Testing for X-Ray–SZ Differences and Redshift Evolution in the X-Ray Morphology of Galaxy Clusters

The Astrophysical Journal, 841:5 (11pp), 2017 May 20 https://doi.org/10.3847/1538-4357/aa6db4

Testing for X-Ray–SZ Differences and Redshift Evolution in the X-Ray Morphology of Galaxy Clusters

D. Nurgaliev[1], M. McDonald[2], B. A. Benson[3,4,5], L. Bleem[5,6], S. Bocquet[5,6], W. R. Forman[7], G. P. Garmire[8], N. Gupta[9,10,11], J. Hlavacek-Larrondo[12], J. J. Mohr[9,10,11], D. Nagai[13,14], D. Rapetti[15,16], A. A. Stark[7], C. W. Stubbs[1,7], and A. Vikhlinin[7]
[1] Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA; nurgaliev@physics.harvard.edu
[2] MIT Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
[3] Fermi National Accelerator Laboratory, Batavia, IL 60510-0500, USA
[4] Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA
[5] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[6] Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA
[7] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[8] Huntingdon Institute for X-Ray Astronomy, LLC, USA
[9] Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany
[10] Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany
[11] Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. D-85748 Garching, Germany
[12] Département de Physique, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, Québec H3C 3J7, Canada
[13] Department of Physics, Yale University, New Haven, CT 06520, USA
[14] Yale Center for Astronomy and Astrophysics, Yale University, New Haven, CT 06520, USA
[15] Center for Astrophysics and Space Astronomy, Department of Astrophysical and Planetary Science, University of Colorado, Boulder, C0 80309, USA
[16] NASA Ames Research Center, Moffett Field, CA 94035, USA
Received 2016 September 1; revised 2017 April 13; accepted 2017 April 16; published 2017 May 16

Abstract

We present a quantitative study of the X-ray morphology of galaxy clusters, as a function of their detection method and redshift. We analyze two separate samples of galaxy clusters: a sample of 36 clusters at $0.35 < z < 0.9$ selected in the X-ray with the *ROSAT* PSPC 400 deg^2 survey, and a sample of 90 clusters at $0.25 < z < 1.2$ selected via the Sunyaev–Zel'dovich (SZ) effect with the South Pole Telescope. Clusters from both samples have similar-quality *Chandra* observations, which allow us to quantify their X-ray morphologies via two distinct methods: centroid shifts (w) and photon asymmetry ($A_{\rm phot}$). The latter technique provides nearly unbiased morphology estimates for clusters spanning a broad range of redshift and data quality. We further compare the X-ray morphologies of X-ray- and SZ-selected clusters with those of simulated clusters. We do not find a statistically significant difference in the measured X-ray morphology of X-ray and SZ-selected clusters over the redshift range probed by these samples, suggesting that the two are probing similar populations of clusters. We find that the X-ray morphologies of simulated clusters are statistically indistinguishable from those of X-ray- or SZ-selected clusters, implying that the most important physics for dictating the large-scale gas morphology (outside of the core) is well-approximated in these simulations. Finally, we find no statistically significant redshift evolution in the X-ray morphology (both for observed and simulated clusters), over the range of $z \sim 0.3$ to $z \sim 1$, seemingly in contradiction with the redshift-dependent halo merger rate predicted by simulations.

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

1. Introduction

Large-scale galaxy cluster surveys can make an important contribution to understanding the growth of structure in the universe, delivering precise constraints on the nature of dark matter and dark energy, and providing insights into astrophysical processes in clusters. The primary interest in studying galaxy clusters from the cosmological point of view is in measuring their abundance as a function of mass and redshift, though alternative approaches to cluster-based cosmology that do not rely on precise masses have recently been proposed (Caldwell et al. 2016; Ntampaka et al. 2016; Pierre et al. 2016). The abundance of galaxy clusters as a function of mass and redshift currently provides constraints on cosmological models and parameters, most importantly matter density $\Omega_{\rm M}$ and the normalization of the matter power spectrum σ_8 (see the review by Allen et al. 2011). There are many subtleties, however, in interpreting abundance information from cluster surveys. First, cluster masses are a necessary ingredient for cosmological analyses based on cluster abundances. The most accurate masses are currently provided by weak lensing studies (e.g., Applegate et al. 2014; Hoekstra et al. 2015), with the most careful analyses achieving <5% systematic uncertainties. When weak lensing mass estimates for all clusters are not available, the absolute mass calibration from weak lensing is then typically tied to a relative mass calibration coming from the selection observable (e.g., X-ray luminosity, Sunyaev–Zeldovich (SZ) signal, optical richness, etc.). Both scatter and potential biases in the weak lensing analysis or cluster scaling relations can have an effect on constraints when fitting cosmological models. Second, one needs to understand the survey's completeness and purity. Finally, more subtle selection effects, such as increased sensitivity to a particular sub-class of clusters, may play a role. An example of such a bias would be an increased sensitivity of X-ray flux-limited samples to cool core clusters: because cool core systems have higher X-ray luminosity than non-cool core systems of the same mass, a flux-limited sample can potentially be biased toward cool core clusters (Hudson et al. 2010; Eckert et al. 2011; Mittal et al. 2011). The ratio of cool core to non-cool core systems at different redshifts is currently a subject of active research (Vikhlinin 2006; Santos et al. 2010; McDonald 2011;

Samuele et al. 2011; Semler et al. 2012; McDonald et al. 2013), so this bias effectively limits our understanding of completeness in the X-ray flux-limited samples.

Many of the biases implicit in the selection of various galaxy cluster samples are well understood. X-ray luminosity is proportional to gas density squared, so X-ray detection is biased toward cool core systems that have high central densities. In contrast, the majority of the SZ signal originates from outside of the core. Consequently, SZ detection is biased toward the large-scale gas properties in the cluster. Because both X-ray and SZ detection methods are based on the physical properties of the ICM, they may have some common biases (Angulo et al. 2012; Maughan et al. 2012; Lin et al. 2015) that are completely different from the optical detection methods which are sensitive to a different component of galaxy clusters (i.e., the galaxies). The finer details of each detection method's sensitivity to specific cluster morphology or dynamical state are not well understood.

It has been suggested that SZ-selected clusters are more often "morphologically disturbed" (i.e., ongoing mergers) than their X-ray-selected counterparts. This line of reasoning stems from (1) the presence of spectacular mergers among the first discovered by the SZ effect, such as El Gordo (Menanteau et al. 2012) and PLCKG214.6+37.0 (Planck Collaboration et al. 2013a); and (2) an extensive discussion of newly discovered mergers in the papers, which originated from the *Planck XMM-Newton* follow-up program (Planck Collaboration et al. 2011, 2012, 2013b). The latter program targeted 51 cluster candidates and led to the confirmation of 43 candidates, two of them being triple systems and four double systems. The 37 remaining objects had (1) lower X-ray luminosity than expected from scaling relations and (2) shallower density profiles than the mean density profiles of X-ray detected clusters. These two observations served as the main arguments for *Planck's* increased sensitivity to mergers. Indeed, recent studies have shown that *Planck* clusters are, on average, more morphologically disturbed (Rossetti et al. 2016) and have a lower occurrence rate of cool cores (Jones et al. 2015) when compared to X-ray-selected clusters.

While systems discovered by *Planck* do have interesting morphological properties, the aforementioned findings do not necessarily indicate an inherent sensitivity of the SZ effect to merging clusters. The lower central density and luminosity of clusters may be related to greater than previously thought intrinsic scatter in these parameters, or factors other than merging processes. Several of the double and triple systems discovered by *Planck* are clusters overlapping in projection, rather than interacting systems (though they still belong to the same supercluster structure). The increased sensitivity of *Planck* to such multiple systems is unsurprising due to its large beam size and consequent inability to resolve multiple systems.

Another question that has been extensively discussed in the literature is whether cluster morphology depends on redshift. The motivation for these studies is the connection of merger rate (and consequently morphology) to the mean matter density Ω_M. The morphology–cosmology connection that was analytically developed by Richstone et al. (1992) and then confirmed in simulations by Evrard et al. (1993) and Jing et al. (1995) predicted that clusters in low Ω_M models are much more regular and spherically symmetric than those in $\Omega_M = 1$ models. Consequently, there were efforts to constrain Ω_M by finding the fraction of clusters with significant level of substructure as defined by various substructure statistics: Mohr et al. (1995) used centroid shifts, Buote & Tsai (1995) used power ratios, and Schuecker et al. (2001) used a trio of tests which quantify mirror symmetry, azimuthal symmetry, and radial elongation. This approach of constraining cosmological parameters through substructure rates has not been as successful as other cluster-based cosmology tests owing to difficulties in robustly defining "significant levels of substructure," connecting observable substructure measures to theoretical merger definitions (Buote & Tsai 1995), and insufficiently low numbers of observed and simulated clusters for these tests.

Modern halo abundance measurements provide much more precise constraints on Ω_M than those obtained by merger fraction studies. Nevertheless, the question of substructure evolution in galaxy clusters is still relevant. The majority of studies have reported statistically significant evolution in cluster morphology (Jeltema et al. 2005; Andersson et al. 2009; Mann & Ebeling 2012); while a smaller number (e.g., Weißmann et al. 2013; Mantz et al. 2015) arrived at the conclusion that clusters at low- and high-z are consistent with no morphological evolution. Weißmann et al. (2013) performed a study of substructure evolution similar to ours, which is described later, using a slightly different cluster sample and substructure statistics, but arriving at similar results (See Section 5 for more details).

Our objectives in this paper are to test for any evidence of a difference in dynamical state between X-ray and SZ-selected clusters, low-z and high-z clusters, and observed and simulated clusters. The difference between X-ray and SZ-selected samples is of particular interest if we wish to combine the X-ray and SZ samples in order to obtain better statistics for various studies of cluster properties. In Section 2, we describe the three cluster samples used in this paper, from the South Pole Telescope (SPT), *ROSAT* PSPC 400 deg^2 survey, and from numerical simulations. In Section 3, we describe our methodology for quantifying X-ray morphology and the various tests that we will perform. In Section 4, we will discuss results of these tests, focusing on the key questions of whether or not X-ray- and SZ-selected clusters are statistically different in terms of their X-ray morphology, whether simulated and real clusters have statistically different morphology, and whether there is any measurable redshift evolution in X-ray morphology. In Section 5, we will discuss these results, placing them in context of previous work and considering their implications. We will conclude in Section 6 with a brief summary and look forward to future studies.

Throughout this work, we assume a flat ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_M = 0.27$.

2. The Data

2.1. Observations

The basis of this study is a subsample of 90 galaxy clusters, drawn from the larger sample of 516 galaxy clusters in the 2500 deg^2 SPT survey of Bleem et al. (2015). These 90 clusters, which were among the most massive of the SPT-selected clusters, all have uniform-depth *Chandra* observations, as summarized in McDonald et al. (2013, 2014). X-ray observations of these clusters were obtained primarily via a *Chandra* X-ray Visionary Project (PI B. Benson). Clusters in the SPT sample span the redshift range of $0.25 < z < 1.2$ and the mass

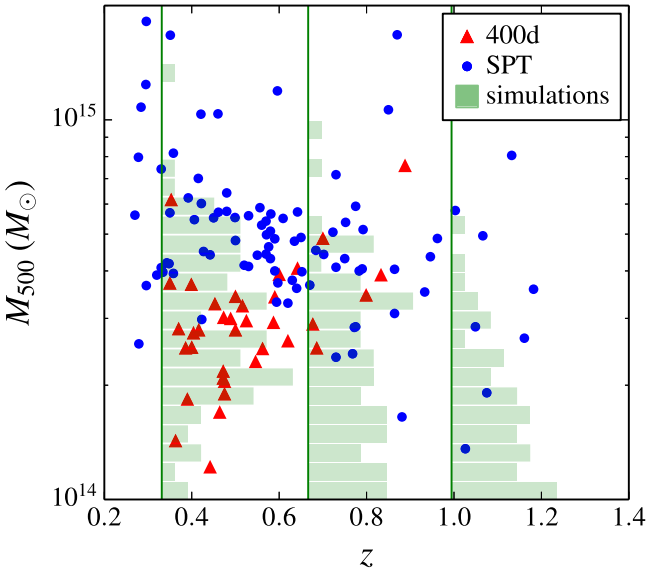

Figure 1. Masses and redshifts for SPT, 400d, and simulated samples. Red and blue dots show individual objects in the 400d and SPT samples, respectively. Green histograms show the distributions of masses for simulated clusters in three arbitrarily chosen redshifts bins—there are a total of 85 simulated clusters used for comparison to the data. The SPT and 400d samples have a reasonable overlap for $z < 0.6$; simulated clusters have a reasonable overlap with the observed clusters at all redshifts. We will show in Section 4.2 that the slightly different mass range for these samples should not significantly bias the distribution of observed morphologies.

range $2 \times 10^{14}\,M_\odot \lesssim M_{500} \lesssim 2 \times 10^{15}\,M_\odot$, where M_{500} is the total mass within r_{500}, and r_{500} is the radius within which the average enclosed density is 500 times the critical density. M_{500} here is derived from the Y_X–M relation, following Andersson et al. (2011). The median redshift and M_{500} for this sample are 0.59 and $4.6 \times 10^{14}\,M_\odot$ respectively.

We use the high-z part of the *ROSAT* PSPC 400 deg^2 cluster survey (Burenin et al. 2007), abbreviated hereafter as 400d, for our X-ray-selected sample. This sample consists of 36 clusters in the redshift range of $0.35 < z < 0.9$ and the mass range of $10^{14}\,M_\odot < M_{500} < 5 \times 10^{14}\,M_\odot$. The median redshift and mass of this sample are 0.48 and $2.6 \times 10^{14}\,M_\odot$ respectively. The masses (M_{500}) for these clusters were determined in the same way as for the SPT-selected clusters, using the same pipeline and scaling relations. These X-ray-selected clusters have a distinct lack of strong central density cusps at $z > 0.5$ (Vikhlinin et al. 2007), similar to what is observed in SPT-selected clusters (McDonald et al. 2013).

Figure 1 shows the distribution of SPT and 400d clusters in the (z, M_{500}) plane. There are fewer clusters at $z > 0.6$ in the 400d sample, due to the fact that it is flux limited. For a fair comparison (which would be free of redshift evolution effects) we will compare morphologies in the $z < 0.6$ subsamples for both catalogs—we will return to this point later when we define comparison samples. In this low-z subsample, the median masses of the 400d and SPT samples are 2.8 and $5.3 \times 10^{14}\,M_\odot$, respectively. This figure highlights the overlap in mass and redshift between the two observational samples, along with the simulated clusters that will be described below.

Both samples have similar-quality *Chandra* observations. Exposures are typically sufficient to obtain ~1500–2000 X-ray source counts (see McDonald et al. 2014). The high-resolution *Chandra* imaging with sufficient photon statistics is crucial to detect substructure in galaxy clusters (Nurgaliev et al. 2013). The 400d and SPT samples are currently among the best available samples of high-redshift clusters with clear selection criteria and high-quality X-ray follow-up. Significant overlap in the redshift and mass ranges allow us to directly compare clusters in these two samples.

The X-ray data reduction steps for both samples are equivalent to those described in Vikhlinin et al. (2009), Andersson et al. (2011), and McDonald et al. (2013). Using CIAO v4.7 and CALDB v4.7.1, we first filter data for flares, before applying the latest calibration corrections. Point sources are identified using an automated routine following a wavelet decomposition technique (Vikhlinin et al. 1998) and then visually inspected. Cluster centers are chosen to be the brightest pixel after convolution with a Gaussian kernel with $\sigma = 40$ kpc, following Nurgaliev et al. (2013).

2.2. Simulations

We analyze mock *Chandra* observations of massive galaxy clusters extracted from the *Omega500* simulation (Nelson et al. 2014). Below, we briefly summarize the main elements of the *Omega500* simulation and mock *Chandra* observations of simulated clusters, and refer the readers to Nelson et al. (2014) and Nagai et al. (2007), respectively, for further details.

The *Omega500* simulation is a large cosmological hydrodynamic simulation performed with the Adaptive Refinement Tree (ART) N-body+hydrodynamics code (Kravtsov 1999; Kravtsov et al. 2002; Rudd et al. 2008). In order to achieve the dynamic ranges necessary to resolve the cores of massive halos, adaptive refinement in space and time and non-adaptive refinement in mass (Klypin et al. 2001) are used. The simulation has a co-moving box length of $500\,h^{-1}$ Mpc and a maximum co-moving spatial resolution of $3.8\,h^{-1}$ kpc (where $h \equiv 0.01\,H_0$) and is performed in a flat ΛCDM model with the *WMAP* five-year cosmological parameters (Komatsu et al. 2009). The simulations include gravity, collisionless dynamics of dark matter and stars, gas dynamics, star formation, metal enrichment, SN feedback, advection of metals, metallicity-dependent radiative cooling, and UV heating due to a cosmological ionizing background. Some relevant physical processes, including AGN feedback, magnetic fields, and cosmic rays were not included.

For each cluster with $M_{500} > 10^{14}\,M_\odot$, of which there are 85, the regions within $5 \times r_{\rm vir}$ are re-simulated with high spatial and mass resolution using the multiple mass resolution technique. The resulting simulation has 2048^3 spatial elements, allowing a corresponding dark matter particle mass of $1.09 \times 10^9\,h^{-1}\,M_\odot$ and maximum physical resolution of 5.4 kpc. Herein we analyze the simulated galaxy clusters at expansion factors $a = 0.5014, 0.6001, 0.7511, 0.9764$ corresponding to redshifts $z = 0.99, 0.66, 0.33, 0.02$. The distribution of masses of simulated clusters at each of these epochs is shown on Figure 1.

For each simulated cluster, we analyze mock *Chandra* data. We first generate X-ray flux maps for each of the simulated clusters. We compute the X-ray emissivity, using the MEKAL plasma code (Mewe et al. 1986; Kaastra & Mewe 1993; Liedahl et al. 1995) and the solar abundance table from Anders & Grevesse (1989). We multiply the plasma spectrum by the photoelectric absorption corresponding to a hydrogen column density of $N_{\rm H} = 2 \times 10^{20}\,{\rm cm}^{-2}$. We then convolve the emission spectrum with the response of the *Chandra* ACIS-I CCDs and draw photons from each position and spectral

channel according to a Poisson distribution with the exposure time necessary to achieve a similar total number of counts as in the SPT and 400d samples. We project X-ray emission from hydrodynamical cells along three perpendicular lines of sight within three times the virial radius around the cluster center, which will allow us to determine how projection affects the determination of X-ray morphology. For $z = 0.33$, 0.66, 0.99, the spatial resolution of the simulations correspond to 4.0, 3.3, and 2.7 kpc, which is of higher resolution than the *Chandra* PSF at these redshifts. For $z = 0.02$, we additionally displaced each photon in the map by a Gaussian noise with width 10 kpc, to approximate the spatial resolution we achieve for high-z systems.

We generate 100 X-ray maps for each cluster, at each redshift, and projection, choosing the exposure times so that the distribution of the total number of photons within an annulus of 0.15–1 R_{500} mimics the corresponding distribution of observed counts in the SPT and 400d samples (see McDonald et al. 2014). This means that for each simulated cluster, we sample a range in redshift, projection, and signal-to-noise that is representative of the observed data. This effectively simulates all possible variations in observation quality that are present in the 400d and SPT samples and minimizes the effect of observational S/N on the distribution of the substructure statistics. For the purpose of substructure comparison, we treat each simulated X-ray map as an independent object, therefore having 85 (clusters) × 3 (projections) × 100 (iterations) = 2550 values of the substructure statistic for each of the four epochs.

3. Methods

3.1. Measuring $A_{\rm phot}$ and w for Data

In this analysis, we utilize two morphological parameters that trace the degree of cluster disturbance: photon asymmetry ($A_{\rm phot}$; Nurgaliev et al. 2013) and centroid shifts (w; Mohr et al. 1993). Briefly, $A_{\rm phot}$ quantifies the amount of asymmetry by comparing the cumulative photon count distribution as a function of azimuth for a given annulus to a uniform distribution, computing a probability that these two distributions are different for each annulus using the nonparametric Watson test (for a complete description of this test, see Feigelson & Babu 2012; Pewsey et al. 2015). On the other hand, w is a measure of how much the X-ray centroid moves over some radial range (see Mohr et al. 1993 for a more detailed description). $A_{\rm phot}$ and w show a significant degree of correlation with each other, and also with by-eye classification of cluster morphology (Nurgaliev et al. 2013). Both are sensitive to spatial irregularities in X-ray emission in the plane of the sky. By design, $A_{\rm phot}$ has more statistical power in resolving substructure and is able to produce more consistent results independent of the quality of observation (such as exposure, background level, etc.). On the other hand, w is better established in the literature as a widely used substructure statistic, so we include it in our tests for comparison to these other works.

In Tables 1 and 2, we provide asymmetry measurements derived from $A_{\rm phot}$ and w for each galaxy cluster in the 400d (Table 1) and SPT (Table 2) samples. These results will be used for the remainder of the study, and are provided here to aid in future studies that wish to isolate a relaxed/disturbed subsample of galaxy clusters over a large range in redshift.

Table 1
X-Ray Asymmetry Measurements for a Sample of 36 X-Ray-selected Clusters from the *ROSAT* PSPC 400 deg^2 Survey

Name	z	M_{500} [$10^{14}\,M_\odot$]	$A_{\rm phot}$	w
cl0302m0423	0.350	3.7	$0.06^{+0.02}_{-0.03}$	$0.005^{+0.001}_{-0.001}$
cl1212p2733	0.353	6.2	$0.05^{+0.03}_{-0.11}$	$0.018^{+0.003}_{-0.004}$
cl0350m3801	0.363	1.4	$0.14^{+0.07}_{-0.11}$	$0.026^{+0.010}_{-0.008}$
cl0318m0302	0.370	2.8	$0.41^{+0.10}_{-0.09}$	$0.025^{+0.004}_{-0.003}$
cl0159p0030	0.386	2.5	$0.01^{+0.03}_{-0.06}$	$0.010^{+0.003}_{-0.004}$
cl0958p4702	0.390	1.8	$0.05^{+0.05}_{-0.09}$	$0.015^{+0.003}_{-0.005}$
cl0809p2811	0.399	3.7	$0.63^{+0.20}_{-0.24}$	$0.059^{+0.008}_{-0.008}$
cl1416p4446	0.400	2.5	$0.11^{+0.02}_{-0.03}$	$0.014^{+0.002}_{-0.002}$
cl1312p3900	0.404	2.8	$0.22^{+0.15}_{-0.27}$	$0.041^{+0.010}_{-0.011}$
cl1003p3253	0.416	2.8	$0.23^{+0.06}_{-0.08}$	$0.028^{+0.003}_{-0.004}$
cl0141m3034	0.442	1.2	$0.29^{+0.18}_{-0.45}$	$0.091^{+0.018}_{-0.020}$
cl1701p6414	0.453	3.3	$0.14^{+0.02}_{-0.03}$	$0.022^{+0.003}_{-0.002}$
cl1641p4001	0.464	1.7	$0.01^{+0.04}_{-0.06}$	$0.015^{+0.004}_{-0.005}$
cl0522m3624	0.472	2.2	$0.06^{+0.05}_{-0.07}$	$0.013^{+0.003}_{-0.005}$
cl1222p2709	0.472	2.1	$0.04^{+0.03}_{-0.05}$	$0.011^{+0.003}_{-0.004}$
cl0355m3741	0.473	3.0	$0.05^{+0.03}_{-0.04}$	$0.016^{+0.004}_{-0.004}$
cl0853p5759	0.475	2.0	$0.32^{+0.10}_{-0.21}$	$0.051^{+0.007}_{-0.009}$
cl0333m2456	0.475	1.9	$0.25^{+0.22}_{-0.52}$	$0.023^{+0.006}_{-0.010}$
cl0926p1242	0.489	3.0	$0.07^{+0.02}_{-0.03}$	$0.010^{+0.003}_{-0.003}$
cl0030p2618	0.500	3.4	$0.07^{+0.08}_{-0.13}$	$0.013^{+0.006}_{-0.005}$
cl1002p6858	0.500	2.8	$0.07^{+0.08}_{-0.07}$	$0.021^{+0.004}_{-0.006}$
cl1524p0957	0.516	3.2	$0.67^{+0.14}_{-0.13}$	$0.055^{+0.007}_{-0.006}$
cl1357p6232	0.525	3.0	$0.20^{+0.06}_{-0.09}$	$0.022^{+0.003}_{-0.004}$
cl1354m0221	0.546	2.3	$0.26^{+0.10}_{-0.14}$	$0.033^{+0.007}_{-0.006}$
cl1120p2326	0.562	2.5	$0.46^{+0.18}_{-0.36}$	$0.041^{+0.007}_{-0.006}$
cl0956p4107	0.587	2.9	$0.83^{+0.18}_{-0.25}$	$0.042^{+0.008}_{-0.008}$
cl0328m2140	0.590	3.4	$0.08^{+0.03}_{-0.06}$	$0.016^{+0.003}_{-0.003}$
cl1120p4318	0.600	3.9	$0.23^{+0.14}_{-0.20}$	$0.024^{+0.005}_{-0.003}$
cl1334p5031	0.620	2.6	$0.09^{+0.18}_{-0.23}$	$0.016^{+0.005}_{-0.007}$
cl0542m4100	0.642	4.1	$0.53^{+0.17}_{-0.27}$	$0.042^{+0.004}_{-0.005}$
cl1202p5751	0.677	2.9	$0.25^{+0.07}_{-0.07}$	$0.053^{+0.008}_{-0.009}$
cl0405m4100	0.686	2.5	$0.03^{+0.03}_{-0.04}$	$0.008^{+0.002}_{-0.002}$
cl1221p4918	0.700	4.9	$0.14^{+0.08}_{-0.19}$	$0.019^{+0.004}_{-0.005}$
cl0230p1836	0.799	3.5	$1.29^{+0.26}_{-0.22}$	$0.098^{+0.007}_{-0.008}$
cl0152m1358	0.833	3.9	$2.27^{+0.22}_{-0.24}$	$0.155^{+0.014}_{-0.021}$
cl1226p3332	0.888	7.6	$0.10^{+0.02}_{-0.02}$	$0.014^{+0.001}_{-0.001}$

3.2. Calibrating $A_{\rm phot}$ with Simulated Clusters

To measure how well photon asymmetry, $A_{\rm phot}$, correlates with the dynamical state of a cluster, and to determine realistic cutoff values for "disturbed" and "relaxed" systems, we consider a set of simulated major mergers ($M_1/M_2 > 0.5$) of massive clusters from the *Omega500* simulations (Kravtsov 1999; Kravtsov et al. 2002; Rudd et al. 2008). For each of the 26 mergers, we produce X-ray photon maps along 3 different projections and at 16 different times, starting at ∼2 Gyr before the first core passage and ending ∼5 Gyr after. At each time step, we compute $A_{\rm phot}$ from the simulated observations in the same way that we do for the data. Figure 2 shows the results of this study. For all 78 mergers (26 clusters × 3 projections), the measured $A_{\rm phot}$ increases dramatically immediately following the merger, with the median cluster having $A_{\rm phot} \sim 0.9$ immediately after the merger and $A_{\rm phot} > 0.6$ for ∼1 Gyr after. We choose $A_{\rm phot} = 0.6$ as a reasonable threshold for a "disturbed" cluster,

Table 2
X-Ray Asymmetry Measurements for a Sample of 90 SZ-selected Clusters from the South Pole Telescope 2500 deg^2 Survey

Name	z	M_{500} [$10^{14}\,M_\odot$]	$A_{\rm phot}$	w
SPT-CLJ0000-5748	0.702	4.4	$0.07^{+0.04}_{-0.05}$	$0.009^{+0.002}_{-0.003}$
SPT-CLJ0013-4906	0.406	5.5	$0.66^{+0.06}_{-0.06}$	$0.029^{+0.002}_{-0.002}$
SPT-CLJ0014-4952	0.752	5.4	$1.44^{+0.22}_{-0.20}$	$0.101^{+0.008}_{-0.008}$
SPT-CLJ0033-6326	0.597	3.7	$0.05^{+0.04}_{-0.06}$	$0.024^{+0.006}_{-0.006}$
SPT-CLJ0037-5047	1.026	1.4	$0.12^{+0.23}_{-0.46}$	$0.066^{+0.014}_{-0.025}$
SPT-CLJ0040-4407	0.350	5.7	$0.06^{+0.03}_{-0.05}$	$0.009^{+0.002}_{-0.003}$
SPT-CLJ0058-6145	0.864	3.1	$0.05^{+0.04}_{-0.07}$	$0.025^{+0.005}_{-0.006}$
SPT-CLJ0102-4603	0.772	2.8	$0.68^{+0.18}_{-0.17}$	$0.074^{+0.010}_{-0.009}$
SPT-CLJ0102-4915	0.870	16.8	$2.46^{+0.02}_{-0.02}$	$0.094^{+0.000}_{-0.000}$
SPT-CLJ0106-5943	0.348	4.2	$0.14^{+0.03}_{-0.04}$	$0.017^{+0.003}_{-0.003}$
SPT-CLJ0123-4821	0.620	3.3	$0.28^{+0.09}_{-0.12}$	$0.039^{+0.005}_{-0.007}$
SPT-CLJ0142-5032	0.730	7.2	$0.33^{+0.11}_{-0.21}$	$0.028^{+0.005}_{-0.006}$
SPT-CLJ0151-5954	1.049	2.9	$1.26^{+0.60}_{-0.33}$	$0.055^{+0.012}_{-0.028}$
SPT-CLJ0200-4852	0.499	5.5	$0.43^{+0.13}_{-0.16}$	$0.033^{+0.006}_{-0.004}$
SPT-CLJ0212-4657	0.640	3.6	$1.11^{+0.15}_{-0.16}$	$0.075^{+0.006}_{-0.005}$
SPT-CLJ0217-5245	0.343	4.2	$1.00^{+0.16}_{-0.17}$	$0.093^{+0.007}_{-0.006}$
SPT-CLJ0232-4421	0.284	10.8	$0.33^{+0.02}_{-0.02}$	$0.017^{+0.001}_{-0.001}$
SPT-CLJ0232-5257	0.556	5.9	$0.30^{+0.06}_{-0.12}$	$0.043^{+0.006}_{-0.006}$
SPT-CLJ0234-5831	0.415	7.0	$0.15^{+0.04}_{-0.06}$	$0.007^{+0.002}_{-0.002}$
SPT-CLJ0235-5121	0.278	8.0	$0.22^{+0.04}_{-0.05}$	$0.031^{+0.003}_{-0.003}$
SPT-CLJ0236-4938	0.334	4.0	$0.19^{+0.06}_{-0.07}$	$0.035^{+0.006}_{-0.004}$
SPT-CLJ0243-5930	0.635	4.8	$0.17^{+0.04}_{-0.04}$	$0.016^{+0.003}_{-0.003}$
SPT-CLJ0252-4824	0.330	4.1	$0.97^{+0.34}_{-0.28}$	$0.034^{+0.008}_{-0.014}$
SPT-CLJ0256-5617	0.580	4.3	$1.53^{+0.72}_{-0.35}$	$0.068^{+0.013}_{-0.018}$
SPT-CLJ0304-4401	0.460	10.4	$0.50^{+0.07}_{-0.08}$	$0.044^{+0.006}_{-0.003}$
SPT-CLJ0304-4921	0.392	6.2	$0.11^{+0.02}_{-0.03}$	$0.010^{+0.002}_{-0.002}$
SPT-CLJ0307-5042	0.550	4.4	$0.08^{+0.03}_{-0.04}$	$0.015^{+0.003}_{-0.003}$
SPT-CLJ0307-6225	0.581	5.7	$2.46^{+0.22}_{-0.28}$	$0.136^{+0.034}_{-0.011}$
SPT-CLJ0310-4647	0.850	10.6	$0.07^{+0.06}_{-0.08}$	$0.014^{+0.003}_{-0.004}$
SPT-CLJ0324-6236	0.730	4.1	$0.02^{+0.03}_{-0.03}$	$0.013^{+0.003}_{-0.003}$
SPT-CLJ0330-5228	0.442	4.4	$0.08^{+0.03}_{-0.07}$	$0.040^{+0.005}_{-0.007}$
SPT-CLJ0334-4659	0.450	5.5	$0.16^{+0.04}_{-0.03}$	$0.018^{+0.003}_{-0.002}$
SPT-CLJ0346-5439	0.530	4.1	$0.12^{+0.02}_{-0.04}$	$0.023^{+0.004}_{-0.004}$
SPT-CLJ0348-4515	0.358	3.9	$0.07^{+0.08}_{-0.13}$	$0.016^{+0.004}_{-0.005}$
SPT-CLJ0352-5647	0.670	3.7	$0.08^{+0.05}_{-0.07}$	$0.017^{+0.003}_{-0.005}$
SPT-CLJ0406-4805	0.590	4.0	$0.37^{+0.14}_{-0.19}$	$0.054^{+0.010}_{-0.008}$
SPT-CLJ0411-4819	0.422	6.0	$0.82^{+0.04}_{-0.04}$	$0.057^{+0.001}_{-0.002}$
SPT-CLJ0417-4748	0.590	4.9	$0.05^{+0.02}_{-0.03}$	$0.010^{+0.002}_{-0.002}$
SPT-CLJ0426-5455	0.630	3.8	$0.22^{+0.27}_{-0.34}$	$0.046^{+0.010}_{-0.010}$
SPT-CLJ0438-5419	0.421	10.3	$0.25^{+0.04}_{-0.03}$	$0.018^{+0.001}_{-0.001}$
SPT-CLJ0441-4855	0.790	4.0	$0.06^{+0.03}_{-0.02}$	$0.013^{+0.003}_{-0.002}$
SPT-CLJ0446-5849	1.182	3.6	$0.21^{+0.36}_{-0.90}$	$0.059^{+0.011}_{-0.012}$
SPT-CLJ0449-4901	0.792	5.1	$0.27^{+0.07}_{-0.10}$	$0.046^{+0.007}_{-0.006}$
SPT-CLJ0456-5116	0.570	4.4	$0.11^{+0.05}_{-0.08}$	$0.023^{+0.005}_{-0.004}$
SPT-CLJ0509-5342	0.461	5.7	$0.07^{+0.02}_{-0.02}$	$0.008^{+0.002}_{-0.002}$
SPT-CLJ0516-5430	0.295	12.4	$0.18^{+0.07}_{-0.16}$	$0.055^{+0.003}_{-0.004}$
SPT-CLJ0522-4818	0.296	3.7	$0.22^{+0.08}_{-0.11}$	$0.016^{+0.003}_{-0.004}$
SPT-CLJ0528-5300	0.768	2.4	$0.07^{+0.04}_{-0.08}$	$0.024^{+0.005}_{-0.007}$
SPT-CLJ0533-5005	0.881	1.6	$0.21^{+0.16}_{-0.24}$	$0.049^{+0.011}_{-0.009}$
SPT-CLJ0542-4100	0.642	5.7	$0.55^{+0.21}_{-0.20}$	$0.041^{+0.004}_{-0.005}$
SPT-CLJ0546-5345	1.066	5.0	$0.10^{+0.03}_{-0.04}$	$0.022^{+0.003}_{-0.004}$
SPT-CLJ0551-5709	0.423	3.0	$0.88^{+0.12}_{-0.16}$	$0.086^{+0.007}_{-0.006}$
SPT-CLJ0555-6406	0.270	5.6	$0.36^{+0.09}_{-0.13}$	$0.039^{+0.004}_{-0.003}$
SPT-CLJ0559-5249	0.609	5.5	$0.39^{+0.10}_{-0.08}$	$0.041^{+0.006}_{-0.007}$
SPT-CLJ0616-5227	0.684	4.5	$0.34^{+0.10}_{-0.06}$	$0.031^{+0.005}_{-0.004}$
SPT-CLJ0655-5234	0.500	4.8	$0.30^{+0.12}_{-0.21}$	$0.018^{+0.005}_{-0.007}$
SPT-CLJ0658-5556	0.296	18.2	$2.41^{+0.01}_{-0.01}$	$0.020^{+0.000}_{-0.000}$
SPT-CLJ2011-5725	0.279	2.6	$0.11^{+0.02}_{-0.02}$	$0.006^{+0.001}_{-0.001}$
SPT-CLJ2031-4037	0.330	7.4	$0.25^{+0.04}_{-0.04}$	$0.017^{+0.001}_{-0.002}$
SPT-CLJ2034-5936	0.946	4.4	$0.05^{+0.05}_{-0.23}$	$0.014^{+0.003}_{-0.005}$
SPT-CLJ2035-5251	0.520	4.1	$0.07^{+0.13}_{-0.35}$	$0.039^{+0.011}_{-0.010}$
SPT-CLJ2043-5035	0.723	5.1	$0.11^{+0.02}_{-0.02}$	$0.009^{+0.002}_{-0.002}$
SPT-CLJ2106-5844	1.132	8.1	$0.20^{+0.04}_{-0.03}$	$0.022^{+0.002}_{-0.002}$
SPT-CLJ2135-5726	0.427	4.5	$0.10^{+0.04}_{-0.08}$	$0.013^{+0.003}_{-0.004}$
SPT-CLJ2145-5644	0.480	5.7	$0.06^{+0.03}_{-0.04}$	$0.011^{+0.003}_{-0.003}$
SPT-CLJ2146-4633	0.933	3.5	$0.05^{+0.05}_{-0.18}$	$0.026^{+0.007}_{-0.009}$
SPT-CLJ2148-6116	0.571	5.0	$0.06^{+0.03}_{-0.07}$	$0.018^{+0.004}_{-0.004}$
SPT-CLJ2218-4519	0.650	4.9	$0.08^{+0.05}_{-0.08}$	$0.027^{+0.005}_{-0.006}$
SPT-CLJ2222-4834	0.652	4.0	$0.09^{+0.03}_{-0.05}$	$0.015^{+0.003}_{-0.004}$
SPT-CLJ2232-5959	0.594	3.3	$0.10^{+0.03}_{-0.03}$	$0.008^{+0.002}_{-0.002}$
SPT-CLJ2233-5339	0.480	6.4	$0.29^{+0.12}_{-0.11}$	$0.026^{+0.005}_{-0.003}$
SPT-CLJ2236-4555	1.161	2.7	$0.90^{+0.18}_{-0.19}$	$0.063^{+0.007}_{-0.009}$
SPT-CLJ2245-6206	0.580	5.1	$0.87^{+0.19}_{-0.19}$	$0.113^{+0.007}_{-0.010}$
SPT-CLJ2248-4431	0.351	16.7	$0.21^{+0.03}_{-0.02}$	$0.006^{+0.001}_{-0.000}$
SPT-CLJ2258-4044	0.864	4.0	$0.31^{+0.15}_{-0.24}$	$0.036^{+0.007}_{-0.004}$
SPT-CLJ2259-6057	0.750	4.3	$0.11^{+0.03}_{-0.03}$	$0.020^{+0.002}_{-0.003}$
SPT-CLJ2301-4023	0.730	2.4	$0.61^{+0.14}_{-0.12}$	$0.028^{+0.003}_{-0.003}$
SPT-CLJ2306-6505	0.530	5.6	$0.05^{+0.04}_{-0.12}$	$0.029^{+0.006}_{-0.008}$
SPT-CLJ2325-4111	0.358	8.2	$0.36^{+0.09}_{-0.10}$	$0.040^{+0.004}_{-0.005}$
SPT-CLJ2331-5051	0.576	4.6	$0.14^{+0.03}_{-0.05}$	$0.034^{+0.002}_{-0.003}$
SPT-CLJ2332-5053	0.560	5.3	$4.10^{+0.48}_{-0.43}$	$0.090^{+0.005}_{-0.007}$
SPT-CLJ2335-4544	0.570	5.4	$0.04^{+0.03}_{-0.07}$	$0.015^{+0.004}_{-0.004}$
SPT-CLJ2337-5942	0.775	5.9	$0.09^{+0.03}_{-0.05}$	$0.016^{+0.002}_{-0.003}$
SPT-CLJ2341-5119	1.003	5.8	$0.07^{+0.02}_{-0.03}$	$0.017^{+0.002}_{-0.003}$
SPT-CLJ2342-5411	1.075	1.9	$0.08^{+0.03}_{-0.04}$	$0.011^{+0.002}_{-0.004}$
SPT-CLJ2344-4243	0.596	11.9	$0.03^{+0.00}_{-0.00}$	$0.002^{+0.000}_{-0.000}$
SPT-CLJ2345-6405	0.962	4.9	$0.33^{+0.19}_{-0.53}$	$0.045^{+0.008}_{-0.007}$
SPT-CLJ2352-4657	0.783	4.0	$0.02^{+0.03}_{-0.04}$	$0.020^{+0.004}_{-0.005}$
SPT-CLJ2355-5055	0.320	3.9	$0.07^{+0.02}_{-0.02}$	$0.013^{+0.003}_{-0.003}$
SPT-CLJ2359-5009	0.775	2.9	$0.24^{+0.08}_{-0.14}$	$0.016^{+0.005}_{-0.006}$

due to the fact that ~70% (54/78) of our simulated systems (combining all three projections) cross this threshold immediately following a major merger—we do not expect to identify 100% of merging systems with a 2D metric such as $A_{\rm phot}$ due to the fact that a significant fraction of systems will be merging along the line of sight. We note that the median value of $A_{\rm phot}$ immediately after a line-of-sight merger is ~0.6.

Before and after the merger, the median value of $A_{\rm phot}$ fluctuates about ~0.2, indicating that this may provide a reasonable threshold for "relaxed," depending on how strict one wants to be with that identifier (see Mantz et al. 2015, for a more strict classification). We note that during this pre- and post-merger phase, the measured $A_{\rm phot}$ is inconsistent with >0.6 for $\gtrsim$85% of systems, meaning that we expect little contamination in the "disturbed" sample from systems that are not currently undergoing major mergers.

Based on these arguments, we infer that an elevated $A_{\rm phot}$ (>0.6) is an adequate proxy of a major merger within the past ~1–2 Gyr.

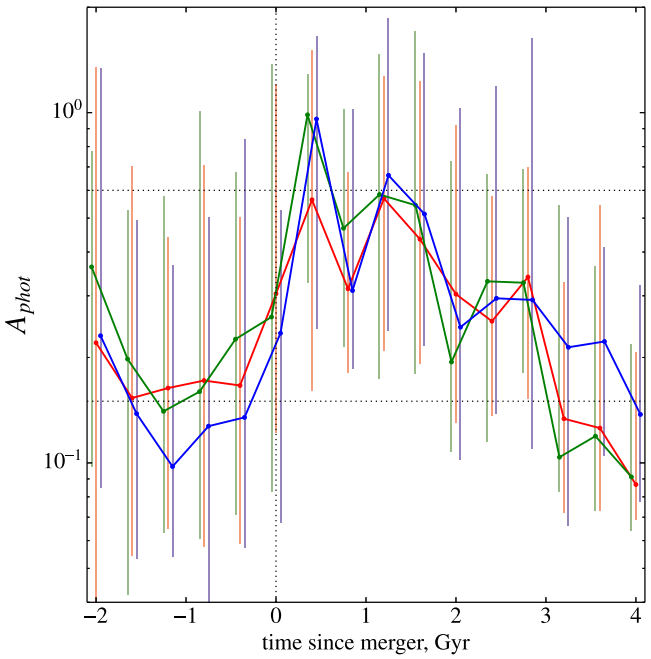

Figure 2. Evolution of $A_{\rm phot}$ for 26 simulated clusters over the course of a major ($M_1/M_2 > 0.5$) merger. The solid colored lines shows the median value of $A_{\rm phot}$ as a function of time since the merger for all 26 clusters along three different sight lines (blue and green are in the plane of the sky, while red is along the line of sight), while the vertical bars show the 1σ scatter. Horizontal lines at 0.15 and 0.6 show our divisions for relaxed and disturbed clusters, respectively. The vertical dotted line corresponds to the start of the merger. We find that, within 1 Gyr of the merger, ∼70% of the simulated images (over all three projections) have $A_{\rm phot} > 0.6$, while <15% of the pre- or post-merger systems have such high values, making $A_{\rm phot} > 0.6$ a reasonable proxy for ongoing mergers.

3.3. Statistical Comparisons

To judge the similarity of cluster subsamples with different selections or redshifts, we use both the two-sample Kolmogorov–Smirnov (KS) and the Anderson–Darling (AD) tests for the empirical distributions of measured substructure parameters. The values for these statistics are converted to the p-value of the null hypothesis (that these two empirical distributions come from the same underlying distribution). Common practice is to reject the null hypothesis for $p < 0.05/N$, where N is the number of tests being conducted (the so-called "Bonferroni correction"). We have tested that using more advanced methods for multiple hypotheses p-value adjustments, such as the Benjamini–Hochberg False Discovery Rate (Benjamini & Hochberg 1995), does not change any of the conclusions of this paper.

We consider the following pairs of subsamples in our comparisons:

(A) *400d low-z versus SPT low-z* ($z < 0.6$; 27 and 50 clusters respectively). Within this redshift range, we find no correlation between mass and $A_{\rm phot}$ (Pearson $r = 0.11$) for simulated clusters, suggesting that poor overlap in mass between the two samples will not drive any result. Therefore, despite the minimal overlap in mass space, these samples allow for the cleanest test of morphology differences between X-ray- and SZ-selected cluster samples.

(B) *400d versus SPT, all redshifts* (36 and 90 clusters). Allows a comparison of cluster morphologies in the complete SZ and X-ray selected catalogs, ignoring that these samples have different redshift and mass ranges.

(C) *SPT low-z versus SPT high-z* ($z < 0.6$, $z > 0.6$; 50 and 40 clusters, respectively). This is the cleanest test for substructure evolution, since the high-z SPT-selected clusters can be considered the progenitors of the low-z SPT-selected systems (see Figure 7 of McDonald et al. 2014). Given the relatively small change in angular size between the low-z and high-z systems, we expect there to be minimal redshift-dependent selection biases between these two subsamples.

(D) *400d low-z versus SPT high-z* ($z < 0.6$, $z > 0.6$; 27 and 40 clusters, respectively). This is a complementary test for substructure evolution, which has the underlying evolution convolved with redshift-dependent, mass-dependent, and method-dependent selection bias. Naively, one might assume that both SZ-selected clusters and high-redshift clusters are more likely to be mergers, because mergers can cause temporarily increased pressure, and the merger rate is higher at early times (e.g., Fakhouri & Ma 2010). This would lead to the high-z, SPT-selected clusters being significantly more disturbed than the low-z, X-ray-selected clusters, if these statements are true. A lack of difference in morphology between high-z SZ-selected and low-z X-ray selected clusters would indicate that the combination of these effects is insignificant.

(E) *400d+SPT low-z versus 400d+SPT high-z* ($z < 0.6$, $z > 0.6$; 75 and 49 clusters, respectively). If selection criteria are indeed not important, this test increases the statistical power of the substructure evolution test due to an increased number of clusters in the combined samples.

(F, G, H) *SPT+400d versus simulations*. For these comparisons, we select real and simulated clusters within the redshift intervals of $z = 0.33 \pm 0.1$, $z = 0.66 \pm 0.1$, and $z = 0.99 \pm 0.15$.

4. Results

4.1. X-Ray–SZ Comparisons

Figure 3(A) shows histograms and cumulative distributions of $A_{\rm phot}$ and w for 400d and SPT subsamples with $z < 0.6$. In general, the distribution of substructure statistics is closer to log-normal than normal. High p-values ($p > 0.1$) of the KS and AD statistics for $A_{\rm phot}$ and w indicate that the two samples (400d and SPT; $z < 0.6$) are indistinguishable in terms of their X-ray morphology. Under the assumption that mergers are characterized by increased values of $A_{\rm phot}$ and w, this means that the fraction of merging systems detected by SPT is similar to the amount of merging systems detected by their X-ray emission. Assuming that ongoing mergers can be identified as having $A_{\rm phot} > 0.6$, following Figure 2 (see also Figure 7 from Nurgaliev et al. 2013), we find a low-z merging fraction of $20^{+7}_{-4}\%$ and $11^{+9}_{-4}\%$ in the SPT and 400d samples, respectively, where the uncertainty range is the 1σ binomial population confidence interval (Cameron 2011). For comparison, Mann & Ebeling (2012) find a merger fraction of $24^{+5}_{-4}\%$ for a sample of 79 X-ray selected clusters spanning the same redshift range ($0.25 < z < 0.6$).

We next compare clusters at low and high redshift, subject to the same selection criteria. We consider $z < 0.6$ and $z > 0.6$ subsamples of the parent SPT sample only, because the 400d sample only has nine clusters at $z > 0.6$. Figure 3(B) shows

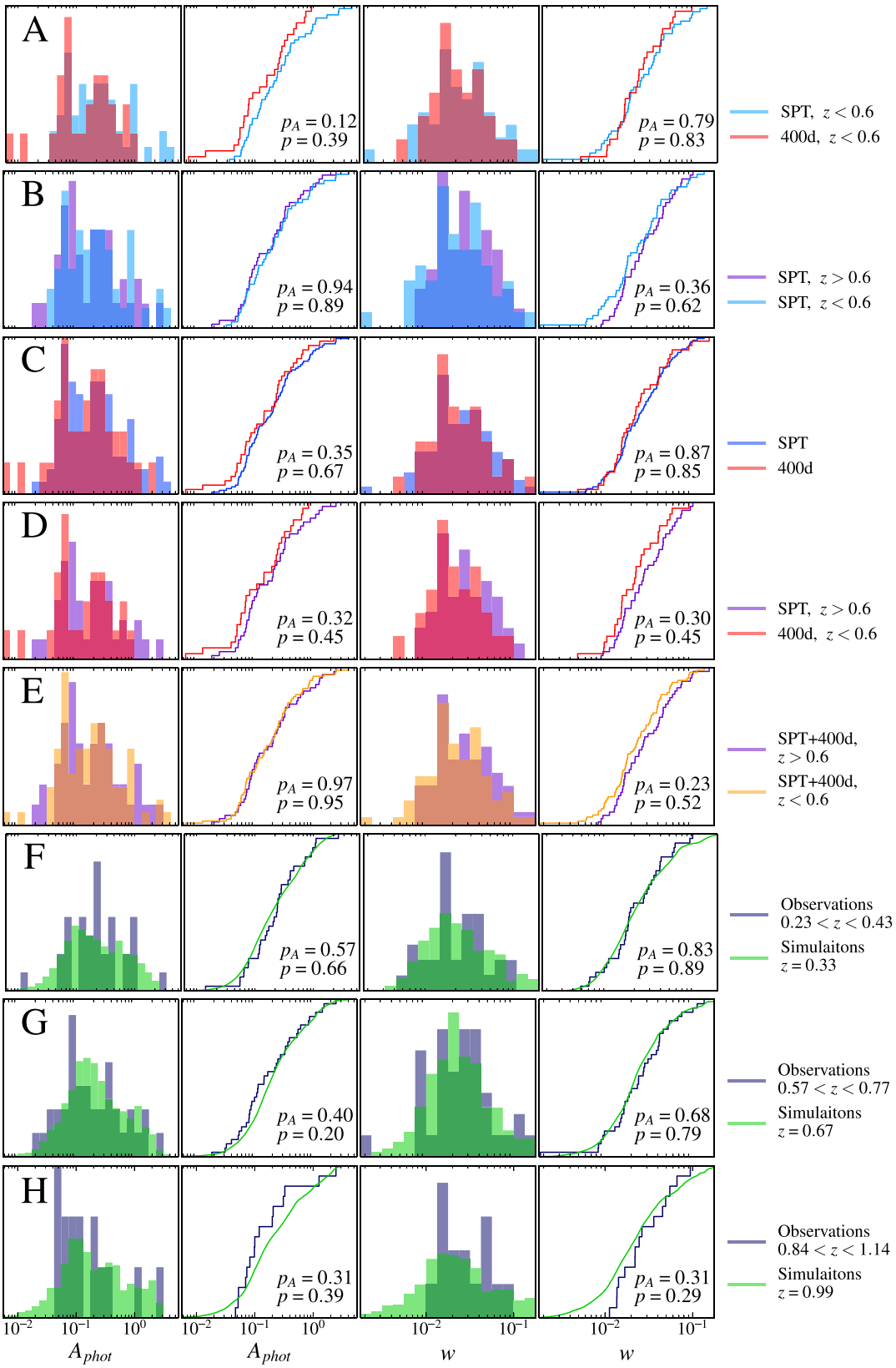

Figure 3. Distribution of morphological statistics for various partitions of clusters in the 400d and SPT catalogs. In each row, the left two plots show the histograms and cumulative distribution functions for $A_{\rm phot}$, and the right two plots show the same for w. The p-value, based on the KS and AD tests, of the hypothesis that the substructure statistic values come from the same distribution are shown in each plot. High values of p indicate that the distributions are similar, low values of p indicate that the distributions are different. It is a common practice to exclude the null hypothesis at $p < 0.05/$(number of tests being conducted). The rightmost column explains what samples are being compared. See Section 3 for the motivation of the subsamples to be compared and Section 4 for the description of the results.

that the distributions of both $A_{\rm phot}$ and w for low-z and high-z systems are indistinguishable ($p > 0.3$). Using the same criterion as above, we find that $20^{+7}_{-4}\%$ and $18^{+8}_{-4}\%$ of clusters at low-z and high-z, respectively, are identified as ongoing mergers.

Given that the SZ-selected clusters demonstrate weak, if any, redshift evolution in their morphology, we next consider the full 400d and SPT catalogs, without any redshift restrictions. Figure 3(C) shows distribution functions for these two samples, which are still indistinguishable ($p > 0.3$) for both $A_{\rm phot}$ and w. The fraction of systems with $A_{\rm phot} > 0.6$ over the full redshift range is $19^{+5}_{-3}\%$ and $14^{+8}_{-4}\%$ for the SPT and 400d samples, respectively.

If there was only a weak dependence of X-ray morphology on both selection and redshift, we might not have sufficient statistical power to detect such dependencies with the tests described above. In an effort to maximize the effects of these two potential biases, we compare the low-z X-ray sample to the high-z SZ sample. Figure 3(D) shows the result of this comparison, demonstrating that there is no statistically significant difference in the distribution of X-ray morphologies between these two extreme subsamples ($p > 0.3$). This means that, if clusters are, on average, more morphologically disturbed in SZ-selected clusters and at high redshift, the combined effects on the measured disturbed fraction are small ($\lesssim$10%).

As a final test, we make the assumption that there is no morphological bias due to selection, and combine the 400d and SPT samples to maximize our ability to detect redshift dependence. Figure 3(E) demonstrates that there is no measurable redshift dependence even when the two samples are combined, for both $A_{\rm phot}$ and w ($p > 0.2$). In this combined sample, we find that the fraction of clusters with $A_{\rm phot} > 0.6$ is $17^{+5}_{-3}\%$ and $19^{+7}_{-4}\%$ for the low-z and high-z subsets, respectively.

In summary, we find no statistically significant dependence on either the distribution of X-ray morphologies or the fraction of clusters classified as morphologically disturbed with redshift or selection methods.

4.2. Data–Simulation Comparisons

Panels (F), (G), and (H) of Figure 3 show comparisons between observed and simulated clusters at three redshifts: $z = 0.33$, 0.66, and 0.99. We note a remarkable agreement in the X-ray morphology between simulations and observations. The lowest measured p-value is 0.20 (for KS test on $A_{\rm phot}$ at $z = 0.67$). Given that we have made 16 individual comparisons, we require $p < 0.003$ to reject the null hypothesis that these two distributions come from the same parent distribution. From the similarity in the data and simulations, we can arrive at two conclusions. First, the lack of evolution in X-ray morphology is observed in both simulations and observations, suggesting that this is not due to a selection bias. Second, the observed morphology is relatively insensitive to complex physics (e.g., cooling, AGN feedback, etc.), and appears to be primarily driven by gravitational processes (i.e., mergers), which simulations adequately describe.

In Figure 4, we show the X-ray morphology, as quantified by $A_{\rm phot}$ and w, as a function of redshift for the 400d and SPT samples, as well as for the *Omega500* simulations. For w, we compare to the low-z REXCESS cluster sample (Böhringer et al. 2010). Our definition of w (Nurgaliev et al. 2013) is slightly different than that used by Böhringer et al. (2010)—to make our results comparable, we compute w for all clusters in the 400d and SPT samples using both methods and found that scaling by a factor of 1.5 brings the two into excellent agreement. After the applied correction, the range and median values of w in the study of Böhringer et al. (2010) are similar to those for low- and high-z clusters in the 400d and SPT samples.

Figure 4 demonstrates that the amount of substructure, as quantified by w and $A_{\rm phot}$, is remarkably similar for four different samples: 400d, SPT, REXCESS, and the *Omega500* simulations. Furthermore, there is no statistically significant redshift evolution over the combined samples. We note that the mild (2σ) redshift evolution measured in w is consistent with our earlier findings that w is biased high for low-quality X-ray data (Nurgaliev et al. 2013). Given that the most distant clusters have higher background and fewer net counts than their low-z counterparts, we expect them to be biased high in w.

In Figure 5, we show the same measurements of $A_{\rm phot}$ and w as in Figure 4, but now as a function of cluster mass. We find no statistically significant dependence of cluster morphology on cluster mass, over the mass range of 10^{14}–$10^{15}\,M_{\odot}$, for both simulated and real clusters. Importantly, the fact that there is no mass trend in the simulated clusters means that we are justified in comparing the simulated and real clusters despite the fact that their mass distributions are not identical.

5. Discussion

5.1. X-Ray, SZ Selection Biases

We have demonstrated in Figures 3–5 that there is no measurable difference between the distribution of X-ray morphology in the 400d X-ray-selected sample and the SPT SZ-selected sample. Furthermore, both of these selections appear to be unbiased with respect to simulations, suggesting that we are probing the full population of massive galaxy clusters. These results are consistent with work by Mantz et al. (2015) who found no difference in the relaxed fraction between the 400d and SPT samples, and with Sifón et al. (2013, 2016) who used dynamical tracers of substructure to show that the dynamical state of SZ-selected clusters from the Atacama Cosmology Telescope (ACT) were consistent with those of simulated massive clusters.

There persists a misconception that SZ-selected clusters are, on average, more often mergers than X-ray-selected clusters. We propose that this perceived bias stems from two facts pertaining to *Planck*-selected clusters that were summarized briefly in the introduction. First, *Planck* has a significantly more extended PSF ($\sim 7'$) compared to ground-based arcminute-resolution SZ experiments such as SPT and the ACT (Swetz et al. 2011). This means that, at high-z, *Planck* does not resolve close pairs (or triplets) of galaxy clusters (e.g., PLCKG214.6+37.0), thereby capturing an "inflated" SZ signal, where an instrument with a smaller beam such as SPT or ACT would see multiple independent systems with lower individual significance. Because the SPT beam ($\sim 1'$) is matched to the angular size of rich clusters at $z > 0.3$, it is unlikely that the SPT sample contains similar blended systems. Second, several major mergers in the *Planck* catalog received a great deal of initial attention, due to the fact that there were very few such systems previously known (i.e., triple clusters). However, the detection of a few dozen previously unknown mergers compared to more than 800 confirmed objects in

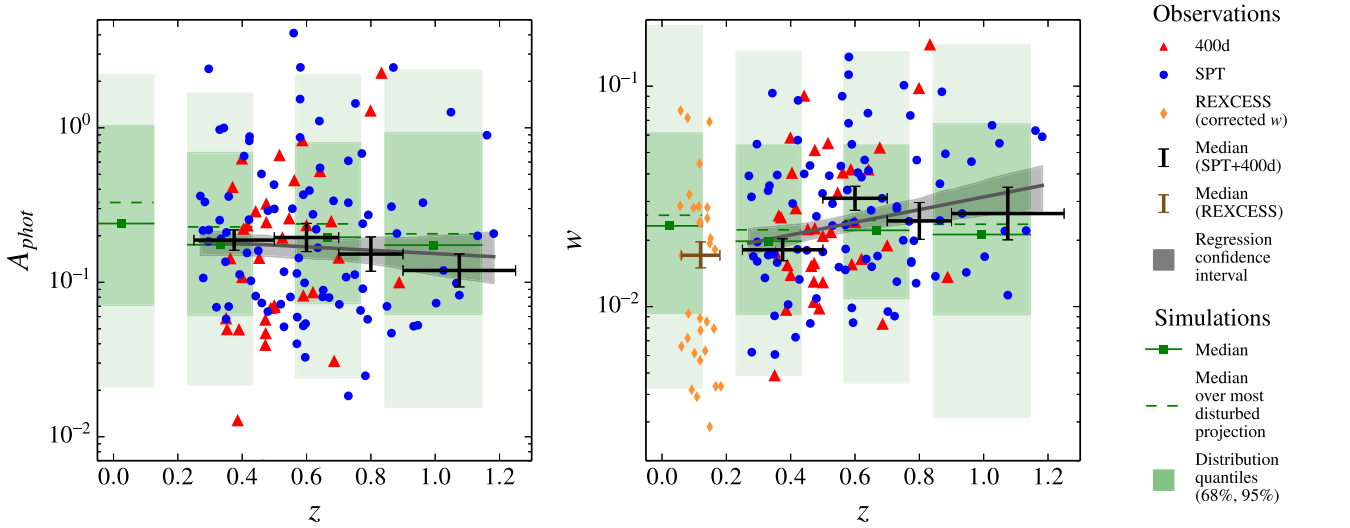

Figure 4. Redshift evolution of $A_{\rm phot}$ and w. Red triangles and blue circles represent 400d and SPT clusters respectively. Black error bars show the median values in four redshift bins ([0.25–0.5], [0.5–0.7], [0.7–0.9], [0.9–1.2]). We measure Pearson R coefficients for the left and right panels of 0.02 and 0.19, respectively, indicating a lack of a statistically significant correlation. The error bars on the median value are obtained by the Median Absolute Devitation method and demonstrate that there is no significant difference in the median value between bins. The slopes, derived by a simple linear regressions of $A_{\rm phot}$ and w with z, are consistent with no redshift evolution at $\sim 0.5\sigma$ and $\sim 2\sigma$ respectively. The shaded green regions show the 68% and 95% ranges for $A_{\rm phot}$ and w in simulated clusters, with the medians shown as green squares. The orange circles in the right panel show the values of w for the REXCESS sample presented in Böhringer et al. (2010), accordingly corrected for the different w definition used in this study.

Planck's catalog is not a statistically significant indication of morphological bias in SZ-selected cluster samples.

The results presented here are also consistent with recent work by Lin et al. (2015), who showed that the bias due to the presence or lack of a cool core is small (<1%) for the majority of systems, with exceptionally rare systems like the Phoenix cluster (McDonald et al. 2012) having biases as high as 10%. Applying the results of Lin et al. (2015) to the HIFLUGCS sample of low-z clusters (Reiprich & Böhringer 2002; Vikhlinin et al. 2007), we estimate that 1 in ~100 clusters has an SZ bias as high as ~10%. This bias is strongly redshift dependent, due to the combined effects of cool cores filling a smaller fraction of the beam at high-z, and being in general less cuspy at early times (McDonald et al. 2013; Mantz et al. 2015). As such, we expect the fraction of cool cores in the SPT-selected cluster sample to be nearly representative of the true underlying population. Convolving this very weak bias with the noisy correlation between X-ray morphology and the presence or lack of a cool core, we do not expect a cool core bias to drive a statistically significant difference in morphology between our X-ray- and SZ-selected samples.

We conclude that the common misconception that SZ-selected samples contain a non-representative fraction of mergers stems from some combination of these two points. We find no morphological bias in an SPT-selected sample, and would not expect any similarly selected samples to be biased either (see also, Motl et al. 2005; Lin et al. 2015).

5.2. Evolution of Substructure with Redshift

In Section 4.1, we demonstrated that low- and high-redshift systems show the same amount of substructure. This is somewhat surprising—in the standard growth of structure scenario, the fraction of disturbed clusters increases with increasing redshift. Additionally, these findings are in contradiction with some earlier works, specifically Jeltema et al. (2005) and Andersson et al. (2009), which both rely on power ratios to quantify morphology. Both Nurgaliev et al. (2013) and Weißmann et al. (2013) have shown that shot noise can strongly influence the measured power ratios. This may explain the results found by Andersson et al. (2009), who did not apply any shot noise correction.

One possible explanation for the lack of observed X-ray morphology evolution may be that there is not a one-to-one correspondence between substructure statistics such as w or P_3/P_0 and the dynamical state of the cluster. Indeed, simulations have shown that substructures statistics may vary significantly on short timescales during cluster mergers (O'Hara et al. 2006; Poole et al. 2006; Hallman & Jeltema 2011). For example, simulations by O'Hara et al. (2006) and Poole et al. (2006) show that w can easily vary in the range of $0.01 < w < 0.1$ over a fraction of characteristic merger time. In addition w shows similarly big fluctuations as a function of line of sight.

Another possibility (as pointed out by Weißmann et al. 2013) is that earlier studies that used power ratios might not account correctly for the insufficient photon statistics of high-z clusters. For example, Jeltema et al. (2005) analyzed redshift evolution using a relatively small sample of 40 clusters divided into low andhigh-redshift subsamples. The low-redshift sample contained 26 clusters with $z < 0.5$ and $\langle z \rangle = 0.24$. The high-redshift sample contained 14 clusters with $z > 0.5$ and $\langle z \rangle = 0.71$. They used the power ratio method and found the amount of substructure to be significantly different between the subsamples as measured by P_3/P_0. They also fit the P_3/P_0–z relation and found the slope to be positive with high ($p \approx 0.005$) significance. Their results are surprising in light of new studies of the properties of the power-ratios method. Both Nurgaliev et al. (2013) and Weißmann et al. (2013) find that P_3/P_0 are fully consistent with zero for a majority of high-redshift clusters due to the insufficient quality of observations.

Weißmann et al. (2013) performed a similar study using the same high-z 400d sample and a subsample of the SPT sample used in Andersson et al. (2011), combined with a low-z sample published in Weißmann et al. (2013). They quantified

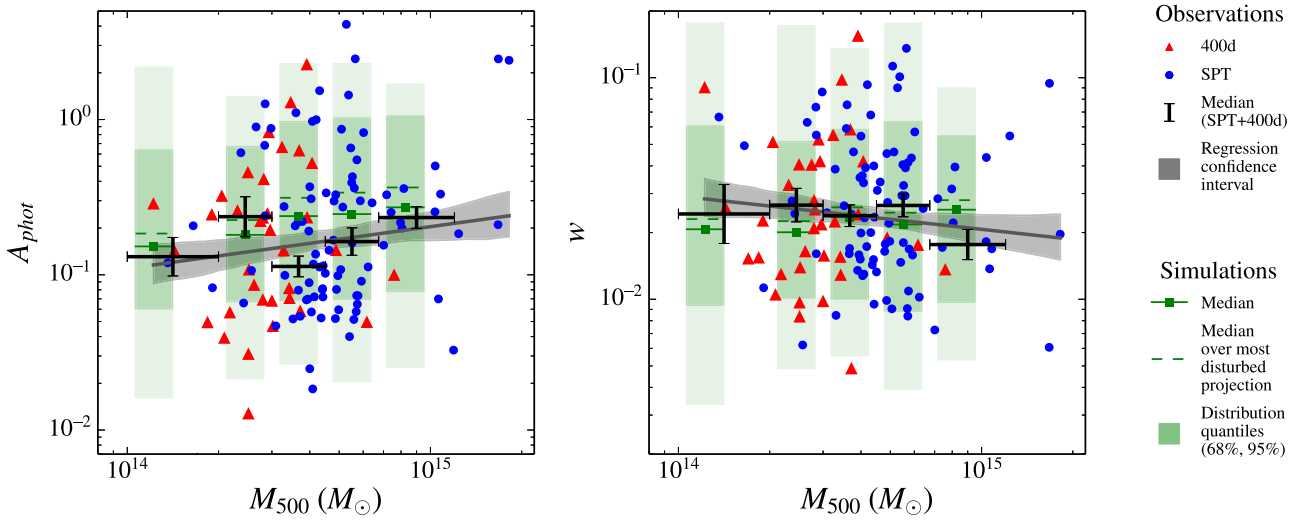

Figure 5. Similar to Figure 4, but now considering cluster mass rather than redshift. We measure Pearson R coefficients for the left and right panels of 0.19 and −0.11, respectively, indicating a lack of a statistically significant correlation. There is no significant correlation between cluster asymmetry and mass, regardless of the asymmetry estimator used or whether we consider real or simulated clusters.

morphology via centroid shifts and power ratios, finding a result consistent with no redshift evolution. Weißmann et al. (2013) emphasized bringing both high-redshift and low-redshift subsamples to the same quality of observations. To achieve that, they artificially degrade higher-quality observations of the low-redshift sample. This is not necessary in our analysis because (1) the observations of both samples were targeted for 2000 counts per cluster and (2) as shown in Nurgaliev et al. (2013) w is not sensitive to number of counts above ∼1000 counts (this is also confirmed in Weißmann et al. 2013) and $A_{\rm phot}$ has even better stability than w with respect to the number of X-ray counts. Indeed, the $A_{\rm phot}$ quantity was derived explicitly to avoid any bias due to data quality. The work presented here extends on that of Weißmann et al. (2013) by including a larger number of distant, SPT-selected clusters, by splitting high-redshift clusters into multiple selection (X-ray, SZ, simulation) bins, and by using a new substructure statistic, $A_{\rm phot}$.

A similar study was also conducted by Mantz et al. (2015), utilizing data from both the SPT and 400 deg^2 surveys along with data for RASS and *Planck* clusters. A direct comparison to the results of this study is not straightforward, because it was focused on the "relaxed fraction" using a conservative estimator for relaxedness, while this work focuses on the evolution of the full distribution of morphologies. Nonetheless, Mantz et al. (2015) find that the relaxed fraction does not evolve significantly between $z \sim 0$ and $z \sim 1$, and that there is no statistically significant difference between the morphologies of SPT-selected and 400d-selected clusters, consistent with this work.

Although we do not find evidence for a change in the amount of substructure with redshift based on a robust nonparametric statistical test, one can speculate about the fact that there are a larger fraction of systems in the low-z subsample with unusually low w and $A_{\rm phot}$, compared to the high-z subsample. Based on the KS and AD tests, and a robust comparison of the medians in different subsamples, however, we cannot call this a significant effect. Both subsamples could be drawn from the same underlying distribution.

5.3. Comparison with Simulated Clusters

In general, the simulated clusters studied here (Nagai et al. 2007) look quantitatively similar to the observed clusters. Distributions of both w and $A_{\rm phot}$ for real and simulated clusters of the same mass and redshift were statistically indistinguishable. Given that these simulations did not include complex astrophysics, such as AGN feedback, we can conclude that, while the development of dense cool cores is sensitive to the specifics of the feedback prescriptions (e.g., Gaspari et al. 2014), overall asymmetry in the ICM is not. This is perhaps not surprising, given that gravitational processes (i.e., mergers) are the dominant source of asymmetry in the ICM.

We can conclude from this work that no more complicated physics is necessary to broadly match the morphology of real and simulated clusters than was included in the *Omega*500 simulations. The combined effects of physical processes including AGN feedback, non-ideal inviscid fluids, and cosmic rays are minimal, and do not significantly bias the observed morphology.

6. Conclusions

Using samples of 36 X-ray selected clusters from the 400 deg^2 *ROSAT* survey, 91 SZ-selected clusters from the SPT 2500 deg^2 survey, and 85 simulated clusters from the *Omega*500 simulations, all observed (or mock observed) to roughly equal depth with the *Chandra X-ray Observatory*, we investigated whether these samples have any bias toward cluster morphological type, and whether high-redshift clusters are more disturbed than their low-redshift counterparts. We considered two well-defined substructure statistics and tested for statistically significant differences in their distributions between different subsamples. In the mass and redshift range studied, we find no evidence for a statistically significant difference in the X-ray morphologies of clusters selected via X-ray or SZ, or at low or high redshift. Furthermore, we found that simulated clusters had quantitatively similar morphology to X-ray- and SZ-selected systems, considering only the asymmetry of the hot gas (i.e., ignoring central cusps).

Our results demonstrate that there is no significant bias for or against preferentially selecting mergers in high resolution ($\sim 1'$) SZ surveys. For SZ surveys with larger beam size (e.g., *Planck*), morphological biases may exist due to the fact that multiple clusters or extended structures can contribute to the integrated signal.

Much of this work was enabled by generous GTO contributions from Steve Murray, and was in progress at the time of his untimely death in 2015. He was a valued member of the Center for Astrophysics and a strong supporter of SPT science; he will be greatly missed by all of us. M.M. acknowledges support for X-ray analysis by NASA through *Chandra* Award Numbers 13800883 and 16800690 issued by the *Chandra* X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of NASA. The South Pole Telescope program is supported by the National Science Foundation through grants ANT-0638937 and PLR-1248097. Argonne National Laboratory's work was supported under U.S. Department of Energy contract DE-AC02-06CH11357. This work was partially completed at Fermilab, operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy. Partial support is also provided by the NSF Physics Frontier Center grant PHY-0114422 to the Kavli Institute of Cosmological Physics at the University of Chicago, the Kavli Foundation, and the Gordon and Betty Moore Foundation. D.N. is supported in part by NSF grant AST-1412768, NASA *Chandra* Theory grant GO213004B, and by the facilities and staff of the Yale Center for Research Computing. D.R. is supported by a NASA Postdoctoral Program Senior Fellowship at the NASA Ames Research Center, administered by the Universities Space Research Association under contract with NASA.

References

Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, ARA&A, 49, 409
Anders, E., & Grevesse, N. 1989, GeCoA, 53, 197
Andersson, K., Benson, B. A., Ade, P. A. R., et al. 2011, ApJ, 738, 48
Andersson, K., Peterson, J. R., Madejski, G., & Goobar, A. 2009, ApJ, 696, 1029
Angulo, R. E., Springel, V., White, S. D. M., et al. 2012, MNRAS, 426, 2046
Applegate, D. E., von der Linden, A., Kelly, P. L., et al. 2014, MNRAS, 439, 48
Benjamini, Y., & Hochberg, Y. 1995, Journal of the Royal Statistical Society. Series B (Methodological), 57, 289
Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, ApJS, 216, 27
Böhringer, H., Pratt, G. W., Arnaud, M., et al. 2010, A&A, 514, A32
Buote, D. A., & Tsai, J. C. 1995, ApJ, 452, 522
Burenin, R. A., Vikhlinin, A., Hornstrup, A., et al. 2007, ApJS, 172, 561
Caldwell, C. E., McCarthy, I. G., Baldry, I. K., et al. 2016, MNRAS, 462, 4117
Cameron, E. 2011, PASA, 28, 128
Eckert, D., Molendi, S., & Paltani, S. 2011, A&A, 526, A79
Evrard, A. E., Mohr, J. J., Fabricant, D. G., & Geller, M. J. 1993, ApJL, 419, L9
Fakhouri, O., & Ma, C. 2010, MNRAS, 401, 2245
Feigelson, E. D., & Babu, G. J. 2012, Modern Statistical Methods for Astronomy (Cambridge: Cambridge Univ. Press)
Gaspari, M., Brighenti, F., Temi, P., & Ettori, S. 2014, ApJL, 783, L10
Hallman, E. J., & Jeltema, T. E. 2011, MNRAS, 418, 2467
Hoekstra, H., Herbonnet, R., Muzzin, A., et al. 2015, MNRAS, 449, 685
Hudson, D. S., Mittal, R., Reiprich, T. H., et al. 2010, A&A, 513, A37
Jeltema, T. E., Canizares, C. R., Bautz, M. W., & Buote, D. A. 2005, ApJ, 624, 606
Jing, Y. P., Mo, H. J., Borner, G., & Fang, L. Z. 1995, MNRAS, 276, 417
Jones, C., Forman, W. R., Andrade-Santos, F., et al. 2015, in American Astronomical Society Meeting Abstracts AA# 225, 321.04
Kaastra, J. S., & Mewe, R. 1993, A&AS, 97, 443
Klypin, A., Kravtsov, A. V., Bullock, J. S., & Primack, J. R. 2001, ApJ, 554, 903
Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330
Kravtsov, A. V. 1999, PhD thesis, New Mexico State Univ.
Kravtsov, A. V., Klypin, A., & Hoffman, Y. 2002, ApJ, 571, 563
Liedahl, D. A., Osterheld, A. L., & Goldstein, W. H. 1995, ApJL, 438, L115
Lin, H. W., McDonald, M., Benson, B., & Miller, E. 2015, ApJ, 802, 34
Mann, A. W., & Ebeling, H. 2012, MNRAS, 420, 2120
Mantz, A. B., Allen, S. W., Morris, R. G., et al. 2015, MNRAS, 449, 199
Maughan, B. J., Giles, P. A., Randall, S. W., Jones, C., & Forman, W. R. 2012, MNRAS, 421, 1583
McDonald, M. 2011, ApJL, 742, L35
McDonald, M., Bayliss, M., Benson, B. A., et al. 2012, Natur, 488, 349
McDonald, M., Benson, B. A., Vikhlinin, A., et al. 2013, ApJ, 774, 23
McDonald, M., Benson, B. A., Vikhlinin, A., et al. 2014, ApJ, 794, 67
Menanteau, F., Hughes, J. P., Sifon, C., et al. 2012, ApJ, 748, 7
Mewe, R., Lemen, J. R., & van den Oord, G. H. J. 1986, A&AS, 65, 511
Mittal, R., Hicks, A., Reiprich, T. H., & Jaritz, V. 2011, A&A, 532, A133
Mohr, J. J., Evrard, A. E., Fabricant, D. G., & Geller, M. J. 1995, ApJ, 447, 8
Mohr, J. J., Fabricant, D. G., & Geller, M. J. 1993, ApJ, 413, 492
Motl, P. M., Hallman, E. J., Burns, J. O., & Norman, M. L. 2005, ApJL, 623, L63
Nagai, D., Kravtsov, A. V., & Vikhlinin, A. 2007, ApJ, 668, 1
Nelson, K., Lau, E. T., Nagai, D., Rudd, D. H., & Yu, L. 2014, ApJ, 782, 107
Ntampaka, M., Trac, H., Cisewski, J., & Price, L. C. 2016, ApJ, 831, 135
Nurgaliev, D., McDonald, M., Benson, B. A., et al. 2013, ApJ, 779, 112
O'Hara, T. B., Mohr, J. J., Bialek, J. J., & Evrard, A. E. 2006, ApJ, 639, 64
Pewsey, A., Neuhäuser, M., & Ruxton, G. D. 2015, Circular Statistics in R (Oxford: Oxford Univ. Press)
Pierre, M., Valotti, A., Faccioli, L., et al. 2016, A&A, submitted (arXiv:160907762P)
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013a, A&A, 550, A132
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013b, A&A, 550, A130
Planck Collaboration, Aghanim, N., Arnaud, M., et al. 2011, A&A, 536, A9
Planck Collaboration, Aghanim, N., Arnaud, M., et al. 2012, A&A, 543, A102
Poole, G. B., Fardal, M. A., Babul, A., et al. 2006, MNRAS, 373, 881
Reiprich, T. H., & Böhringer, H. 2002, ApJ, 567, 716
Richstone, D., Loeb, A., & Turner, E. L. 1992, ApJ, 393, 477
Rossetti, M., Gastaldello, F., Ferioli, G., et al. 2016, MNRAS, 457, 4515
Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, ApJ, 672, 19
Samuele, R., McNamara, B. R., Vikhlinin, A., & Mullis, C. R. 2011, ApJ, 731, 31
Santos, J. S., Tozzi, P., Rosati, P., & Böhringer, H. 2010, A&A, 521, A64
Schuecker, P., Böhringer, H., Reiprich, T. H., & Feretti, L. 2001, A&A, 378, 408
Semler, D. R., Šuhada, R., Aird, K. A., et al. 2012, ApJ, 761, 183
Sifón, C., Battaglia, N., Hasselfield, M., et al. 2016, MNRAS, 461, 248
Sifón, C., Menanteau, F., Hasselfield, M., et al. 2013, ApJ, 772, 25
Swetz, D. S., Ade, P. A. R., Amiri, M., et al. 2011, ApJS, 194, 41
Vikhlinin, A. 2006, ApJ, 640, 710
Vikhlinin, A., Burenin, R., Forman, W. R., et al. 2007, in Heating versus Cooling in Galaxies and Clusters of Galaxies, ed. H. Böhringer et al. (Berlin; Heidelberg: Springer-Verlag), 48
Vikhlinin, A., Burenin, R. A., Ebeling, H., et al. 2009, ApJ, 692, 1033
Vikhlinin, A., McNamara, B. R., Forman, W., et al. 1998, ApJL, 498, L21
Weißmann, A., Böhringer, H., & Chon, G. 2013, A&A, 555, A147