaller than 1.5.

The X-ray Compton parameter $Y_{\rm X}$, $M_{\rm gas}(< R_{500}) \times T_{\rm X}$ (0.15–1.0 R_{500}), is also derived. We convert our $T_{\rm X}$ (0.15–0.75 R_{500}) to $T_{\rm X}$ (0.15–1.0 R_{500}) with the correction factor of 0.96 based on Morandi et al. (2015). The gas mass $M_{\rm gas}$ within R_{500} is summed from the gas density profile, which is derived from deprojecting the surface brightness profile. For clusters with surface brightness profiles not reaching R_{500}, we also extrapolate the gas density profiles using a power-law model measured from the profile at large radii, similar to what we did for luminosity extrapolation. The resulting $Y_{\rm X}$ value is only included in the analysis if the extrapolated correction factor of $M_{\rm gas}$ ([data + extrapolation]/data) is smaller than 1.5. Table 3 includes the results of X-ray properties.

3.4 Multiple cluster systems

Most clusters in the sample are mergers. Some clusters with multiple components, e.g. J150, A781, A750, and A1319, require special attention as the association of the X-ray emission with the optical cluster becomes ambiguous. We apply different strategies to deal with such systems.

For systems of one optical cluster corresponding to multiple X-ray clusters (e.g. J150, A750, A1319), we use two different methods to associate the X-ray emission with the optical cluster and study the scaling relations.

(i) Add. The X-ray properties are derived individually for each cluster after masking the R_{500} region of the nearby cluster. Then we derive L_X-weighted T_X, add L_X or Y_X together, and then assign to the corresponding optical counterpart from maxBCG or redMaPPer catalogue. For instance, A750 is a close pair composed of A750E and A750W (Fig. C2, Appendix C). Both maxBCG and redMaPPer catalogue mix them as one cluster. We add the L_X values from A750E and A750W, as well as their Y_X values. The T_X is from the L_X-weighted T_X of A750E and A750W, which should be similar to the emission-weighted.

(ii) Mix. The multiple subclusters are treated as one cluster. We centre the cluster on the peak of the X-ray brightest subcluster and derived the X-ray properties within R_{500} determined from the $M-N$ relation based on the optical catalogue, under the assumption of spherical symmetry. In the case of A750, we centre the cluster in A750E and do not mask out A750W.

We compare the results in Table 4 from these two methods and find that the L_X are comparable from add and mix, while the Y_X from add is smaller than the Y_X from mix. For clusters in the early merge stage (J150, A750, A1319, and J175), we use the add method. For clusters in the late merge stage (Z5247, A1560, and A815) where the individual R_{500} regions heavily overlap, we use the mix method. We emphasize that switching the method for multiple cluster systems in Table 4 does not affect any of our conclusions for the whole sample.

For systems of one X-ray cluster corresponding to multiple optical clusters (e.g. A781, A1760, A2631), we match the X-ray cluster with the richest optical cluster.

4 RESULTS

We first study the scaling relations between X-ray and optical properties as T_X-N, L_X-N, and Y_X-N relations. We then study the X-ray scaling relations as L_X-T_X and L_X-Y_X relations. Finally we classify the clusters using quantitative X-ray morphology parameters.

4.1 X-ray–optical scaling relations

As the T_X is a robust mass proxy with low scatter, the temperature-richness relation $P(T_X|N_{200})$ offers an insight into halo mass selection properties of the maxBCG algorithm. Fig. 1, left-hand panel shows the T_X-N_{200} relation for this sample, along with the expected relation from the X-ray $M-T_X$ relation and the weak-lensing $M-N_{200}$ relation using the multivariate scaling relation in Appendix A. While the richness threshold of 78 (dash–dotted line) corresponds to ∼5 keV, a wide temperature range of from ∼2 to ∼10 keV is observed from the data, with approximately half of the systems cooler than 5 keV. The temperature distribution is not symmetric around the expected relation (dashed line), but with more outliers towards lower temperatures. Indeed, 29 clusters in the sample are detected in the PSZ2 catalogue and the maxBCG clusters without *Planck* detections all have low temperatures (Fig. 1). Since the optical richness $N_{\rm gal}$ is a ‘projected measure’ with only moderate redshift resolution, the mapping of M to $N_{\rm gal}$ is skewed to high $N_{\rm gal}$ when surrounding filamentary projections are favourable. Explicit simulation studies (Cohn et al. 2007; Noh & Cohn 2011; Rozo et al. 2011) have shown that $P(N_{\rm gal}|M)$ is asymmetric, with a tail extending to high richness. While the overall amplitude of this asymmetry may be modest (e.g. involving 10 per cent of all haloes), it will be significantly larger when one selects only the highest richness systems from a large population. As we study only the richest 0.3 per cent of the maxBCG catalogue, the fraction of projected or ‘blended’ haloes (Cohn et al. 2007) could be boosted by a factor of ∼2. Fig. C1 and detail in Fig. C2 show that 14/38 clusters are in multisystems, which could potentially boost the richness estimate. In fact, 5/38 (∼13 per cent) maxBCG clusters are picked apart by redMaPPer. Moreover, the $P(T_X|\lambda)$ relation in Fig. 1 right-hand panel shows more low-mass clusters are identified by redMaPPer and the whole sample disperses along the mean relation with lower scatter. This demonstrates that the update algorithm of redMaPPer improves the richness estimate, which is less affected by the projection and with lower scatter. These improvements are mostly due to the multicolour filter (Rozo et al. 2009b).

Fig. 2 shows the L_X-N relation for maxBCG (left-hand panel) and redMaPPer (right-hand panel). The dashed line shows the expected L_X-N relation combined from the L_X-M relation and the $M-N_{200}$ or the $M-\lambda$ relation. As miscentring and substructure can also contribute to the mismatch problem, we attempt to classify clusters in this sample. The red dots (marked as ‘*M*’ in Table 3) are clusters with miscentring problems (the offset between X-ray peak and optical centre larger than 50 kpc), with most of them representing merging clusters (e.g. Hudson et al. 2010). The blue dots (marked as ‘*P*’ in Table 3) are clusters with projection problems (another nearby X-ray cluster with similar redshift in a 2 Mpc projected radius). The half-red half-blue dots (marked as ‘*M*, *P*’ in Table 3) are with both miscentring and projection problems. The grey dots (marked as ‘*R*’ in Table 3) are clusters without miscentring and projection problems. The L_X-N distribution is similar to the T_X-N distribution. Similarly, there are more outliers towards lower X-ray luminosity.

Fig. 3 shows the Y_X-N relation for maxBCG (left-hand panel) and redMaPPer (right-hand panel). The Y_X-N distribution is similar to the previous two distributions. The SZ signal $Y_{\rm SZ}$ is characterized by the integrated Compton parameter $y_{\rm c} = \int n_{\rm e} T_{\rm e} {\rm d}l$. As an X-ray analogue of the integrated Compton parameter, Y_X is closely related to $Y_{\rm SZ}$ with a linear relation (Rozo et al. 2014a). Therefore, Fig. 3 may be converted to the $Y_{\rm SZ}-N$ relation, which also recovers previous stacking results of lower observed SZ signal compared

Table 3. X-ray properties.

Cluster	RA	Dec.	$z_{\rm spec}$	T_X keV	L_X 10^{44} erg s^{-1}	Y_X 10^{14} M$_\odot$ keV	R_{500} Mpc	offset^N kpc	offset$^\lambda$ kpc	M_{500} 10^{14} M$_\odot$	M_{500}^N 10^{14} M$_\odot$	M_{500}^λ 10^{14} M$_\odot$	opt^N	opt$^\lambda$
A2142	239.58563	27.22967	0.091	8.20 ± 0.05 (X)	32.24 ± 0.06	11.44 ± 0.33	1.43	25.7	25.8	9.1	12.2	12.6	R	R
J150M	150.61145	20.51861	0.320	3.63 ± 0.48 (X)	3.24 ± 0.12	0.98 ± 0.15	0.81	1273.2	888.0	2.1	10.6	10.9	M, P	M, P
J150E	150.68539	20.49089	0.309	2.44 ± 0.35 (X)	1.08 ± 0.05	0.28 ± 0.05	0.65			1.1				
J150W	150.55545	20.53728	0.320	3.11 ± 0.71 (X)	1.00 ± 0.09	0.55 ± 0.14	0.74			1.6				
A1689	197.87247	− 1.34129	0.183	10.70 ± 0.26 (C)	41.73 ± 1.06	13.89 ± 0.63	1.58	5.4	5.8	13.4	10.0	12.2	R	R
				8.01 ± 0.13 (X)	33.59 ± 0.11	8.86 ± 0.37	1.34			8.3				
A1443	180.30410	23.10630	0.264	8.78 ± 0.60 (C)	19.97 ± 2.08	11.58 ± 1.02	1.35	216.3	112.2	9.3	8.9	9.1	M	M
A781	140.11223	30.49477	0.299	5.55 ± 0.17 (X)	10.03 ± 0.08	3.88 ± 0.22	1.03	67.3	179.0	4.3	8.0	1.9	M, P	M, P
A781M	140.22101	30.46535	0.293	3.12 ± 0.14 (X)	0.75 ± 0.06	2.32 ± 0.15	0.75		71.7	1.7		8.8		M, P
A781E	140.29418	30.46433	0.429	3.34 ± 0.34 (X)	3.26 ± 0.09	0.83 ± 0.10	0.72		84.5	1.7		1.9		M, P
A781W	139.89517	30.53303	0.427	3.24 ± 0.24 (X)	2.81 ± 0.06	0.62 ± 0.06	0.71			1.6				
A1986	223.28780	21.89560	0.117	3.65 ± 0.47 (C)	2.16 ± 0.38	1.09 ± 0.17	0.90	639.5	23.5	2.4	7.6	3.9	M	R
A1882E	213.77818	− 0.49216	0.139	3.35 ± 0.40 (C)	0.85 ± 0.10		0.85	13.0		2.0	7.3		P	
				2.44 ± 0.17 (X)	0.68 ± 0.10		0.72			1.2				
A1882W	213.60034	− 0.37877	0.138	1.91 ± 0.16 (X)	0.33 ± 0.01		0.63		10.0	0.8		6.8		P
A1882M	213.74050	− 0.34915	0.139	1.37 ± 0.08 (X)	0.27 ± 0.01		0.52			0.5				
A1882N	213.53597	− 0.27103	0.140	1.36 ± 0.14 (X)	0.11 ± 0.01		0.52			0.5				
A1758N	203.19958	50.53972	0.279	10.15 ± 0.43 (C)	17.01 ± 2.31	11.76 ± 0.77	1.45	491.0	492.0	11.7	7.2	11.1	M, P	M, P
A1758S	203.12317	50.40931	0.273	5.87 ± 0.24 (C)	11.61 ± 0.67	4.40 ± 0.29	1.08		203.7	4.7		2.7		M, P
A1760	203.53705	20.23874	0.173	6.71 ± 0.92 (C)	7.10 ± 1.02	5.41 ± 0.84	1.23	138.4	1001.5	6.2	7.1	4.3	M	M
A1622	192.42163	49.87525	0.283	4.70 ± 0.89 (C)	2.84 ± 0.14	1.31 ± 0.36	0.95	477.3	604.1	3.3	6.7	8.9	M, P	M, P
A750E	137.30282	10.97472	0.176	5.02 ± 0.26 (C)	7.44 ± 0.56	2.52 ± 0.19	1.05		23.7	3.9		13.1		P
				4.76 ± 0.16 (X)	5.83 ± 0.06	2.04 ± 0.12	1.02			3.6				
A750W	137.21662	11.02747	0.164	3.36 ± 0.18 (X)	1.96 ± 0.04	0.86 ± 0.07	0.84	13.0		2.0	6.7		P	
A1682	196.70864	46.55945	0.218	7.90 ± 0.67 (C)	10.42 ± 0.52	5.38 ± 0.67	1.31	159.7	3.6	8.0	6.7	8.3	M	R
A1246	170.98878	21.48200	0.193	6.23 ± 0.66 (C)	9.77 ± 0.49	3.93 ± 0.60	1.17	312.3	68.4	5.5	6.5	7.5	M	M
A1961	221.13280	31.22670	0.233	4.12 ± 1.28 (C)	3.63 ± 0.48	1.35 ± 0.44	0.91	1005.8	3.8	2.7	6.3	7.6	M	R
A2034	227.55129	33.50973	0.112	5.98 ± 0.17 (X)	8.82 ± 0.06	3.96 ± 0.20	1.19	174.0	174.0	5.3	6.2	6.3	M	M
A655	126.37410	47.13130	0.129	4.31 ± 0.35 (C)	3.83 ± 0.55	1.91 ± 0.20	0.98	25.0	25.4	3.1	6.2	9.1	R	R
A1914	216.50380	37.82557	0.170	9.33 ± 0.66 (C)	33.10 ± 1.57	11.41 ± 0.94	1.47	174.9	174.9	10.8	6.1	6.6	M	M
				7.84 ± 0.19 (X)	29.39 ± 0.14	8.71 ± 0.39	1.34			8.1				
Z5247	188.57520	9.77170	0.231	5.91 ± 0.64 (C)	5.85 ± 0.66	3.68 ± 0.71	1.11	79.1	35.3	4.9	6.1	5.8	M, P	P
A657	125.83066	15.96316	0.153	2.89 ± 0.17 (X)	1.28 ± 0.02	0.53 ± 0.05	0.78	5.2	5.2	1.6	5.9	4.9	R	R
J229	229.33991	− 0.71679	0.114	2.50 ± 0.12 (X)	0.61 ± 0.01	0.29 ± 0.03	0.73	178.3	12.5	1.3	5.8	3.1	M, P	P
A2051	229.18317	− 0.97167	0.120	3.78 ± 0.13 (X)	2.28 ± 0.03	1.18 ± 0.08	0.92			2.5				
A2051S	229.24147	− 1.11056	0.117	2.45 ± 0.26 (X)	0.27 ± 0.01		0.73	4.6	233.6	1.2	1.1	2.8	R	M
A2051N	229.10108	− 0.83036	0.376	3.43 ± 0.52 (X)	1.70 ± 0.13	0.68 ± 0.14	0.76		6.2	1.8		1.1		

Table 3 – *continued*

Cluster	RA	Dec.	z_{spec}	T_X keV	L_{X} 10^{44} erg s^{-1}	Y_X 10^{14} M$_{\odot}$ keV	R_{500} Mpc	offset^N kpc	offset$^{\lambda}$ kpc	M_{500} 10^{14} M$_{\odot}$	M_{500}^N 10^{14} M$_{\odot}$	M_{500}^{λ} 10^{14} M$_{\odot}$	opt^N	opt$^{\lambda}$
A2050	229.07404	0.08937	0.121	4.46 ± 0.12 (X)	4.05 ± 0.04	1.69 ± 0.10	1.01	5.9	5.8	3.3	4.4	4.6	R	R
A1423	179.32222	33.61106	0.214	6.99 ± 0.48 (C)	11.60 ± 0.58	5.68 ± 0.64	1.23	1.6	93.5	6.5	5.7	5.2	R	M
A801	142.01788	20.52813	0.192	4.31 ± 0.64 (C)	4.95 ± 0.67	1.74 ± 0.29	0.95	22.2	22.3	3.0	5.7	5.8	R	R
A773	139.47280	51.72909	0.217	8.47 ± 0.52 (C)	18.68 ± 0.19	8.79 ± 0.69	1.36	26.0	25.3	9.0	5.6	11.1	R	R
				6.53 ± 0.17 (X)	14.68 ± 0.11	5.31 ± 0.27	1.18			5.8				
A1576	189.24296	63.18712	0.301	7.99 ± 0.55 (C)	14.44 ± 2.17	7.95 ± 0.87	1.26	29.4	9.2	7.8	5.5	6.6	R	R
A2631	354.41084	0.26799	0.277	7.83 ± 0.54 (C)	20.24 ± 1.01	8.18 ± 0.96	1.26	87.9	87.8	7.6	5.5	8.6	M	M
A1703	198.77059	51.81964	0.282	9.63 ± 0.75 (C)	21.80 ± 1.09	9.81 ± 1.19	1.41	36.9	35.8	10.7	5.4	11.0	R	R
A2219	250.08197	46.71113	0.225	12.59 ± 0.29 (C)	61.02 ± 3.05	26.43 ± 0.99	1.69	6.7	7.3	17.2	5.3	15.6	R	R
A1319M	174.05510	40.04070	0.290	2.95 ± 0.59 (C)	2.00 ± 0.39	0.57 ± 0.13	0.73	41.2	41.2	1.5	5.3	11.3	P	P
A1319NW	173.99769	40.08607	0.295	2.95 ± 0.44 (C)	2.59 ± 0.49		0.73			1.5				
A1560	188.47084	15.19453	0.284	5.18 ± 0.77 (C)	4.95 ± 0.25	2.15 ± 0.48	1.00	1.8	308.8	3.8	5.2	9.6	P	M, P
J175N	175.56923	10.45006	0.117	1.66 ± 0.32 (X)	0.20 ± 0.01		0.59	1205.2	47.3	0.6	5.2	1.0	M, P	P
J175S	175.52926	10.14652	0.119	1.42 ± 0.40 (X)	0.14 ± 0.02		0.54		24.8	0.5	–	1.8		P
J249SW	249.90682	47.05049	0.227	3.51 ± 0.28 (X)	1.76 ± 0.04	0.71 ± 0.07	0.83	338.3	33.9	2.1	5.2	3.4	M, P	P
J249NE	250.03190	47.16119	0.224	2.67 ± 0.64 (X)	0.77 ± 0.03	0.30 ± 0.08	0.72			1.3	–	–		
A1201	168.22715	13.43506	0.168	6.12 ± 0.30 (C)	7.54 ± 0.38	3.83 ± 0.26	1.17	8.1	452.8	5.4	5.1	4.6	R	M
A2009	225.08160	21.36920	0.153	6.30 ± 0.37 (C)	14.62 ± 0.84	4.21 ± 0.31	1.20	3.2	3.3	5.7	5.1	5.7	R	R
				5.91 ± 0.21 (X)	14.04 ± 0.10	3.96 ± 0.24	1.16			5.1				
A2111	234.92407	34.41675	0.228	8.57 ± 0.59 (C)	12.87 ± 0.73	8.60 ± 0.08	1.36	114.4	114.4	9.1	5.0	10.9	M	M
A815	143.03589	29.06782	0.300	2.58 ± 0.30 (X)	1.47 ± 0.05		0.68	0.2	0.0	1.2	4.9	5.7	P	P
Z1450	223.05214	37.88133	0.282	2.30 ± 0.40 (X)	1.48 ± 0.05	0.34 ± 0.07	0.64	60.2	28.6	1.0	4.9	6.5	M, P	P
A1765	204.20616	10.44110	0.155	2.22 ± 0.16 (X)	0.49 ± 0.02	0.20 ± 0.02	0.67	11.1	11.1	1.0	4.8	5.0	P	P
A1902	215.41900	37.29160	0.163	7.21 ± 1.24 (C)	9.72 ± 1.31	4.21 ± 0.79	1.28	16.3	16.3	7.1	4.8	5.0	R	R

Note. The individual cluster name is marked the same as in Fig. C2. RA and Dec. are the X-ray peak position. z_{spec} is the spectroscope redshift of BCG. T_X is the X-ray temperature within 0.15–0.75 R_{500}. L_X is the bolometric luminosity within R_{500}. Y_X is the X-ray Compton parameter, $Y_X = M_{\mathrm{gas}}\,(< R_{500}) \times T_X\,(0.15\text{–}1.0\,R_{500})$, and is not listed if the extrapolation factor for M_{gas} is larger than 1.5. R_{500} and M_{500} are estimated from the $M-T_X$ relation (Sun et al. 2009). offset^N or offset$^{\lambda}$ is the offset between X-ray peak and maxBCG or redMaPPer centre, respectively. M_{500}^N and M_{500}^{λ} are from $M-N$ relation (Rozo et al. 2009a) for maxBCG and $M-\lambda$ relation (Simet et al. 2017, with erratum in Simet et al. 2018 applied below) for redMaPPer, respectively. We apply $\sim$−10 per cent and $\sim$−3 per cent correction for the maxBCG and redMaPPer masses because we cite $\tilde{a}_{m|n}$ from $M-N$ relation and correct to $a_{m|n}$ with the relation of $\tilde{a}_{m|n} = a_{m|n} + \frac{1}{2}\sigma^2_{m|n}$ (Appendix A). The label of '*M*' means clusters with miscentring problems (the offset between X-ray peak and optical centre larger than 50 kpc), '*P*' means clusters with projection problems (another X-ray cluster in a 2 Mpc radius), and '*R*' means relaxed cluster without both miscentring and projection problems.

Table 4. Assignment table for multiple cluster system.

Cluster	Subclusters	maxBCG T_X, L_X, Y_X	redMaPPer T_X, L_X, Y_X
J150	J150E, J150M, J150W (add)	$3.29 \pm 0.33, 5.32 \pm 0.16, 1.81 \pm 0.21$	$3.29 \pm 0.33, 5.32 \pm 0.16, 1.81 \pm 0.21$
	J150E, J150M, J150W (mix)	$2.89 \pm 0.28, 5.54 \pm 0.15, 2.95 \pm 0.35$	$2.94 \pm 0.29, 5.47 \pm 0.14, 2.88 \pm 0.35$
A750	A750E, A750W (add)	$4.41 \pm 0.13, 7.79 \pm 0.07, 2.90 \pm 0.68$	$4.41 \pm 0.13, 7.79 \pm 0.07, 2.90 \pm 0.68$
	A750E, A750W (mix)	$4.01 \pm 0.11, 7.16 \pm 0.07, 3.88 \pm 0.21$	$3.66 \pm 0.12, 7.35 \pm 0.07, 4.09 \pm 0.24$
Z5247	Z5247NE, Z5247SW (mix)	$6.03 \pm 1.23, 7.52 \pm 0.74, 5.58 \pm 1.79$	$6.03 \pm 1.23, 7.52 \pm 0.74, 5.58 \pm 1.79$
A1319	A1319M, A1319NW, A1319SW (add)	$2.95 \pm 0.36, 4.59 \pm 0.63$, –	$2.95 \pm 0.36, 4.59 \pm 0.63$, –
A1560	A1560NW, A1560SE (mix)	$5.31 \pm 0.71, 4.33 \pm 0.33, 2.24 \pm 0.42$	$5.31 \pm 0.71, 4.33 \pm 0.33, 2.24 \pm 0.42$
J175	J175N, J175S (add)	$1.56 \pm 0.25, 0.34 \pm 0.02$, –	–
A815	A815N, A815S (mix)	$2.63 \pm 0.30, 1.57 \pm 0.06$, –	$2.58 \pm 0.30, 1.47 \pm 0.05$, –

Note. For systems of one optical cluster corresponding to multiple X-ray clusters, we use two different methods to assign the X-ray properties for clusters in different merge stages. The add method adds L_X or Y_X of multiple clusters, the T_X is from the L_X-weighted T_X of individual cluster. The mix method treats the multiple clusters as one cluster and uses the same routines for the single cluster, then derives the X-ray properties centring on the brightest X-ray peak within R_{500} from M–N relations.

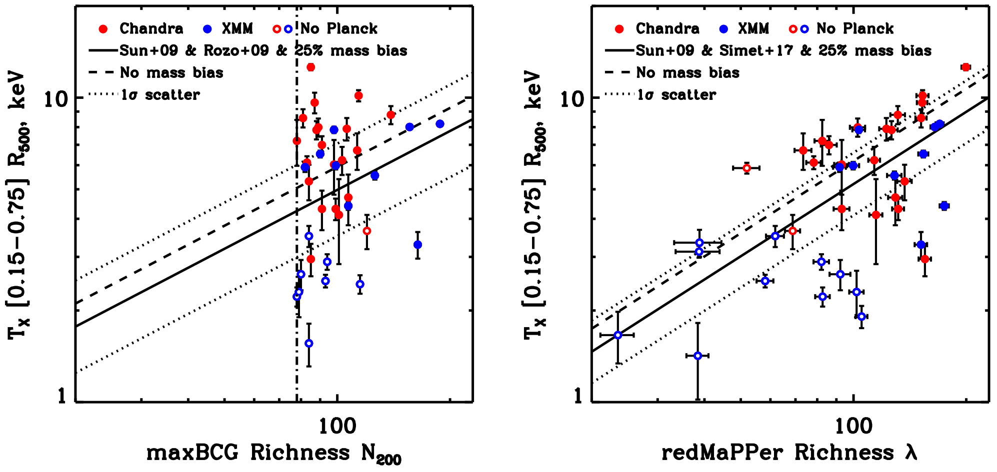

Figure 1. *Left*: The T_X–N_{200} relation for the richest maxBCG sample, with T_X measured from *Chandra* (red) or *XMM–Newton* (blue). The empty circles are clusters without *Planck* SZ detection. The dashed line are the expected T_X–N_{200} relations from M–T_X relation (Sun et al. 2009) and weak-lensing M–N_{200} relation (Rozo et al. 2009a) on the full maxBCG sample (see Appendix A for the detail and bias correction). The solid line shows 25 per cent hydrostatic bias (the X-ray mass is 25 per cent lower than the weak-lensing mass), and the dotted line shows expected 1σ scatter from multivariate scaling relation. The dash–dotted line shows the richness threshold of $N_{200} = 78$. *Right*: The T_X–λ relation for the same sample. The dashed line is from X-ray M–T_X relations and weak-lensing M–λ relation (Simet et al. 2017). There are more *XMM* outliers than *Chandra*, which reflects the fact that many *XMM* clusters were selected for observations with non-X-ray reasons, including seven clusters in our new *XMM* project. Moreover, most of these low T_X outliers have no *Planck* SZ detection. We also note that the highest T_X system, A2219, shifts from $N_{200} = 85$ to $\lambda = 199.5$, more in line with the overall population.

with the model prediction (e.g. Planck Collaboration XII 2011; Sehgal et al. 2013; Saro et al. 2017). Moreover, we convert the *Planck* mean stacking $Y_{\rm SZ}$ (Planck Collaboration XII 2011) in the two richest maxBCG bins to the expected Y_X, using the $D_{\rm A}^2 Y_{\rm SZ}$–CY_X relation (Rozo et al. 2014a; more detail in Appendix A). We also evaluate the observed mean Y_X (with *XMM–Newton* based Y_X converted to *Chandra* using the cross-calibration in Appendix B) in the same two bins. The expected Y_X and observed Y_X are shown as orange and green bowties in Fig. 3. They are consistent with each other, and both are lower than the model prediction from the combination of X-ray mass proxy with optical weak-lensing mass calibration, which is shown as a dashed line.

4.2 X-ray scaling relation

X-ray imaging is good at identifying individual haloes. Each extended X-ray source would map cleanly on to a single halo, while optical clusters offer a dirtier mapping to haloes mainly due to projection effect. So long as the hot gas properties of the involved haloes are not influenced by the projection effect in optical, we expected that the low-mass haloes associated with blended (projection-dominated) optical systems will have X-ray properties comparable to those of the overall massive halo population. If M is lower so will be L_X, T_X, and Y_X. Figs 4 and 5 show the L_X–T_X relation and the L_X–Y_X relation for this sample, respectively. Established relations from Vikhlinin et al. (2009a) (*Chandra* calibration) and Pratt et al. (2009) (*XMM–Newton* calibration) are also shown. With the cross-calibration derived in Appendix B, we also rescale all *Chandra* properties to *XMM–Newton*. Indeed, we find that these optically selected clusters follow the ‘normal’ L_X–T_X and L_X–Y_X relations calibrated from X-ray selected samples, ‘normal’ means the behaviour of the underlying true population of massive haloes.

4.3 X-ray morphology parameters

Besides the dynamical criterion based on the BCG/X-ray offset and cluster environment as presented in Table 3, we also apply several quantitative X-ray morphology parameters in Table 5: the centroid shift (w; Mohr, Fabricant & Geller 1993), the surface brightness concentration ($c_{\rm SB}$; Santos et al. 2008), and the symmetry-

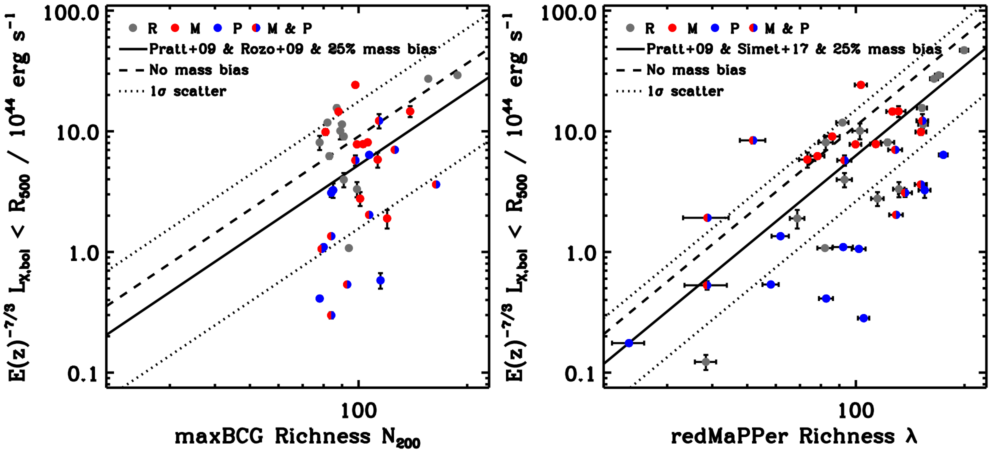

Figure 2. *Left*: The bolometric L_X–N_{200} relation for the sample. Red dots are clusters with miscentring problems (the offset between the X-ray peak and the optical centre larger than 50 kpc), with most of them are merging clusters. Blue dots are clusters with projection problems (another X-ray cluster within 2 Mpc, which corresponds to 9 arcmin at median redshift = 0.23 for the maxBCG sample, comparable to *Planck*'s resolution). Half-red half-blue dots are clusters with both miscentring and projection problems. Grey dots are relaxed clusters without miscentring and projection problems. The dashed line shows the expected L_X–N_{200} relation from the X-ray L_X–M relation (Pratt et al. 2009) and the weak-lensing M–N_{200} relation (Rozo et al. 2009a) on the full maxBCG sample (Appendix A). The solid line and the dotted line show 25 per cent hydrostatic bias and expected 1σ scatter, respectively. *Right*: L_X–λ relation for the sample. The L_X–M relation (Pratt et al. 2009) and the M–λ relation (Simet et al. 2017) are combined. These two plots show the roles of miscentring and projection on the scatter of the optical-richness relations, while the cluster mass bias can also reduce the discrepancy.

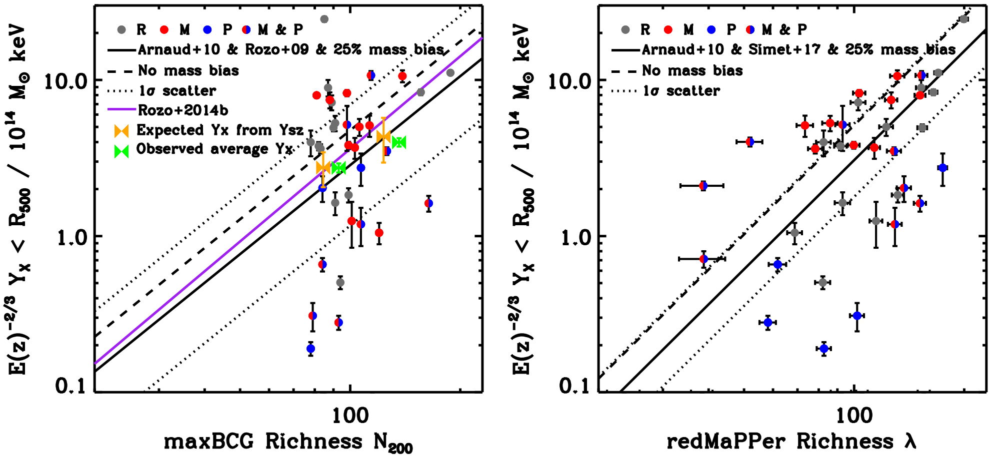

Figure 3. The relations between Y_X and the optical richness for the sample, with the same symbol set as in Fig. 2. Orange bowties are the expected Y_X, which is inferred from the *Planck* mean staking Y_{SZ} in the two richest maxBCG bins (Planck Collaboration XII 2011). The $D_A^2 Y_{SZ}$–CY_X relation from Rozo et al. (2014a) is assumed. Green bowties are the observed mean Y_X in the same two bins. The bowties are misplaced to the average richness of each bin to avoid overlapping. The expected Y_X and observed Y_X are consistent with each other, and both of them are lower than the model prediction shown by the dashed line. The black dashed line combines the M–Y_X relation (Arnaud et al. 2010) and the M–N_{200} relation (Rozo et al. 2009a) on the full maxBCG sample (Appendix A) for the *left*, and the M–Y_X relation (Arnaud et al. 2010) and the M–λ relation (Simet et al. 2017) for the *right*. We also plot the predicted Y_X–N_{200} from Rozo et al. (2014b) as a purple line in the left-hand panel. Most relaxed clusters are more close to the expected scaling relation.

peakiness-alignment (SPA; Mantz et al. 2015). Generally, the w, s, and a measure the bulk asymmetry, while c_{SB} and p judge the CC. The distribution of these parameters are shown in Fig. 6.

5 DISCUSSION

5.1 Scaling relation from optical, X-ray, and SZ

There is mismatch or discrepancy between optical and SZ: the expected SZ signals from the model are higher than the observed ones for stacked optically selected clusters (e.g. Planck Collaboration XII 2011; Sehgal et al. 2013; Saro et al. 2017). However, the scaling relations between SZ and X-ray are consistent with each other. For example, Planck Collaboration X (2011) extracts the SZ signal individually for the Meta-catalogue of X-ray detected Clusters of galaxies and find that the SZ measurements and the X-ray based predictions are in excellent agreement. Similar results are found based on both X-ray selected clusters (Melin et al. 2011; Planck Collaboration XII 2011; Sehgal et al. 2013) and SZ selected clusters (Andersson et al. 2011; Planck Collaboration IX 2011; Planck Collaboration XI 2011). Moreover even for optically selected clusters, Planck Collaboration XII (2011) compares the stacked SZ signal with the stacked *ROSAT* X-ray luminosity (Rykoff et al. 2008) around maxBCG clusters in the same richness bins and finds the mean SZ signal and mean X-ray luminosity are consistent with the model predictions.

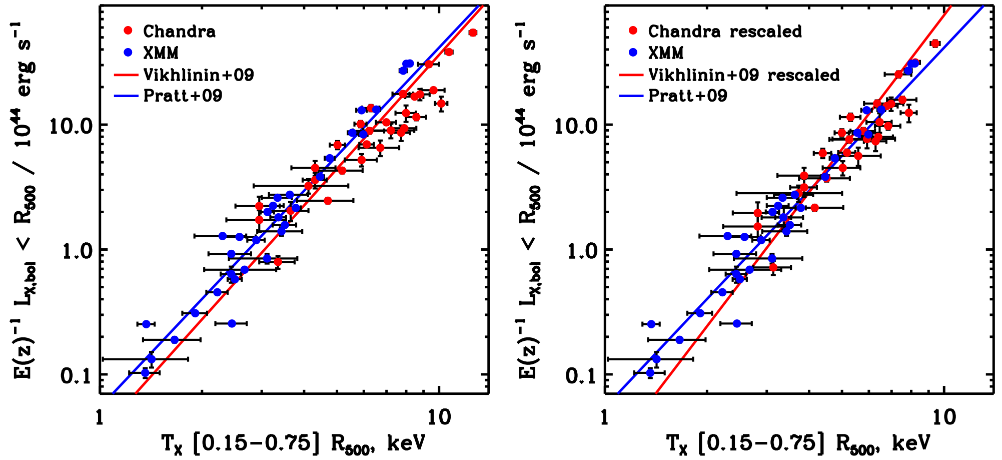

Figure 4. *Left*: The L_X-T_X relation for individual clusters from Table 3 (red: *Chandra*, blue: *XMM–Newton*). The red line is the L_X-T_X relation (Vikhlinin et al. 2009a) based on *Chandra* calibration, while the blue line (Pratt et al. 2009) is based on *XMM–Newton* calibration. *Right*: *Chandra* data (L_X and T_X) and L_X-T_X relation are rescaled to *XMM–Newton* with the in-house cross-calibration in Appendix B. These clusters follow the normal L_X-T_X relation with less scatters than those in Fig. 2.

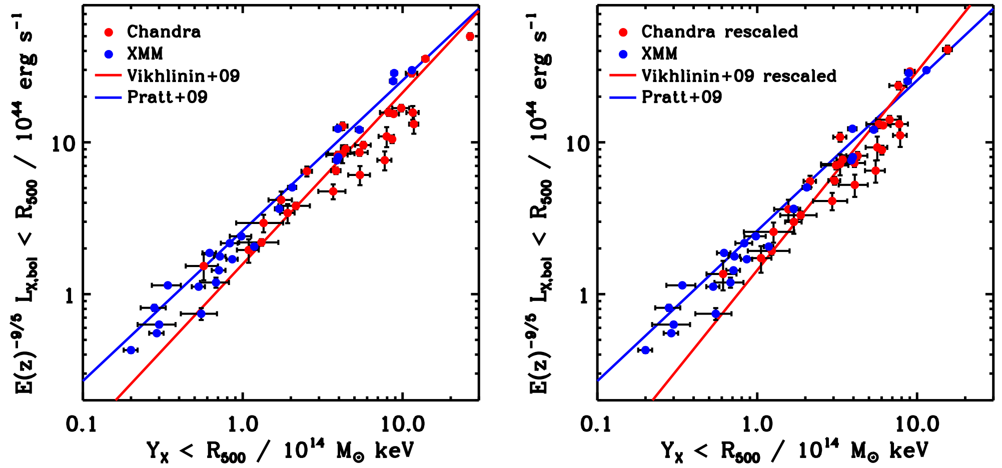

Figure 5. *Left*: The L_X-Y_X relation for individual clusters from Table 3. The red line is the L_X-Y_X relation (Vikhlinin et al. 2009a) based on *Chandra* calibration, while the blue line (Pratt et al. 2009) is based on *XMM–Newton* calibration. *Right*: *Chandra* data (L_X and Y_X) and L_X-Y_X relation are rescaled to *XMM–Newton* with the in-house cross-calibration in Appendix B. These clusters follow the normal L_X-Y_X relation with less scatters than those in Fig. 2.

The consistency between SZ and X-ray observations is expected, because they are both from the same ICM. The problem is that optically selected clusters post a challenge for robust stacking of the X-ray or SZ signals.

5.2 Possible solutions for the discrepancy

As mentioned in Section 1, there are some possible solutions for the discrepancy between the stacked *Planck* SZ fluxes and the model expectations for the optically selected clusters. First, we focus on the ICM part to rule out some solutions. The consistency between SZ and X-ray observations demonstrates that the calibration and contamination of the SZ signal are not significantly biased. The assumed pressure profile used to estimate the SZ signal (Arnaud et al. 2010) and the X-ray scaling relation are not significantly biased. These richest maxBCG clusters follow normal L_X-T_X and L_X-Y_X relations, which suggests that the fraction of X-ray-dark-but-optical-normal clusters is not significant (at least for massive ones), consistent with some previous studies (e.g. Andreon & Moretti 2011).

Secondly, we focus on galaxies to discuss some remaining solutions. The miscentring (e.g. A1986, A1961) decreases the X-ray luminosity/SZ signal, as some flux moves outside of the aperture. Sehgal et al. (2013) demonstrated the effect of miscentring on decreasing SZ signal. They also proposed that the miscentring effect causes their lower measured SZ signal compared to *Planck* due to the finer resolution of *ACT*. However, as they point out, the miscentring distribution from their sample alone can only explain part of the discrepancy between optical and SZ, unless an unrealistic larger offset exists. Moreover, most miscentred clusters are merging or disturbed clusters with lower X-ray or SZ surface brightness, which makes them more difficult to be detected. The discrepancy of SZ signal is at a level of 10 per cent in our richness range (Biesiadzinski et al. 2012). The projection (e.g. A750, A1319) will increase the optical richness. As the cluster mass-richness relation is close to a linear relation, projection causes the projected clusters to simply slide up and down the mass-richness relation,

Table 5. X-ray morphological parameters.

Cluster	w	$c_{\rm SB}$	s	p	a
A2142	0.007 ± 0.001	0.075 ± 0.001	0.788 ± 0.014	-0.921 ± 0.005	1.159 ± 0.017
J150M	0.011 ± 0.005	0.054 ± 0.010	0.843 ± 0.179	-1.259 ± 0.323	0.870 ± 0.308
J150E	0.025 ± 0.008	0.050 ± 0.015	0.772 ± 0.196	-1.596 ± 0.412	0.663 ± 0.284
J150W	0.026 ± 0.011	0.031 ± 0.014	0.475 ± 0.218	-1.694 ± 0.697	0.280 ± 0.240
A1689	0.002 ± 0.001	0.116 ± 0.002	1.166 ± 0.053	-0.829 ± 0.018	1.423 ± 0.035
A1443	0.037 ± 0.010	0.025 ± 0.003	0.650 ± 0.164	-1.501 ± 0.142	0.789 ± 0.090
A781	0.046 ± 0.002	0.023 ± 0.002	0.446 ± 0.023	-1.881 ± 0.825	0.941 ± 0.032
A781M	0.007 ± 0.002	0.029 ± 0.004	1.172 ± 0.150	-1.772 ± 0.349	1.203 ± 0.220
A781E	0.009 ± 0.003	0.020 ± 0.005	1.181 ± 0.216	-1.770 ± 0.492	1.381 ± 0.326
A781W	0.010 ± 0.003	0.073 ± 0.007	0.966 ± 0.151	-1.158 ± 0.071	1.030 ± 0.180
A1986	0.019 ± 0.005	0.047 ± 0.010			
A1882E	0.024 ± 0.005	0.051 ± 0.004	0.575 ± 0.091	-1.843 ± 0.540	0.995 ± 0.221
A1882W	0.020 ± 0.005	0.051 ± 0.010	0.552 ± 0.238	-1.674 ± 0.274	0.137 ± 0.234
A1882M	0.012 ± 0.005	0.271 ± 0.041	0.883 ± 0.150	-0.895 ± 0.036	0.819 ± 0.209
A1882N	0.034 ± 0.011	0.058 ± 0.016			
A1758N	0.026 ± 0.019	0.016 ± 0.005	0.517 ± 0.240	-1.604 ± 0.281	1.438 ± 0.090
A1758S	0.015 ± 0.001	0.041 ± 0.002	0.497 ± 0.034	-1.376 ± 0.402	0.971 ± 0.053
A1760	0.029 ± 0.005	0.037 ± 0.007			
A1622	0.036 ± 0.001	0.046 ± 0.001			
A750E	0.007 ± 0.001	0.075 ± 0.003	1.021 ± 0.098	-1.086 ± 0.022	1.298 ± 0.072
A750W	0.042 ± 0.004	0.056 ± 0.006	0.178 ± 0.056	-1.172 ± 0.208	0.827 ± 0.274
A1682	0.034 ± 0.006	0.039 ± 0.005	0.429 ± 0.072	-1.414 ± 0.332	0.444 ± 0.140
A1246	0.011 ± 0.004	0.047 ± 0.007	0.671 ± 0.214	-1.265 ± 0.111	0.745 ± 0.291
A1961	0.026 ± 0.008	0.051 ± 0.013	0.564 ± 0.124	-1.496 ± 0.798	0.624 ± 0.231
A2034	0.015 ± 0.002	0.027 ± 0.001	0.904 ± 0.039	-1.316 ± 0.081	0.911 ± 0.035
A655	0.018 ± 0.006	0.047 ± 0.005	0.466 ± 0.144	-1.427 ± 0.147	0.774 ± 0.236
A1914	0.010 ± 0.001	0.061 ± 0.002	0.660 ± 0.034	-1.000 ± 0.014	1.352 ± 0.069
Z5247	0.056 ± 0.013	0.029 ± 0.006	0.530 ± 0.149	-1.767 ± 0.673	1.247 ± 0.319
A657	0.011 ± 0.003	0.114 ± 0.007	0.571 ± 0.104	-1.073 ± 0.083	0.638 ± 0.187
J229	0.012 ± 0.003	0.061 ± 0.006	1.148 ± 0.195	-1.607 ± 0.517	1.100 ± 0.289
A2051	0.012 ± 0.001	0.050 ± 0.003	0.718 ± 0.074	-1.349 ± 0.212	1.001 ± 0.114
A2051S	0.009 ± 0.005	0.105 ± 0.019	0.888 ± 0.291	-1.353 ± 0.264	1.084 ± 0.455
A2051N	0.015 ± 0.005	0.069 ± 0.009	0.480 ± 0.138	-2.324 ± 1.295	0.757 ± 0.374
A2050	0.011 ± 0.001	0.052 ± 0.002	0.777 ± 0.050	-1.230 ± 0.072	1.335 ± 0.126
A1423	0.007 ± 0.004	0.101 ± 0.008	1.005 ± 0.111	-1.003 ± 0.036	1.034 ± 0.122
A801	0.014 ± 0.004	0.082 ± 0.009	0.772 ± 0.153	-1.141 ± 0.066	0.863 ± 0.186
A773	0.007 ± 0.002	0.045 ± 0.003	1.147 ± 0.115	-1.185 ± 0.052	1.264 ± 0.082
A1576	0.017 ± 0.004	0.048 ± 0.005	0.805 ± 0.097	-1.138 ± 0.051	1.021 ± 0.081
A2631	0.030 ± 0.004	0.033 ± 0.003	0.636 ± 0.083	-1.282 ± 0.123	0.988 ± 0.077
A1703	0.014 ± 0.003	0.047 ± 0.003	0.830 ± 0.166	-1.130 ± 0.046	1.396 ± 0.080
A2219	0.007 ± 0.001	0.027 ± 0.001	1.008 ± 0.073	-1.159 ± 0.033	1.442 ± 0.051
A1319M	0.013 ± 0.006	0.070 ± 0.022			
A1319NW	0.022 ± 0.009	0.089 ± 0.018			
A1560	0.065 ± 0.007	0.060 ± 0.007			
J175N	0.023 ± 0.014	0.043 ± 0.014			
J175S	0.017 ± 0.006	0.051 ± 0.015			
J249SW	0.014 ± 0.004	0.039 ± 0.006	0.751 ± 0.125	-1.586 ± 0.399	1.288 ± 0.218
J249NE	0.009 ± 0.004	0.072 ± 0.010	0.907 ± 0.249	-1.476 ± 0.566	0.792 ± 0.360
A1201	0.023 ± 0.001	0.100 ± 0.002	0.539 ± 0.048	-1.034 ± 0.022	0.696 ± 0.119
A2009	0.003 ± 0.001	0.163 ± 0.007	1.141 ± 0.141	-0.894 ± 0.026	1.381 ± 0.088
A2111	0.043 ± 0.014	0.040 ± 0.004	0.495 ± 0.108	-1.479 ± 0.567	0.792 ± 0.051
A815	0.027 ± 0.012	0.020 ± 0.005			
Z1450	0.009 ± 0.003	0.036 ± 0.009	0.878 ± 0.151	-1.525 ± 0.158	0.799 ± 0.240
A1765	0.005 ± 0.002	0.065 ± 0.009	0.908 ± 0.228	-1.604 ± 0.629	0.897 ± 0.400
A1902	0.008 ± 0.003	0.148 ± 0.009	0.796 ± 0.135	-0.939 ± 0.020	0.897 ± 0.172

Note. The centroid shift (w within R_{500}, Mohr et al. 1993), the surface brightness concentration ($c_{\rm SB}$ in 40–400 kpc range; Santos et al. 2008), and the SPA (Mantz et al. 2015).

without deviating from it (Simet et al. 2017). However, the L_X–M or $Y_{\rm SZ}$–M relation is a power-law relation with an index greater than one (∼1.6; Rozo et al. 2014a). For example, if we have two clusters with $N_{200} = 80$ ($M_{500|N=80} = 5 \times 10^{14}\,{\rm M}_\odot$, bolometric X-ray luminosity $L_{X|N=80} = 11 \times 10^{44}\,{\rm erg\,s}^{-1}$) projected together, they will be detected as a $N = 160$ cluster, the corresponding expected mass and X-ray luminosity are $M_{500|N=160} = 10 \times 10^{14}\,{\rm M}_\odot$ and $L_{X|N=160} = 43 \times 10^{44}\,{\rm erg\,s}^{-1}$, respectively. Though the mass is equal to the sum of two subclusters, the total X-ray luminosity is overestimated by a factor of ∼2 than the linear combination

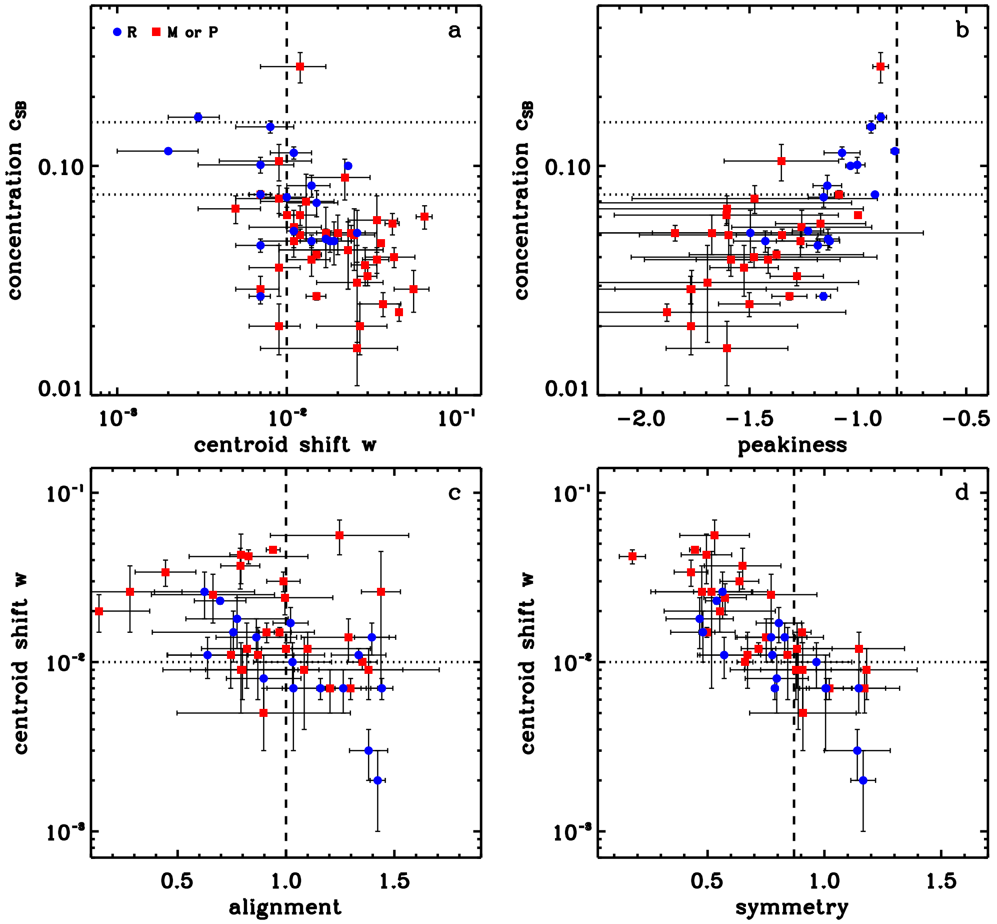

Figure 6. The distribution of X-ray morphology parameters. Red squares are disturbed clusters with the same miscentring and projection problems as in Figs 2 and 3, while blue dots are relaxed clusters without miscentring and projection problems. The lines mark the classification threshold: $w \le 0.01$ for relaxed clusters; $c_{\rm SB} \ge 0.075$ for moderate CC clusters and $c_{\rm SB} \ge 0.155$ for pronounced CC clusters; $s > 0.87$, $p > -0.82$, and $a > 1.00$ for relaxed clusters. The CC indicators $c_{\rm SB}$ and p are correlated with each other. The asymmetry indicators w, s, and a are also correlated.

of $2L_{X|N=80} = 22 \times 10^{44}\,{\rm erg\,s^{-1}}$. Thus, projection causes the expected X-ray luminosity or the SZ signal from the summed optical richness to be higher than the actual summed values. The projection fraction of samples extending to much lower richness is around 10 per cent (Simet et al. 2017), even higher for these richest maxBCG clusters (Fig. 2). The above two effects can act together, especially in super clusters and large-scale filaments (Fig. D1). The contamination of low-mass haloes, whose true halo mass is far below the value suggested by the optical richness, would also dilute and reduce the mean mass of the sample. We note that there is contamination of such low-mass systems based on Fig. A2. These low-mass haloes are mostly blended systems with boosted richness affected by nearby large-scale structure. Thus there is a mixture of halo masses at very high N_{200}, the clean and the blended, and the PDF of M given N_{200} (or λ) will be asymmetric with a low-mass tail. A skew-normal or Hermite polynomial expansion (Shaw, Holder & Dudley 2010) are good alternatives to mixture modelling. Next, we roughly estimate the contamination fraction. There are four low L_X systems beyond the 2σ line of L_X–N relation from Rozo et al. (2014b) in Fig. A2 (and seven beyond the 1.5σ line towards lower L_X versus 1 beyond the 1.5σ line towards higher L_X). Taking these numbers at face value, the contamination is 10–15 per cent.

Thirdly, the discrepancy may be induced by the mass bias and covariance bias between the ICM and galaxy scaling relations. The mass bias is either from the X-ray HE mass or the lensing mass. Rozo et al. (2014b) found the discrepancy problem could be solved by lowering the galaxy weak-lensing mass by 10 per cent while raising the X-ray mass by 21 per cent. The weak-lensing mass could be overestimated due to LOS contamination and triaxial halo (e.g. Corless & King 2007). The red-sequence cluster finding algorithm tends to find more prolate clusters (major axis aligned with LOS) than oblate clusters due to higher galaxy density and background contrast. This orientation bias can lead to 3–6 per cent overestimate of lensing mass (Dietrich et al. 2014). The covariance bias between $M_{\rm wl}$ and N at fixed mass can also induce a 10 per cent overestimate of lensing mass (equation 12 of Rozo et al. 2014b and Appendix A). We compare the weak-lensing mass with the mass estimated from the galaxy velocity dispersion (Farahi et al. 2016), and CMB weak-lensing (Geach & Peacock 2017; Baxter et al. 2018). Indeed, the mean galaxy weak-lensing mass is higher than the mass from the other two methods, although uncertainty with the latter two relations is substantial. Meanwhile, the X-ray HE mass could be underestimated due to gas bulk motion and turbulence (e.g. Nagai, Kravtsov & Vikhlinin 2007; Lau, Kravtsov & Nagai 2009). We also include the 25 per cent HE bias as solid lines in Figs 1–3. Compared with the dashed line of no mass bias gives an idea of how much the mass bias could reduce the discrepancy. Moreover, the covariance bias from the multivariate scaling relations (Appendix A) leads an additional ∼10 per cent correction.

5.3 Can we further increase the robustness of the optical richness

From the ICM–galaxy scaling relations, we find that optically selected clusters have large scatter with more clusters biased to lower mass. Both Planck Collaboration XII (2011) and Sehgal et al. (2013) selected a BCG-dominated subsample of their optically selected clusters, defined as the BCG luminosity ratio $L_{\rm BCG}/(L_{\rm tot}-L_{\rm BCG})$, being larger than the average ratio for a given richness bin. They find that the BCG-dominated sample has a higher normalization, closer to the predicted relation.

Planck Collaboration XI (2013) compared the scaling relation between the SZ signal and the stellar mass for a large sample of locally brightest galaxies, analogous to a BCG-dominated sample. The relation is close to the self-similar prediction extending from rich clusters down to groups ($M_{500} \sim 2 \times 10^{13}\,{\rm M}_{\odot}$), but with a normalization ∼20 per cent lower than the X-ray selected clusters. This discrepancy is mainly from the Malmquist bias from the X-ray sample and the miscentring from satellite contamination (Planck Collaboration XI 2013). Meanwhile, Hand et al. (2011) and Greco et al. (2015) found consistent results, at least down to the group mass scale (but necessarily to the lower mass systems). Moreover, Gralla et al. (2014) found that the stacked SZ signal from radio selected sources is also consistent with the self-similar prediction. The SZ signal is mainly from the AGN hosted haloes of giant galaxies instead of galaxy clusters or groups. The miscentring and projection problems are insignificant for such giant galaxies compared with the massive optically selected clusters.

We also select a BCG-dominated subsample with the maxBCG and redMaPPer catalogues locating the same BCG. Fig. A1 shows that the BCG-dominated clusters tend to be relaxed clusters and agree better with the model prediction than the full sample. Three of the four low-luminosity outliers have the projection problems (see Fig. 2). Therefore, the BCG-dominated subsample more closely correspond to the X-ray selected and SZ selected samples.

Another potentially useful information could improve the optical mass proxy is the optical luminosity of either member galaxies, or the BCG, or the intracluster light (e.g. Gonzalez, Zaritsky & Zabludoff 2007; Reyes et al. 2008). Though the stars are a minor fraction of baryon in clusters, they are still scaled with the total halo mass. Moreover, the velocity dispersion from the upcoming Dark Energy Spectroscopic Instrument survey will also increase the robustness of the optical richness.

5.4 Dynamical state and cool core fraction

As an optically selected sample, this sample also provides an opportunity to study the ICM dynamical state and CC fraction using the X-ray morphology parameters presented in Section 4.3, without any ICM selection bias. Due to the diversity in the recent merger histories of individual clusters, which is further complicated by projection, the morphological parameters should be treated with caution. However, they provide useful statistical tools to characterize trends of properties in large cluster samples (e.g. Böhringer et al. 2010). We use w and $c_{\rm SB}$ to compare the fractions of relaxed ($w \le 0.01$) and CC ($c_{\rm SB} \ge 0.075$) clusters among optical, SZ, and X-ray selected samples (Table 6). Both the fractions of relaxed and CC clusters change as: optical < SZ < X-ray. Moreover, the combination of asymmetry and CC indicators provide a more rigorous definition for relaxed clusters, e.g. the SPA criterion ($s > 0.87$, $p > -0.82$, and $a > 1.00$; Mantz et al. 2015). Only 1 out of 55 clusters in our sample (2 per cent) is close to the SPA criterion, compared with 57/361 (16 per cent) for an X-ray selected sample (Mantz et al. 2015). Simulations also tend to find less relaxed clusters compared with an X-ray selected sample (e.g. Böhringer et al. 2010).

Table 6. The fractions of relaxed and CC clusters.

	optical	SZ	X-ray
Relaxed ($w \le 0.01$)	29% (55[a])	35% (120[b])	61% (31[c])
CC ($c_{\rm SB} \ge 0.075$)	22% (55[a])	36% (164[d])	61% (100[d])

Note. This optically selected sample shows lower fractions of relaxed and CC clusters compared with SZ and X-ray selected samples.

[a] This work from Table 5, the number in brackets is the total cluster number in the sample.

[b] Lovisari et al. (2017).

[c] Böhringer et al. (2010).

[d] Andrade-Santos et al. (2017).

In Table A1, we further compare the amplitudes (fix the slope to the model prediction) of the whole (all) sample, as well as the relaxed (R) and disturbed ($M + P$) subsamples, with the model prediction including 25 per cent mass bias. We find that the amplitude of whole sample is close to the model prediction, while the relaxed subsample is higher and the disturbed subsample is lower than the model. This fact indicates that cluster with different dynamical state may have different level of mass bias.

6 CONCLUSIONS

Great progress on galaxy clusters has been made in the last decade with X-ray, SZ, and optical surveys. The scaling relations between X-ray and SZ are consistent with each other for X-ray, SZ and even optically selected clusters, because both the X-ray emission and SZ signal are from the same ICM. However, discrepancies emerge when we compare the ICM scaling relation based on X-ray and SZ data with the galaxy scaling relation based on optical data.

In order to study the discrepancies, we directly compare the optical and X-ray scaling relations for a complete sample of 38 richest maxBCG clusters. We list these factors contributing to the discrepancies: (1) miscentring, ∼10 per cent; (2) projection, ∼10 per cent; (3) contamination of low-mass systems of optical selection, ∼10–15 per cent; (4) hydrostatic mass bias, ∼25 per cent; (5) weak-lensing mass bias, ∼10 per cent; and (6) covariance bias, ∼10 per cent. These biases mix in some cases and can compensate with each other, but the dominant one is the mass bias. More studies are required to constrain these biases better.

This sample offers insights into maxBCG mass selection $P(M|N)$ from the distribution of $P(X|N)$, where X is the low scatter X-ray mass proxy like T_X, L_X, and Y_X. In such top richest (∼0.3 per cent of whole maxBCG sample) range, a significant amount of blended haloes with boosted richness mixes with the clean haloes. However, all the blended haloes are resolved by the X-ray imaging, as they follow the X-ray scaling relations calibrated from X-ray selected haloes. The fraction of blended system could be reduce when including more information from optical data, e.g. whether hosts a dominant BCG or not.

This optically selected sample also provides an unbiased perspective to the ICM properties. We find a rising fraction of relaxed or

CC clusters from optical (∼26 per cent), to SZ (∼36 per cent), and to X-ray (∼61 per cent) selected samples. Moreover, the disturbed subsample shows higher mass bias than the relaxed subsample.

Optical surveys and algorithms are very successful and efficient for finding clusters, more works need to be done to better understand the halo selection properties of optical catalogue and the mass bias, with the aid of simulations and mock catalogue, before we implement the resulting scaling relations to study cosmology.

ACKNOWLEDGEMENTS

We thank the referee, Gus Evrard, for important comments and suggestions. We thank Andrea Morandi for his early work on the *Chandra* data. Support for this work was provided by the National Aeronautics and Space Administration grants NNX16AH32G and NNX16AH26G. Support for this work was also provided by the National Aeronautics and Space Administration through *Chandra* Award Number GO4-15119B and GO4-15115X issued by the *Chandra* X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under contract NAS8-03060. ER is supported by DOE grant DE-SC0015975 and by the Sloan Foundation grant FG-2016-6443. NS acknowledges support from NSF grant 1513618. DN acknowledges support from NSF grant AST-1412768. This research has made use of data and/or software provided by the High Energy Astrophysics Science Archive Research Center (HEASARC), which is a service of the Astrophysics Science Division at NASA/GSFC and the High Energy Astrophysics Division of the Smithsonian Astrophysical Observatory.

REFERENCES

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
Andersson K. E., Madejski G. M., 2004, ApJ, 607, 190
Andersson K. et al., 2011, ApJ, 738, 48
Andrade-Santos F. et al., 2017, ApJ, 843, 76
Andreon S., Moretti A., 2011, A&A, 536, A37
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Arnaud K. A., 1996, in Jacoby G. H., Barnes J., eds, Proc. ASP Conf. Ser. Vol. 101, Astronomical Data Analysis Software and Systems V. Astron. Soc. Pac., San Francisco, p. 17
Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Bacchi M., Feretti L., Giovannini G., Govoni F., 2003, A&A, 400, 465
Barrena R., Boschin W., Girardi M., Spolaor M., 2007, A&A, 467, 37
Barrena R., Girardi M., Boschin W., 2013, MNRAS, 430, 3453
Baxter E. J. et al., 2018, MNRAS, 476, 2674
Biesiadzinski T., McMahon J., Miller C. J., Nord B., Shaw L., 2012, ApJ, 757, 1
Böhringer H. et al., 2010, A&A, 514, A32
Bonafede A. et al., 2015, MNRAS, 454, 3391
Boschin W., Girardi M., Barrena R., Biviano A., Feretti L., Ramella M., 2004, A&A, 416, 839
Bower R. G., McCarthy I. G., Benson A. J., 2008, MNRAS, 390, 1399
Buote D. A., Gastaldello F., Humphrey P. J., Zappacosta L., Bullock J. S., Brighenti F., Mathews W. G., 2007, ApJ, 664, 123
Canning R. E. A. et al., 2017, MNRAS, 464, 2896
Carlstrom J. E., Holder G. P., Reese E. D., 2002, ARA&A, 40, 643
Cassano R., Ettori S., Giacintucci S., Brunetti G., Markevitch M., Venturi T., Gitti M., 2010, ApJ, 721, L82
Cohn J. D., Evrard A. E., White M., Croton D., Ellingson E., 2007, MNRAS, 382, 1738
Corless V. L., King L. J., 2007, MNRAS, 380, 149
Croton D. J. et al., 2006, MNRAS, 365, 11
Dahle H., Kaiser N., Irgens R. J., Lilje P. B., Maddox S. J., 2002, ApJS, 139, 313
David L. P., Kempner J., 2004, ApJ, 613, 831
Dietrich J. P. et al., 2014, MNRAS, 443, 1713
Durret F., Laganá T. F., Haider M., 2011, A&A, 529, A38
Edge A. C., Smith G. P., Sand D. J., Treu T., Ebeling H., Allen S. W., van Dokkum P. G., 2003, ApJ, 599, L69
Eke V. R., Cole S., Frenk C. S., Patrick Henry J., 1998, MNRAS, 298, 1145
Evrard A. E., Arnault P., Huterer D., Farahi A., 2014, MNRAS, 441, 3562
Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, MNRAS, 460, 3900
Farahi A., Evrard A. E., McCarthy I., Barnes D. J., Kay S. T., 2018, MNRAS, 478, 2618
Foster A. R., Ji L., Smith R. K., Brickhouse N. S., 2012, ApJ, 756, 128
Freeland E., Stilp A., Wilcots E., 2009, AJ, 138, 295
Ge C., Wang Q. D., Tripp T. M., Li Z., Gu Q., Ji L., 2016, MNRAS, 459, 366
Geach J. E., Peacock J. A., 2017, Nat. Astron., 1, 795
Giovannini G., Bonafede A., Feretti L., Govoni F., Murgia M., Ferrari F., Monti G., 2009, A&A, 507, 1257
Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147
Govoni F., Feretti L., Giovannini G., Böhringer H., Reiprich T. H., Murgia M., 2001, A&A, 376, 803
Govoni F., Markevitch M., Vikhlinin A., van Speybroeck L., Feretti L., Giovannini G., 2004, ApJ, 605, 695
Govoni F., Murgia M., Giovannini G., Vacca V., Bonafede A., 2011, A&A, 529, A69
Gralla M. B. et al., 2014, MNRAS, 445, 460
Greco J. P., Hill J. C., Spergel D. N., Battaglia N., 2015, ApJ, 808, 151
Hahn O., Martizzi D., Wu H.-Y., Evrard A. E., Teyssier R., Wechsler R. H., 2017, MNRAS, 470, 166
Hand N. et al., 2011, ApJ, 736, 39
Henriksen M., Wang Q. D., Ulmer M., 1999, MNRAS, 307, 67
Henry J. P., Arnaud K. A., 1991, ApJ, 372, 410
Henson M. A., Barnes D. J., Kay S. T., McCarthy I. G., Schaye J., 2017, MNRAS, 465, 3361
Hickson P., 1982, ApJ, 255, 382
Hinshaw G. et al., 2013, ApJS, 208, 19
Hudson D. S., Mittal R., Reiprich T. H., Nulsen P. E. J., Andernach H., Sarazin C. L., 2010, A&A, 513, A37
Johnston D. E. et al., 2007, preprint (arXiv:0709.1159)
Kaiser N., 1986, MNRAS, 222, 323
Kale R. et al., 2015, A&A, 579, A92
Koester B. P. et al., 2007, ApJ, 660, 239
Kravtsov A. V., Nagai D., Vikhlinin A. A., 2005, ApJ, 625, 588
Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128
Kuntz K. D., Snowden S. L., 2008, A&A, 478, 575
Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Lovisari L. et al., 2017, ApJ, 846, 51
Ma C.-J., Owers M., Nulsen P. E. J., McNamara B. R., Murray S. S., Couch W. J., 2012, ApJ, 752, 139
Macario G. et al., 2013, A&A, 551, A141
Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010, MNRAS, 406, 1773
Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., von der Linden A., Urban O., 2015, MNRAS, 449, 199
Mantz A. B. et al., 2016, MNRAS, 463, 3582
Markevitch M. et al., 2000, ApJ, 541, 542
Melin J.-B., Bartlett J. G., Delabrouille J., Arnaud M., Piffaretti R., Pratt G. W., 2011, A&A, 525, A139
Miyatake H. et al., 2018, preprint (arXiv:1804.05873)
Mohr J. J., Fabricant D. G., Geller M. J., 1993, ApJ, 413, 492
Monteiro-Oliveira R., Cypriano E. S., Machado R. E. G., Lima Neto G. B., Ribeiro A. L. B., Sodré L., Dupke R., 2017, MNRAS, 466, 2614
Morandi A., Pedersen K., Limousin M., 2011, ApJ, 729, 37

Morandi A., Sun M., Forman W., Jones C., 2015, MNRAS, 450, 2261

Nagai D., Kravtsov A. V., Vikhlinin A., 2007, ApJ, 668, 1

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

Noh Y., Cohn J. D., 2011, MNRAS, 413, 301

Okabe N., Umetsu K., 2008, PASJ, 60, 345

Okabe N., Takada M., Umetsu K., Futamase T., Smith G. P., 2010, PASJ, 62, 811

Owers M. S., Nulsen P. E. J., Couch W. J., Markevitch M., Poole G. B., 2009, ApJ, 692, 702

Owers M. S., Nulsen P. E. J., Couch W. J., 2011, ApJ, 741, 122

Owers M. S. et al., 2013, ApJ, 772, 104

Owers M. S. et al., 2014, ApJ, 780, 163

Patej A., Loeb A., 2016, ApJ, 824, 69

Penna-Lima M., Bartlett J. G., Rozo E., Melin J.-B., Merten J., Evrard A. E., Postman M., Rykoff E., 2017, A&A, 604, A89

Planck Collaboration IX, 2011, A&A, 536, A9

Planck Collaboration X, 2011, A&A, 536, A10

Planck Collaboration XI, 2011, A&A, 536, A11

Planck Collaboration XII, 2011, A&A, 536, A12

Planck Collaboration XI, 2013, A&A, 557, A52

Planck Collaboration XXVII, 2016, A&A, 594, A27

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

Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716

Reyes R., Mandelbaum R., Hirata C., Bahcall N., Seljak U., 2008, MNRAS, 390, 1157

Richard J., Pei L., Limousin M., Jullo E., Kneib J. P., 2009, A&A, 498, 37

Rossetti M., Eckert D., De Grandi S., Gastaldello F., Ghizzardi S., Roediger E., Molendi S., 2013, A&A, 556, A44

Rozo E., Rykoff E. S., 2014, ApJ, 783, 80

Rozo E. et al., 2009a, ApJ, 699, 768

Rozo E. et al., 2009b, ApJ, 703, 601

Rozo E. et al., 2010, ApJ, 708, 645

Rozo E., Rykoff E., Koester B., Nord B., Wu H.-Y., Evrard A., Wechsler R., 2011, ApJ, 740, 53

Rozo E., Evrard A. E., Rykoff E. S., Bartlett J. G., 2014a, MNRAS, 438, 62

Rozo E., Bartlett J. G., Evrard A. E., Rykoff E. S., 2014b, MNRAS, 438, 78

Rykoff E. S. et al., 2008, ApJ, 675, 1106

Rykoff E. S. et al., 2014, ApJ, 785, 104

Santos J. S., Rosati P., Tozzi P., Böhringer H., Ettori S., Bignamini A., 2008, A&A, 483, 35

Saro A. et al., 2017, MNRAS, 468, 3347

Sato K., Matsushita K., Yamasaki N. Y., Sasaki S., Ohashi T., 2014, PASJ, 66, 85

Schellenberger G., Reiprich T. H., Lovisari L., Nevalainen J., David L., 2015, A&A, 575, A30

Sehgal N., Hughes J. P., Wittman D., Margoniner V., Tyson J. A., Gee P., dell'Antonio I., 2008, ApJ, 673, 163

Sehgal N. et al., 2011, ApJ, 732, 44

Sehgal N. et al., 2013, ApJ, 767, 38

Sereno M., Ettori S., Umetsu K., Baldi A., 2013, MNRAS, 428, 2241

Shaw L. D., Holder G. P., Dudley J., 2010, ApJ, 716, 281

Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, MNRAS, 466, 3103

Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2018, MNRAS, 480, 5385

Snowden S. L., Mushotzky R. F., Kuntz K. D., Davis D. S., 2008, A&A, 478, 615

Sun M., 2012, New J. Phys., 14, 045004

Sun M., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., 2009, ApJ, 693, 1142

Sun M., Sehgal N., Voit G. M., Donahue M., Jones C., Forman W., Vikhlinin A., Sarazin C., 2011, ApJ, 727, L49

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Truong N. et al., 2018, MNRAS, 474, 4089

Umetsu K. et al., 2015, ApJ, 806, 207

Vacca V., Govoni F., Murgia M., Giovannini G., Feretti L., Tugnoli M., Verheijen M. A., Taylor G. B., 2011, A&A, 535, A82

Venturi T., Giacintucci S., Dallacasa D., Cassano R., Brunetti G., Bardelli S., Setti G., 2008, A&A, 484, 327

Venturi T. et al., 2017, A&A, 603, A125

Vikhlinin A. et al., 2009a, ApJ, 692, 1033

Vikhlinin A. et al., 2009b, ApJ, 692, 1060

Wang Q. D., Ulmer M. P., Lavery R. J., 1997, MNRAS, 288, 702

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rev., 530, 87

Willingale R., Starling R. L. C., Beardmore A. P., Tanvir N. R., O'Brien P. T., 2013, MNRAS, 431, 394

APPENDIX A: MULTIVARIATE SCALING RELATIONS

Due to the slope or asymmetry of the mass function and the mass variance, the mean of the correlated multivariate scaling relation is not equal to the naive 'plug in' expectation (e.g. Allen et al. 2011; Evrard et al. 2014). Evrard et al. (2014) model is aimed at describing the underlying massive halo population. The $P(N_{\rm gal}|M)$ kernel can be separated into intrinsic halo scatter, $P(N_{\rm int}|M)$, and an LOS/noise component, $P(N_{\rm gal}|N_{\rm int}, M)$. Asymmetry in the latter PDF is expected from non-linear clustering. This generates a corresponding skewness in the mass selection, $P(M|N_{\rm gal})$. The selection of the 38 (0.3 per cent of whole sample) richest maxBCG clusters will exacerbate the skewness relative to a lower richness selection. However, as a first-order approximation and largely for illustrative purposes here, we assume the conditional probability is a lognormal distribution.

Following Rozo et al. (2014a), we define $m = \ln(M/M_0)$ and the mean distribution of log-observable, e.g. temperature $t = \ln(T_X/T_{X,0})$, is

$$\langle t|m\rangle = a_{t|m} + s_{t|m}\, m, \tag{A1}$$

where $a_{t|m}$ is the amplitude and $s_{t|m}$ is the slope of scaling relation. Using Bayes Theorem, we can convert the T_X-M relation to $M-T_X$ relation with the equation A5 of Rozo et al. (2014a),

$$\langle m|t\rangle = [\frac{t - a_{t|m}}{s_{t|m}}] - \beta\sigma^2_{m|t}, \tag{A2}$$

The first term in square brackets is the naive expected mean from equation (A1), while the second term is the Eddington bias correction. β is the slope of the halo mass function ($\mathrm{d}n/\mathrm{d}\ln M \propto M^{-\beta}$) and $\sigma_{m|t}$ is the scatter in m at fixed t. Assuming the Tinker et al. (2008) mass function and the *WMAP9* cosmology (Hinshaw et al. 2013), in the typical mass range of this sample (0.6–$13.3 \times 10^{14}\,\mathrm{M}_\odot$), $\beta = 1.4$–4.5.

L_X–N relation is derived from the L_X-M relation (Pratt et al. 2009; L_1-M_Y MB in their table 2) and the $M-N$ relation (Rozo et al. 2009a; Simet et al. 2017) using equation A13 of Rozo et al. (2014a),

$$\langle l|n\rangle = a_{l|m} + s_{l|m}(\langle m|n\rangle + r\beta\sigma_{m|l}\sigma_{m|n}). \tag{A3}$$

Thus, the amplitude and slope of L_X-N relation are

$$a_{l|n} = [a_{l|m} + s_{l|m}a_{m|n}] + r_{l,n|m}\beta s_{l|m}\sigma_{m|l}\sigma_{m|n}, \tag{A4}$$

$$s_{l|n} = s_{l|m}s_{m|n}. \tag{A5}$$

The scatter of L_X-N relation is estimated using equation A14 of Rozo et al. (2014a):

$$\sigma^2_{l|n} = s^2_{l|m}(\sigma^2_{m|n} + \sigma^2_{m|l} - 2r_{l,n|m}\sigma_{m|n}\sigma_{m|l}). \tag{A6}$$

On the amplitude, the term in square brackets is the naive 'plug

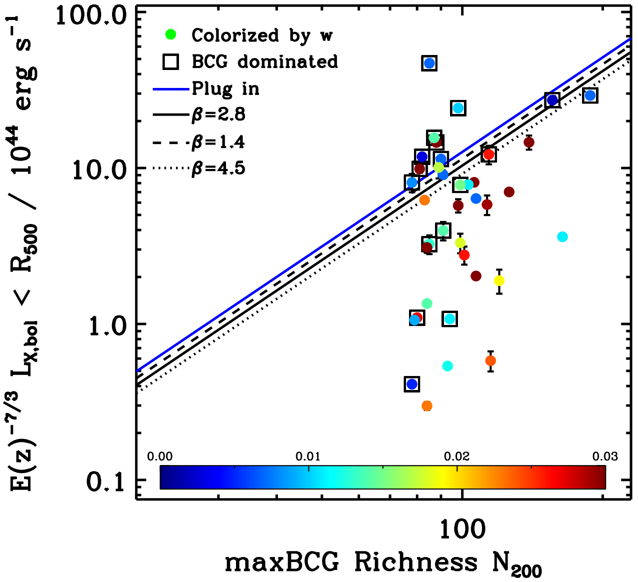

Figure A1. Comparison of the 'plug in' L_X–N_{200} relation (blue solid line) with the bias-corrected ones. The 'plug in' relation simply combines the X-ray L_X–M relation (Pratt et al. 2009) and the weak-lensing M–N_{200} relation (Rozo et al. 2009a). The black solid, dashed, and dotted lines show bias-corrected relation for β of 2.8, 1.4, and 4.5, respectively. The dots are colourized by the centroid shift w. The black boxes represent a BCG dominated subsample with the maxBCG and redMaPPer locating the same galaxy as the BCG (see Section 5.3). The BCG dominated clusters tend to be relaxed clusters and follow more closely to the model prediction of solid line.

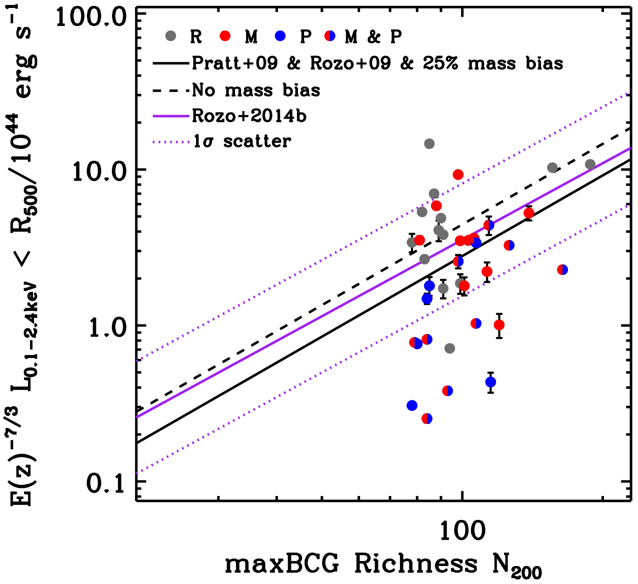

Figure A2. The same plot as Fig. 2, but for soft band 0.1–2.4 keV luminosity. The purple solid line with 1σ dashed lines are from Rozo et al. (2014b).

in' value. The other term is the covariance bias, $r_{l,n|m}$ is the correlation coefficient between l and n at fixed m. We note the binned masses of Rozo et al. (2009a) and Simet et al. (2017) measure the $\ln\langle M|N\rangle$ rather than $\langle \ln M|N\rangle$, and $\ln\langle M|N\rangle = \langle \ln M|N\rangle + \frac{1}{2}\sigma^2_{m|n}$ assuming a lognormal distribution, thus the amplitude is related as $\tilde{a}_{m|n} = a_{m|n} + \frac{1}{2}\sigma^2_{m|n}$. We quote their $\tilde{a}_{m|n}$ and get $a_{m|n}$ by subtracting $\frac{1}{2}\sigma^2_{m|n}$. To get an order of magnitude estimate for the bias term, we set $\beta = 2.8$. To the first order, the hot gas–galaxy correlation coefficient is zero, which results in no bias. However, recent studies suggest negative hot gas–galaxy correlation and we also take $r_{l,n|m} = -0.5$ (Farahi et al. 2018 and private communication with Gus Evrard) to examine its impact. We also apply $\sigma_{m|l} = 0.28$ (Angulo et al. 2012) and $\sigma_{m|n} = 0.45$ (Rozo et al. 2009a) or $\sigma_{m|\lambda} = 0.25$ (Rozo & Rykoff 2014), which yields a correction term ~ -0.34 (~ -0.17 for λ), corresponds to a 28 per cent (15 per cent) down offset for the amplitude of L_X. If we set β in the range of 1.4–4.5, the down offset is 15–42 per cent (8–23 per cent for λ) instead. Fig. A1 compares the difference between the 'plug in' method and the bias-corrected method with different β values for the maxBCG sample. In order to compare with previous publications, especially the ones from *ROSAT*, we also present the L_X in 0.1–2.4 keV in Fig. A2. The $L_{0.1\text{–}2.4\ \mathrm{keV}}$ is converted from $L_{X,\mathrm{bol}}$ using an APEC model with metallicity fixed to 0.3 $\mathrm{Z}_\odot$. The black line is the predicted L_X–N relation from the L_X–M relation (Pratt et al. 2009; $L[0.1\text{–}2.4] - M_Y$ MB in their table B2) and the M–N relation (Rozo et al. 2009a) using equations (A4) and (A5). The purple line is from the preferred L_X–N relation (Rozo et al. 2014b; in their table 4).

$\boldsymbol{T_X}$**–*N*** relation is derived from the M–T_X relation (Sun et al. 2009; Tier 1 + 2 + clusters in their table 6) and the M–N relation (Rozo et al. 2009a; Simet et al. 2017). From equation A13 of Rozo et al. (2014a), we find

$$\langle t|n\rangle = a_{t|m} + s_{t|m}(\langle m|n\rangle + r_{t,n|m}\beta\sigma_{m|t}\sigma_{m|n}). \tag{A7}$$

The M–T_X relation is related to the T_X–M relation from equation (A2),

$$a_{t|m} = -\frac{a_{m|t} + \beta\sigma^2_{m|t}}{s_{m|t}},\ \ s_{t|m} = \frac{1}{s_{m|t}}, \tag{A8}$$

substitute $a_{t|m}$ and $s_{t|m}$ into equation (A7), we get the amplitude and slope of T_X–N relation as

$$a_{t|n} = [\frac{a_{m|n} - a_{m|t}}{s_{m|t}}] + \frac{\beta\sigma_{m|t}}{s_{m|t}}(r_{t,n|m}\sigma_{m|n} - \sigma_{m|t}), \tag{A9}$$

$$s_{t|n} = \frac{s_{m|n}}{s_{m|t}}. \tag{A10}$$

Similar to the scatter of L_X–N relation, but substitute $s_{t|m} = 1/s_{m|t}$,

$$\sigma^2_{t|n} = \frac{1}{s^2_{m|t}}(\sigma^2_{m|n} + \sigma^2_{m|t} - 2r_{t,n|m}\sigma_{m|n}\sigma_{m|t}). \tag{A11}$$

We assume $r_{t,n|m} = -0.5$ (Farahi et al. 2018 and private communication with Gus Evrard) and $\sigma_{m|t} = 0.20$ (Kravtsov, Vikhlinin & Nagai 2006) to estimate the bias, which yields a correction term of ~ -0.13, corresponding to a 12 per cent down offset for n or λ.

$\boldsymbol{Y_X}$**–*N*** relation is derived from the M–Y_X relation (Arnaud et al. 2010; their equation 2) and the M–N relation (Rozo et al. 2009a; Simet et al. 2017) as

$$a_{x|n} = [\frac{a_{m|n} - a_{m|x}}{s_{m|x}}] + \frac{\beta\sigma_{m|x}}{s_{m|x}}(r_{x,n|m}\sigma_{m|n} - \sigma_{m|x}), \tag{A12}$$

$$s_{x|n} = \frac{s_{m|n}}{s_{m|x}}. \tag{A13}$$

$$\sigma^2_{x|n} = \frac{1}{s^2_{m|x}}(\sigma^2_{m|n} + \sigma^2_{m|x} - 2r_{x,n|m}\sigma_{m|n}\sigma_{m|x}). \tag{A14}$$

Assuming $r_{x,n|m} = -0.5$ (Farahi et al. 2018 and private communication with Gus Evrard) and $\sigma_{m|x} = 0.087$ (Arnaud et al. 2010), we

Table A1. Scaling relation.

Relation	Amplitude	Slope	Scatter	Sample
		$M-X$ relation from literature		
L_X-M_{500}	$a_{l\|m} = -0.68 \pm 0.08$	$s_{l\|m} = 1.90 \pm 0.11$	$\sigma_{m\|l} = 0.28^{\mathrm{a}}$	Pratt et al. (2009)
$M_{500}-T_X$	$a_{m\|t} = -1.34 \pm 0.06$	$s_{m\|t} = 1.65 \pm 0.04$	$\sigma_{m\|t} = 0.20^{\mathrm{b}}$	Sun et al. (2009)
$M_{500}-Y_X$	$a_{m\|x} = 0.84 \pm 0.02$	$s_{m\|x} = 0.56 \pm 0.02$	$\sigma_{m\|x} = 0.09$	Arnaud et al. (2010)
		$M-N$ relation from literature		
$M_{500}-N_{200}{}^{\mathrm{c}}$	$a_{m\|n} = 0.86 \pm 0.12$	$s_{m\|n} = 1.06 \pm 0.11$	$\sigma_{m\|n} = 0.45 \pm 0.19$	Rozo et al. (2009a)
$M_{500}-\lambda^{\mathrm{d}}$	$a_{m\|n} = 0.66 \pm 0.07$	$s_{m\|n} = 1.30 \pm 0.10$	$\sigma_{m\|n} = 0.25 \pm 0.05$	Simet et al. (2017)
		$X-N$ relation from model		
T_X-N_{200}	$a_{t\|n} = 1.02 \pm 0.09$	$s_{t\|n} = 0.64 \pm 0.07$	$\sigma_{t\|n} = 0.35 \pm 0.11$	maxBCG (Fig. 1 left)
$T_X-\lambda$	$a_{t\|n} = 0.93 \pm 0.06$	$s_{t\|n} = 0.79 \pm 0.06$	$\sigma_{t\|n} = 0.24 \pm 0.03$	redMaPPer (Fig. 1 right)
L_X-N_{200}	$a_{l\|n} = 0.08 \pm 0.28$	$s_{l\|n} = 2.01 \pm 0.24$	$\sigma_{l\|n} = 1.22 \pm 0.33$	maxBCG (Fig. 2 left)
$L_X-\lambda$	$a_{l\|n} = -0.16 \pm 0.16$	$s_{l\|n} = 2.47 \pm 0.24$	$\sigma_{l\|n} = 0.87 \pm 0.10$	redMaPPer (Fig. 2 right)
Y_X-N_{200}	$a_{x\|n} = -0.61 \pm 0.22$	$s_{x\|n} = 1.89 \pm 0.21$	$\sigma_{x\|n} = 0.90 \pm 0.33$	maxBCG (Fig. 3 left)
$Y_X-\lambda$	$a_{x\|n} = -0.93 \pm 0.13$	$s_{x\|n} = 2.32 \pm 0.20$	$\sigma_{x\|n} = 0.54 \pm 0.09$	redMaPPer (Fig. 3 right)
		$X-N$ relation from data		
T_X-N_{200}	$a_{t\|n} = 1.05 \pm 0.16$	$s_{t\|n} = 0.64$ (fixed)	$\sigma_{t\|n} = 0.51 \pm 0.03$	maxBCG (all)
T_X-N_{200}	$a_{t\|n} = 1.31 \pm 0.28$	$s_{t\|n} = 0.64$ (fixed)	$\sigma_{t\|n} = 0.46 \pm 0.10$	maxBCG (*R*)
T_X-N_{200}	$a_{t\|n} = 0.92 \pm 0.28$	$s_{t\|n} = 0.64$ (fixed)	$\sigma_{t\|n} = 0.55 \pm 0.08$	maxBCG ($M+P$)
L_X-N_{200}	$a_{l\|n} = 0.01 \pm 0.16$	$s_{l\|n} = 2.01$ (fixed)	$\sigma_{l\|n} = 1.18 \pm 0.02$	maxBCG (all)
L_X-N_{200}	$a_{l\|n} = 0.73 \pm 0.28$	$s_{l\|n} = 2.01$ (fixed)	$\sigma_{l\|n} = 0.95 \pm 0.05$	maxBCG (*R*)
L_X-N_{200}	$a_{l\|n} = -0.37 \pm 0.20$	$s_{l\|n} = 2.01$ (fixed)	$\sigma_{l\|n} = 1.13 \pm 0.02$	maxBCG ($M+P$)
Y_X-N_{200}	$a_{x\|n} = -0.56 \pm 0.17$	$s_{x\|n} = 1.89$ (fixed)	$\sigma_{x\|n} = 1.13 \pm 0.02$	maxBCG (all)
Y_X-N_{200}	$a_{x\|n} = -0.10 \pm 0.28$	$s_{x\|n} = 1.89$ (fixed)	$\sigma_{x\|n} = 0.98 \pm 0.05$	maxBCG (*R*)
Y_X-N_{200}	$a_{x\|n} = -0.85 \pm 0.22$	$s_{x\|n} = 1.89$ (fixed)	$\sigma_{x\|n} = 1.14 \pm 0.03$	maxBCG ($M+P$)

Note. The scaling relation takes the form $\langle \ln \psi \rangle = a + s \ln(\chi/\chi_0)$, they are evaluated at $z = 0.23$, the median redshift of the maxBCG cluster sample. The units are 10^{14} M$_\odot$ for M_{500}, 10^{44} ergs s^{-1} for L_X, keV for T_X, and 10^{14} M$_\odot$ keV for Y_X, all the amplitude of $M-X$ relations are transfer to the values at unit pivot, while the pivot value of $M-N$ and $X-N$ relations is at richness = 40. A mass bias of $b = 0.25$ is included when derive the $X-N$ relation, through adding the amplitude of $M-N$ relation with a value of $\ln(1-b)$.

[a] $\sigma_{m|l} = 0.28$ is from Angulo et al. (2012).

[b] $\sigma_{m|t} = 0.20$ is from Kravtsov et al. (2006).

[c] The binned masses of Rozo et al. (2009a) and Simet et al. (2017) measure the $\ln\langle M|N\rangle$ rather than $\langle \ln M|N\rangle$, and $\tilde{a}_{m|n} = a_{m|n} + \frac{1}{2}\sigma^2_{m|n}$. We quote their $\tilde{a}_{m|n}$ and get $a_{m|n}$ by subtracting $\frac{1}{2}\sigma^2_{m|n}$.

[d] Simet et al. (2017) measure $M_{200m}-\lambda$ relation, we convert M_{200m} to M_{500c} with a typical ratio of 1.67. We note that the $X-N$ amplitude of 'all' sample is close to the model prediction with 25% mass bias, while the '*R*' sample is higher and the '$M+P$' sample is lower than the model. This may indicate that the dynamical state of cluster affects the level of mass bias.

derive a correction term of ~ -0.12, corresponding to a 11 per cent down offset for n or λ.

$\boldsymbol{Y_X-N}$ relation is also from the $Y_{SZ}-Y_X$ relation (Rozo et al. 2014a; M10 data set in their table 1) and the $Y_{SZ}-N$ relation (Rozo et al. 2014b; in their table 4) as

$$a_{x|n} = \left[\frac{a_{sz|n} - a_{sz|x}}{s_{sz|x}}\right] + \frac{\beta\sigma_{sz|x}}{s_{sz|x}}(r_{x,n|sz}\sigma_{sz|n} - \sigma_{sz|x}), \tag{A15}$$

$$s_{x|n} = \frac{s_{sz|n}}{s_{sz|x}}. \tag{A16}$$

We assume $r_{x,n|sz} = 0$. The scatter of Y_X-N relation is estimated using equation A14 of (Rozo et al. 2014a):

$$\sigma^2_{x|n} = s^2_{x|m}[\sigma^2_{m|n} + \sigma^2_{m|x} - 2r_{x,n|m}\sigma_{m|n}\sigma_{m|x}]. \tag{A17}$$

The resultant $\sigma_{x|n} = 0.69$ is very close to the $\sigma_{sz|n} = 0.70$ (Rozo et al. 2014b; in their table 4).

The binned Y_X presented as green bowties in Fig. 3 is $\langle \ln Y_X|N\rangle$, which is evaluated from $\ln\langle Y_X|N\rangle - \frac{1}{2}\sigma^2_{x|n}$ assuming a lognormal distribution. When we present the expected Y_X inferred from *Planck* stacking Y_{SZ}, we include additional corrections listed as below. (1) From stacking $\ln\langle Y_{SZ}|N\rangle$ (Planck Collaboration XII 2011; in their table 1) to $\langle \ln Y_{SZ}|N\rangle$, $-\frac{1}{2}\sigma^2_{sz|n} = -22$ per cent. (2) Aperture-induced correction due to covariance, as the *Planck* Y_{SZ} is measured within R_{500}, which is based on maxBCG N_{200}. From equation (A4) and replacing L_X with Y_{SZ}, the amplitude is overbiased with a factor of $r_{sz,n|m}\beta s_{sz|m}\sigma_{m|sz}\sigma_{m|n}$ with $r_{sz,n|m} = 0.47$ (Angulo et al. 2012), which is different from the case of Y_{SZ} measured from R_{500} independent of N_{200}. In that case, $r_{sz,n|m}$ should be 0 or even negative, suggested by a negative hot gas–galaxy correlation when we derive the predicted multivariate scaling relation. The amplitude bias of $r_{sz,n|m}\beta s_{sz|m}\sigma_{m|sz}\sigma_{m|n}$ is further divided by a factor of 2, as *Planck* measurements are template–amplitude fits rather than cylindrically integrated Y_{SZ} measurements and the inner radii weight more than the cylindrical integration. The final correction is -5 per cent (Rozo et al. 2014b). (3) Miscentring correction at a level of 10 per cent in the richness range we present (Biesiadzinski et al. 2012). (4) Eddington bias correction based on equation (A2), as we convert Y_{SZ} to Y_X using $Y_{SZ}-Y_X$ relation (Rozo et al. 2014a; M10 data set in their table 1; $\sigma_{x|sz} = \sigma_{sz|x}/s_{sz|x} = 0.15$) with a correction of $-\beta\sigma^2_{x|sz} = -6$ per cent. (5) Aperture-induced correction due to hydrostatic or weak-lensing mass bias is ignored, because it is typically small relative with the mass bias itself.

In summary, we list all the cited and derived relations in Table A1.

APPENDIX B: *Chandra* AND *XMM–Newton* CROSS-CALIBRATION

The X-ray data of our sample are from *Chandra* and *XMM–Newton*. There are some cross-calibration issues between these two instruments reported by the International Astronomical Consortium for High Energy Calibration (e.g. Schellenberger et al. 2015). We use six clusters in our sample with both the *Chandra* and *XMM–Newton* data and spanning a wide temperature range to do the in-house cross-calibration. We first measure the X-ray properties such as T_X, L_X (bolometric), and Y_X individually from six clusters and independently from *Chandra* and *XMM–Newton*. All the X-ray properties are derived with the same procedures as detailed in Section 3. Note that R_{500} can be different for the *Chandra* data and the *XMM–Newton* data because of the different temperatures.

Fig. B1 compares the temperatures from *Chandra* and *XMM–Newton*, with *Chandra* temperatures systematically higher than *XMM–Newton*'s. We then fit the relation with a power-law function at $T_X = 2$–10.5 keV as

$$\frac{kT_{\rm XMM}}{1\ {\rm keV}} = a_T\left(\frac{kT_{\rm Ch}}{1\ {\rm keV}}\right)^{s_T}, \quad a_T = 1.15 \pm 0.09,\ s_T = 0.83 \pm 0.04. \tag{B1}$$

Similarly, we have the power-law relations on L_X and Y_X:

$$\frac{L_{\rm XMM}}{10^{44}\ {\rm erg\ s^{-1}}} = a_L\left(\frac{L_{\rm Ch}}{10^{44}\ {\rm erg\ s^{-1}}}\right)^{s_L}, \quad a_L = 0.90 \pm 0.01,\ s_L = 0.98 \pm 0.01. \tag{B2}$$

$$\frac{Y_{\rm XMM}}{10^{14}\ {\rm M_\odot\ keV}} = a_Y\left(\frac{Y_{\rm Ch}}{10^{14}\ {\rm M_\odot\ keV}}\right)^{s_Y}, \quad a_Y = 0.98 \pm 0.08,\ s_Y = 0.84 \pm 0.04. \tag{B3}$$

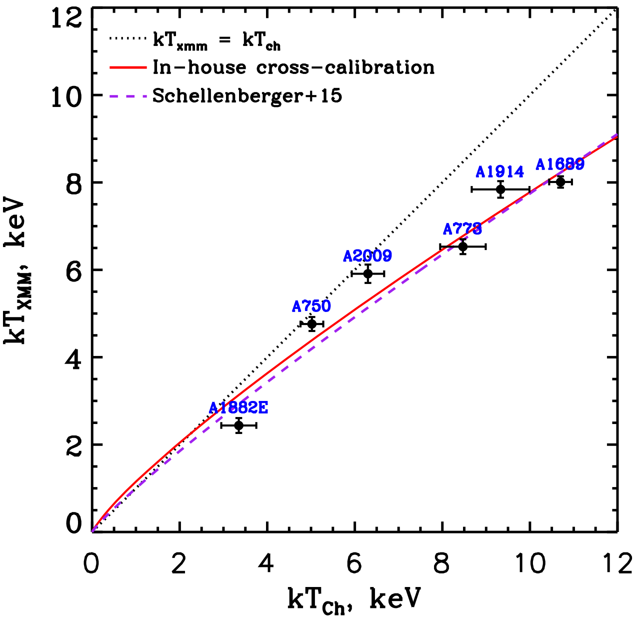

Figure B1. Comparison of temperatures from *Chandra* and *XMM–Newton* for six clusters in our sample. Temperatures are derived from 0.15 to 0.75 R_{500} for *Chandra* and *XMM–Newton*, respectively, with R_{500} determined iteratively from the $M-T_X$ relation. The dotted line is the line of equality. The red line is our power-law best-fitting relation (see Appendix B) used in this work, while the purple dashed line is from Schellenberger et al. (2015) ($a_T = 1.00$ and $s_T = 0.89$), with older *Chandra* and *XMM–Newton* calibrations than what we used.

APPENDIX C: FEATURES OF INDIVIDUAL CLUSTERS

Here, we briefly comment clusters in the sample (in order of the maxBCG richness N_{200}), with emphasis on the substructure and dynamical state.

A2142 is the richest maxBCG cluster. There is an ongoing merger as indicated by multiple cold fronts (Markevitch et al. 2000; Rossetti et al. 2013), radio haloes (Venturi et al. 2017), and group-scale substructures (Owers, Nulsen & Couch 2011). However, its global X-ray emission appears relaxed and both maxBCG and redMaPPer positions agree with each other.

J150 is the second richest maxBCG cluster. One may expect J150 as a massive ($\sim$1.1 × 10^{15} M$_\odot$) and high-temperature ($kT \sim 10$ keV) cluster inferred from its optical richness. However from the *XMM* data, we find a ∼5 Mpc filament interlaced with three $kT \sim 3$ keV clusters (J150E, J150M, and J150W) and one group (J150EE; $kT \sim 1.3$ keV). The maxBCG centre is on J150E. redMaPPer identifies a cluster centred on J150W and a group on J150EE. However, the *Planck* SZ centre is located at the X-ray brightest cluster (J150M). Thus, this large-scale filament suffers problems of miscentring and projection when selected as a single optical cluster. We then compare the optical richness inferred from T_X with the one in optical catalogue. Based on the mass derived from the X-ray temperature of different components, we estimate that the total maxBCG richness for three clusters and the group is 82 (versus 164 from maxBCG), and the total redMaPPer richness for three

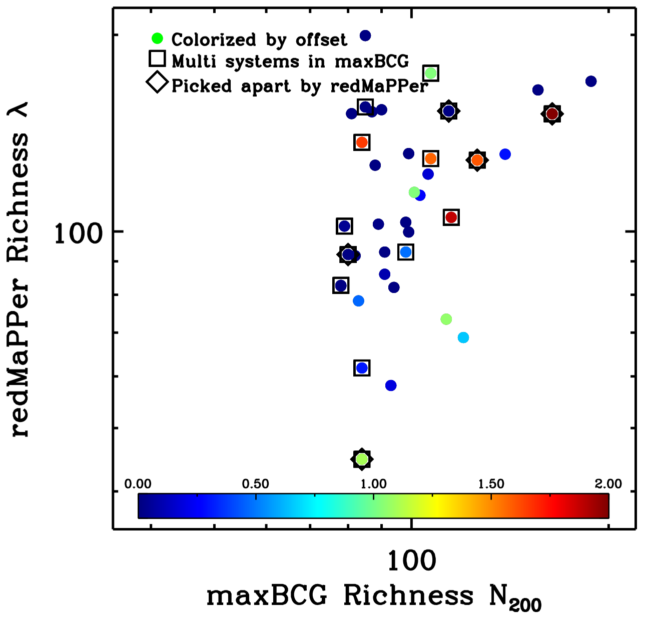

Figure C1. Comparison of richness from maxBCG and redMaPPer of the sample. The dots are colourized by the centroid offset between maxBCG and redMaPPer, in units of Mpc. The box shows the maxBCG cluster with multiple X-ray clumps. The diamond shows the same maxBCG cluster, but picked apart as multiclusters by redMaPPer. The clusters with larger centroid offset are more likely to be multicluster systems.

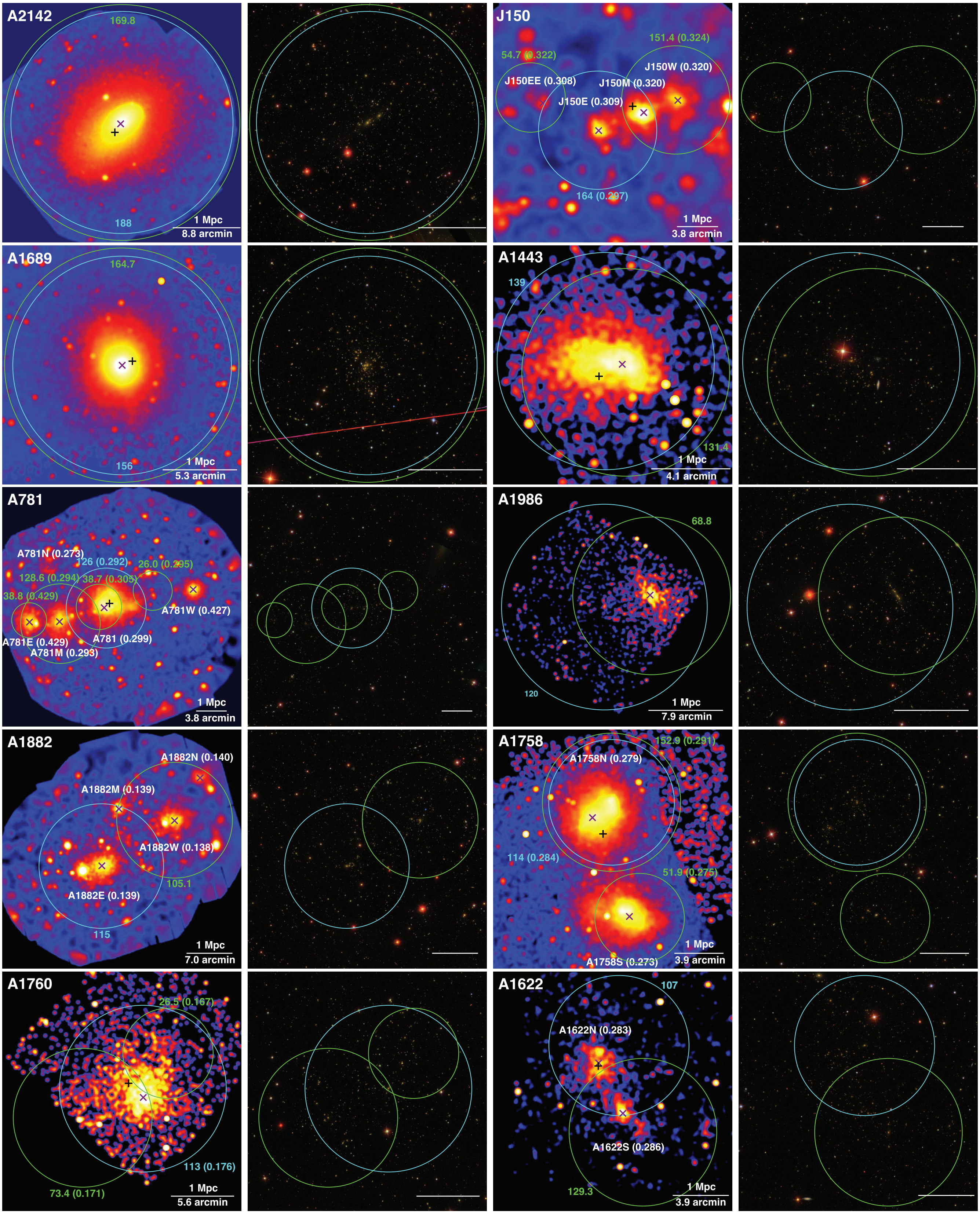

Figure C2. *Chandra* 0.7–2 keV/*XMM–Newton* 0.5–2 keV images of the 38 richest maxBCG clusters (in order of the maxBCG richness). All images are instrumental background subtracted, exposure corrected, and smoothed. The side-by-side images are *SDSS* RGB images with the same FOV. Purple × marks the X-ray peak, while black + marks the SZ centre from PSZ2 catalogue (Planck Collaboration XXVII 2016). The centres of cyan and green circles are from the maxBCG (Koester et al. 2007) and redMaPPer (Rykoff et al. 2014) catalogue, respectively, while the number marks the optical richness and the circles are R_{500} from the mass-richness relation (Rozo et al. 2009a for maxBCG and Simet et al. 2017 for redMaPPer). For multiple cluster systems, we also show the name and the spectroscopic redshift from *SDSS* for each X-ray cluster, as well as the photometric redshifts for the optical cluster from the optical catalogues.

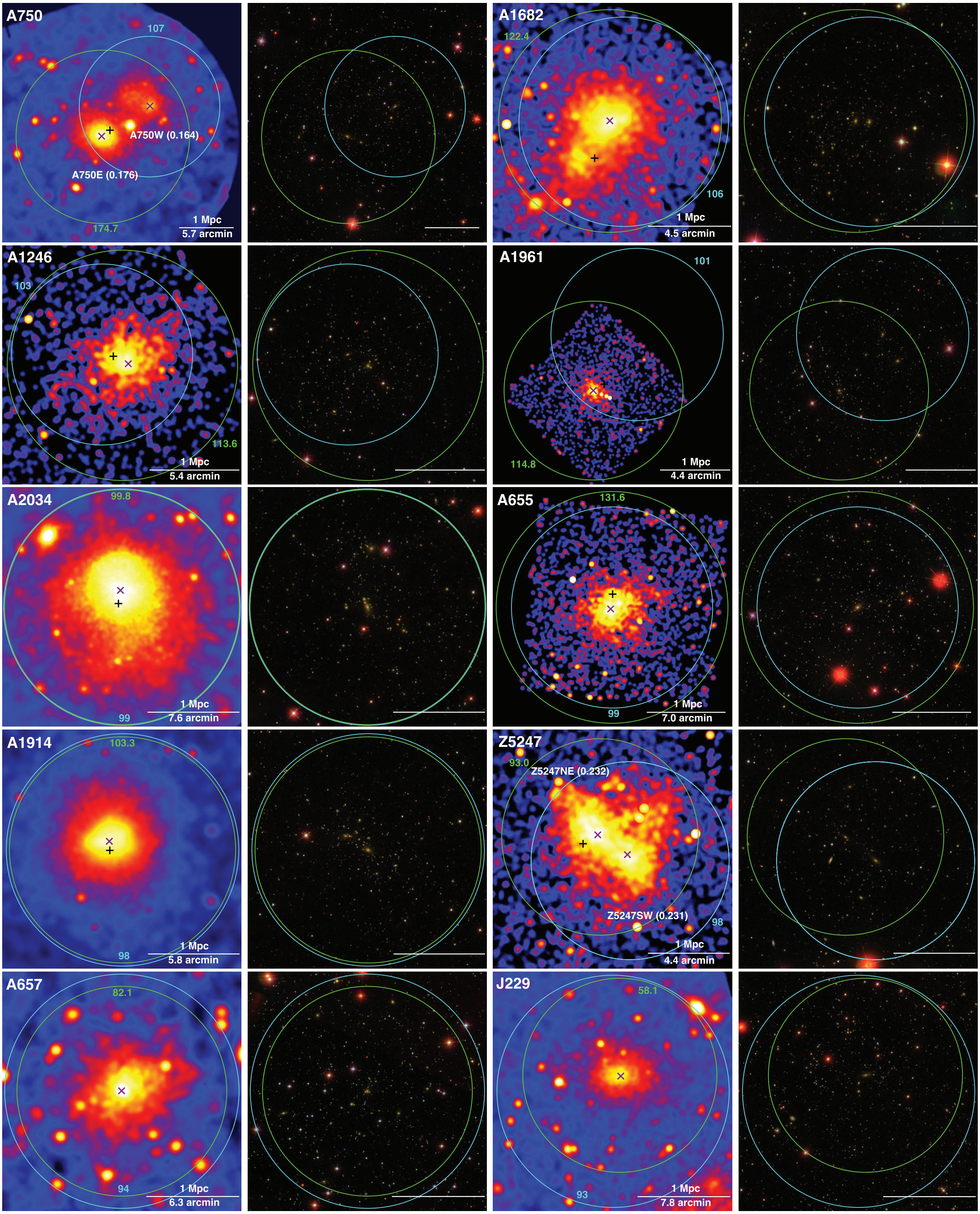

Figure C2 – *Continued*

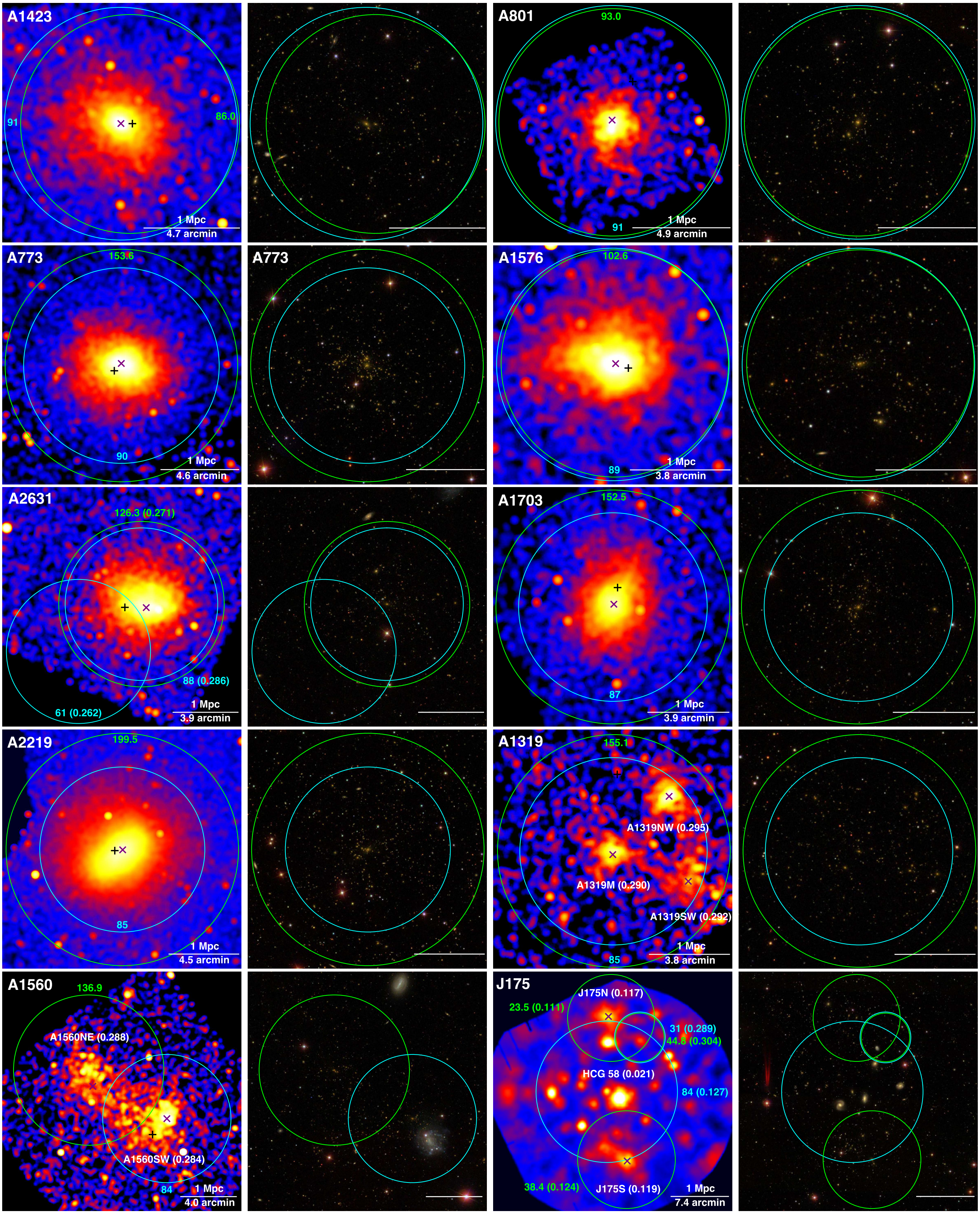

Figure C2 – *Continued*

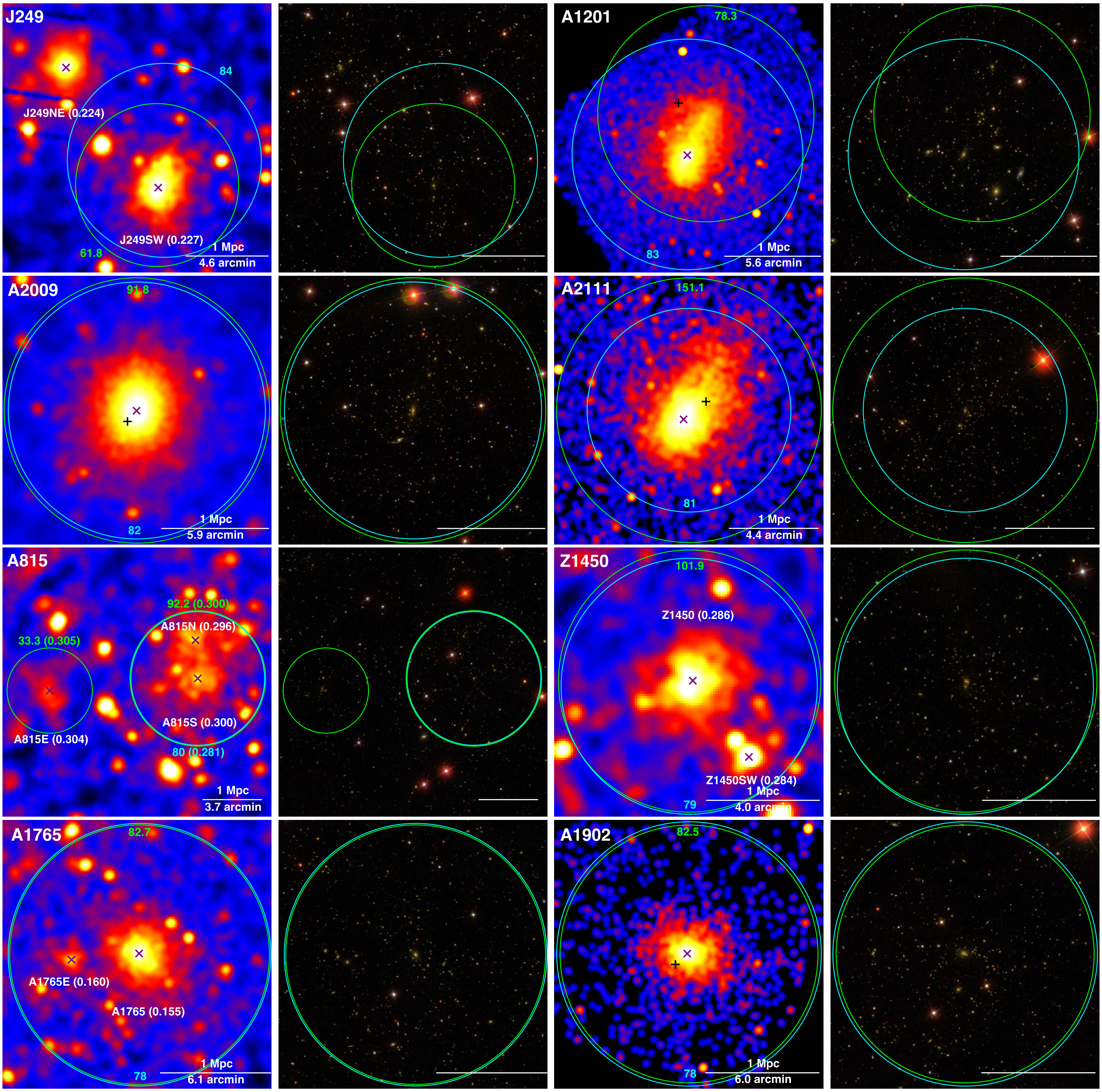

Figure C2 – *Continued*

clusters is 105.5 (versus 151.4 from redMaPPer). Thus, this system is a particularly rich one in optical. Less X-ray luminous groups in this large-scale structure and projection contamination may also contribute.

A1689 has a large concentration parameter and its gravitational lensing mass is higher than the X-ray mass by a factor of 2, which could be explained by its triaxial shape with the major axis nearly orientated along the line of sight (e.g. Morandi, Pedersen & Limousin 2011; Sereno et al. 2013; Umetsu et al. 2015). The triaxiality may be induced by a merger along the line of sight as indicated by the asymmetric temperature distribution (Andersson & Madejski 2004) and the diffuse radio emission (Vacca et al. 2011). However in the plane of sky, A1689 appears relaxed based on the X-ray morphology. The X-ray peak is also consistent with both maxBCG and redMaPPer centres.

A1443 is a merging cluster from the X-ray E–W elongation and the diffuse radio emission (Bonafede et al. 2015). maxBCG and redMaPPer selected different galaxies as the BCG of the cluster.

A781 in fact contains two large filaments projected on the sky, one ∼2 Mpc filament (A781, A781M, A781N; $z \sim 0.29$) and another ∼7 Mpc filament (A781E, A781W; $z \sim 0.43$). All these five clusters, including the X-ray faint A781N ($kT = 1.3$ keV, $L_X = 1.6 \times 10^{43}$ erg s^{-1} in a 2 arcmin radius), are confirmed by weak-lensing detections (Sehgal et al. 2008). A radio halo and a candidate radio relic was also discovered to be associated with the hottest cluster A781 (Govoni et al. 2011). The maxBCG only detects the brightest A781, while the redMaPPer detects A781E but mixes A781 with A781M.

A1986 is not detected by *Planck*. The redMaPPer centre is on the X-ray peak. The mass estimated from the redMaPPer richness ($\sim 3.9 \times 10^{14}\,\mathrm{M}_{\odot}$) is more consistent with the mass suggested by the X-ray temperature ($\sim 2.4 \times 10^{14}\,\mathrm{M}_{\odot}$). The maxBCG centre is ~640 kpc from the X-ray centre and the maxBCG richness is biased high by ~2.

A1882 is in a ~3 Mpc filament at $z \sim 0.14$ with at least four groups (A1882E, A1882M, A1882W, A1882N, $kT \sim 1.4$–2.5 keV), which are corresponding to four concentrations of galaxy distribution (Owers et al. 2013). The merger is considered to be at an early stage from the large projected separation and the dearth of evidence for a recent major interaction in X-ray data (Owers et al. 2013). The maxBCG and redMaPPer centres are on A1882E and A1882W, respectively. There is also no *Planck* SZ detection in this region.

A1758 is composed of A1758N and A1758S separated by ~2 Mpc in projection with similar redshift ($z \sim 0.28$) and X-ray luminosity. A1758N is in the late stage of a merger of two 7 keV subclusters near the plane of sky, while A1758S is in the early stage of a merger of two 5 keV subclusters close to the line of sight from detailed multiwavelength observations (e.g. David & Kempner 2004; Giovannini et al. 2009; Durret, Laganá & Haider 2011; Monteiro-Oliveira et al. 2017). These two systems are most likely gravitationally bound and will eventually merge into a ~12 keV cluster (David & Kempner 2004). The X-ray mass of A1758N and A1758S is 11.7×10^{14} and $4.7 \times 10^{14}\,\mathrm{M}_{\odot}$, respectively, which can be compared with the masses suggested by their redMaPPer richness values ($11.1 \times 10^{14}\,\mathrm{M}_{\odot}$ for A1758N and $2.7 \times 10^{14}\,\mathrm{M}_{\odot}$ for A1758S). The maxBCG only identifies A1758N with an optical mass of $7.2 \times 10^{14}\,\mathrm{M}_{\odot}$.

A1760 is divided into two clusters by redMaPPer and the total richness of these two clusters is compatible with maxBCG.

A1622 is composed of two clusters as shown in the X-ray image. Only the X-ray properties of the northern one can be constrained from the shallow *Chandra* data. The maxBCG and redMaPPer identified different galaxies as the BCG. Optical richness values from both maxBCG and redMaPPer suggest mass values of 2–3 times higher than the mass estimated from the X-ray temperature.

A750 is composed of A750E/A750W at $z \sim 0.16$. A750W is possibly falling into A750E, as indicated by the highly disturbed X-ray morphology and the large offset of two X-ray peaks. Both maxBCG and redMaPPer mix these two clusters as one with different galaxies as the BCG.

A1682 is a merging cluster as shown by the disturbed X-ray morphology. Both galaxy distribution and weak-lensing mass map show two peaks (Dahle et al. 2002) coincident with the X-ray peaks. The diffuse radio emission is complex with possibly one halo and two relics (e.g. Macario et al. 2013). The maxBCG and redMaPPer identified different galaxies as the BCG.

A1246 has been observed to the virial radius with *Suzaku* (Sato et al. 2014). The X-ray mass within R_{500} ($5.5 \times 10^{14}\,\mathrm{M}_{\odot}$) from the $M-T_X$ relation of the *Chandra* data is consistent with the mass derived from the *Suzaku* data ($4.3 \times 10^{14}\,\mathrm{M}_{\odot}$; Sato et al. 2014). The maxBCG and redMaPPer identified different galaxies as the BCG.

A1961 is a poor cluster ($2.7 \times 10^{14}\,\mathrm{M}_{\odot}$) from its X-ray temperature. However, the optical richness values from maxBCG and redMaPPer suggest mass values of ~3 times higher. The maxBCG and redMaPPer identified different galaxies as the BCG. The miscentring of maxBCG is large, ~1 Mpc.

A2034 is a merging cluster with the merger axis along S–N, as indicated by the northern shock (Owers et al. 2014) and the diffuse radio emission (Giovannini et al. 2009). The complex dynamics is also shown by the galaxy distribution (Owers et al. 2014) and weak-lensing mass distribution (Okabe & Umetsu 2008). The maxBCG and redMaPPer results are very similar.

A655 hosts a dominated cD galaxy at the centre, identified by both maxBCG and redMaPPer. However, the optical richness values from maxBCG and redMaPPer suggest mass values of 2–3 times higher than that from the X-ray temperature. The tentacle-like outskirts based on the X-ray morphology may suggest connection with other large-scale filaments and infalling galaxy groups as also suggested by Patej & Loeb (2016).

A1914 is a merging cluster as indicated by two substructures along the NE–SW direction from the galaxy distribution (Barrena, Girardi & Boschin 2013) and the weak-lensing mass distribution (Okabe & Umetsu 2008), as well as the diffuse radio emission (Bacchi et al. 2003). Both maxBCG and redMaPPer identified the same BCG, 175 kpc south to the X-ray peak.

Z5247 is a merging cluster as indicated by the disturbed X-ray morphology. There are two X-ray peaks, corresponding to two substructures from the galaxy distribution and the weak-lensing mass distribution (Dahle et al. 2002). The cluster also hosts a radio relic and a candidate radio halo (Kale et al. 2015). The maxBCG and redMaPPer identified different galaxies as the BCG.

A657 is not detected by *Planck*. While maxBCG and redMaPPer identified the same BCG close to the X-ray peak, the optical richness values from maxBCG and redMaPPer suggest mass values of 3–4 times higher than that from the X-ray temperature.

J229 is a ~2.5 keV system without *Planck* SZ detection. The maxBCG and redMaPPer identified different galaxies as the BCG. The optical richness values from maxBCG and redMaPPer suggest mass values of 2–4 times higher than that from the X-ray temperature.

A1423 is a relaxed cluster from the smooth X-ray morphology, galaxy distribution, and weak-lensing mass distribution (Dahle et al. 2002).

A801 appears relaxed in X-rays. The richness values from maxBCG and redMaPPer suggest a cluster mass similar to that derived from its X-ray temperature.

A773 is a merging cluster as shown by evidence such as two X-ray peaks in the centre along NE–SW, asymmetric X-ray temperature distribution (Govoni et al. 2004), two peaks of galaxy distribution and their velocity distribution (Barrena et al. 2007), two peaks in the weak-lensing mass distributions (Dahle et al. 2002), and diffuse radio emission (Govoni et al. 2001).

A1576 is a disturbed cluster as indicated by its lopsided X-ray morphology and the presence of multiple peaks in galaxy distribution and weak-lensing mass distribution (Dahle et al. 2002).

A2631 is classified as a disturbed cluster based on the multiple morphology parameters (Cassano et al. 2010). However, the weak-lensing map shows only one single peak (Okabe et al. 2010) and there is not any significant extended radio emission (Venturi et al. 2008).

A1703 is a relaxed, unimodal cluster from the strong-lensing model (Richard et al. 2009).

A2219 is a merging cluster with the main merger axis along NW–SE direction, as shown by a series shocks and a possible cold front (Canning et al. 2017), two luminous BCGs in the cluster centre and substructure in galaxy distribution (Boschin et al. 2004), and a diffuse radio halo (Bacchi et al. 2003).

A1319 is composed of three clusters at similar redshifts, A1319M, A1319NW, and A1319SW. Both maxBCG and redMaPPer only identified one cluster centred on the BCG of A1319M.

A1560 is a merging cluster with two subclusters. The maxBCG centre is on A1560SW and the redMaPPer centre is on A1560NE.

J175 is in a complex field with both foreground and background sources. X-ray emission mainly shows two clusters separated by $\sim$2.5 Mpc in projection along the N–S direction, J175N at $z_{\rm spec} = 0.117$ and J175S at $z_{\rm spec} = 0.119$, which is also confirmed by redMaPPer. However, the maxBCG cluster is centred on a luminous galaxy at $z_{\rm spec} = 0.117$, nearly mid-way between J175N and J175S. The maxBCG centre is very close to the foreground galaxy group HCG 58 (Hickson 1982; Freeland, Stilp & Wilcots 2009) that hosts the brightest X-ray source in the *XMM* field, the X-ray AGN of NGC 3822. The X-ray diffuse luminosity from this maxBCG region, excluding the NGC 3822 AGN, other point sources and diffuse emission from J175N/J175S is only $\sim 3.3 \times 10^{42}$ erg s^{-1} at the maxBCG cluster redshift. There is no *Planck* SZ source in this region. There is also a faint background cluster at $z_{\rm spec} = 0.280$, detected by both maxBCG and redMaPPer.

J249 is a cluster pair with comparable X-ray temperatures. There is no *Planck* SZ detection. The maxBCG and redMaPPer identified different galaxies in J249SW as the BCG.

A1201 is a merging cluster with the merger axis along NW–SE from the lopsided X-ray morphology. It hosts two cold fronts and an offset remnant core with a stripped tail (Ma et al. 2012). Substructures are identified from the spatial and velocity distribution of member galaxies (Owers et al. 2009). The strong-lensing arc indicates large mass elongation (Edge et al. 2003). maxBCG and redMaPPer identified different galaxies as the BCG.

A2009 appears relaxed from the smooth X-ray morphology.

A2111 is a merging cluster from the lopsided X-ray morphology and early X-ray observations (Wang, Ulmer & Lavery 1997; Henriksen, Wang & Ulmer 1999). The member galaxy and weak-lensing mass distribution shows the same elongation as the X-rays in the NW–SE direction (Dahle et al. 2002). No significant diffuse radio emission has been detected (Venturi et al. 2008).

A815 is a merging cluster from the disturbed X-ray morphology. A815N and A815S each has a luminous galaxy on its X-ray peak. There is no *Planck* SZ detection. There is a group to the east detected by the redMaPPer at a similar redshift. Based on the *XMM* data, we derive $kT = 1.1$ keV and $L_X = 1.5 \times 10^{43}$ erg s^{-1} within R_{500} that is determined from redMaPPer's richness.

Z1450 is a merging cluster from the asymmetric X-ray morphology. The maxBCG and redMaPPer identified different galaxies as the BCG. The optical richness values from maxBCG and redMaPPer suggest mass values of 5–7 times higher than that from the X-ray temperature. There is no *Planck* SZ detection. There is a group to the SW at a similar redshift. The X-ray emission of this region is dominated by two bright point sources, which leaves insufficient amount of data to constrain the gas properties of this group.

A1765 is a low-mass ($1.0 \times 10^{14}\,{\rm M}_{\odot}$) cluster from the X-ray temperature ($kT = 2.2$ keV), while the optical richness values from maxBCG and redMaPPer suggest mass values of $\sim$5 times higher. There is no *Planck* SZ detection. There is a group to the east at a similar redshift. Based on the *XMM* data, we derive $kT = 1.5$ keV, $L_X = 4.3 \times 10^{42}$ erg s^{-1} within a 1.5 arcmin radius.

A1902 appears relaxed. Masses from optical and X-ray are consistent.

APPENDIX D: LARGE-SCALE STRUCTURES

We also searched for large-scale structures around each maxBCG cluster in our sample, by looking at the maxBCG and redMaPPer catalogues, and examining *Chandra* and *XMM–Newton* data archives in a 1 deg radius (corresponding to $\sim$13 Mpc at $z = 0.23$). Besides merging clusters like J150, A781, A1882, and A1319 as detailed in Appendix C, there are two additional large-scale structures, as shown in Fig. D1.

J229-LS contains at least four clusters (A2050, J229, A2051, A2051S; $z \sim 0.12$) in a $\sim$10 Mpc filament. The X-ray properties of these clusters are also listed in Table 3. Note that A2051N is a background cluster at $z \sim 0.38$. Interestingly, the poor cluster J229 is identified as a rich optical cluster, while the hotter cluster A2051 is missed by both maxBCG and redMaPPer.

A2219-LS is composed of both A2219 and J249 in a $\sim$6 Mpc filament at $z \sim 0.23$. Both clusters are in this sample with similar maxBCG richness, while J249 is actually composed of two subclusters from the X-ray data.

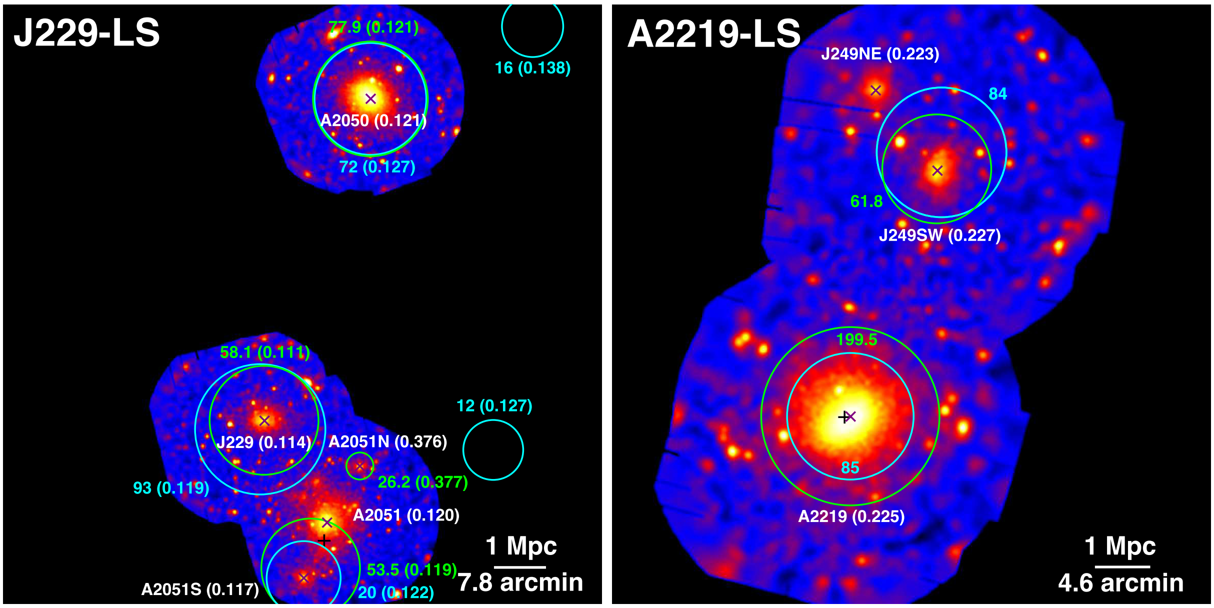

Figure D1. Large-scale structures identified around maxBCG clusters in our sample (the marks are the same as in Fig. C2).

This paper has been typeset from a TEX/LATEX file prepared by the author.