CHANDRA CLUSTER COSMOLOGY PROJECT. II. SAMPLES AND X-RAY DATA REDUCTION

A. VIKHLININ[1,2], R. A. BURENIN[2], H. EBELING[3], W. R. FORMAN[1], A. HORNSTRUP[4], C. JONES[1], A. V. KRAVTSOV[5], S. S. MURRAY[1],
D. NAGAI[6], H. QUINTANA[7], AND A. VOEVODKIN[2,8]

[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[2] Space Research Institute (IKI), Profsoyuznaya 84/32, Moscow, Russia
[3] Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[4] Danish National Space Center, Juliane Maries Vej 30, Copenhagen 0, DK-2100, Denmark
[5] Department of Astronomy and Astrophysics, Kavli Institute for Cosmological Physics, Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA
[6] Department of Physics and Yale Center for Astronomy & Astrophysics, Yale University, New Haven, CT 06520, USA
[7] Departamento de Astronomia y Astrofisica, Pontificia Universidad Catolica de Chile, Casilla 306, Santiago, 22, Chile
[8] Los Alamos National Laboratory, Los Alamos, NM 87545, USA

ABSTRACT

We discuss the measurements of the galaxy cluster mass functions at $z \approx 0.05$ and $z \approx 0.5$ using high-quality *Chandra* observations of samples derived from the *ROSAT* PSPC All-Sky and 400 deg^2 surveys. We provide a full reference for the data analysis procedures, present updated calibration of relations between the total cluster mass and its X-ray indicators (T_X, M_{gas}, and Y_X) based on a subsample of low-z relaxed clusters, and present a first measurement of the evolving L_X–M_{tot} relation (with M_{tot} estimated from Y_X) obtained from a well defined statistically complete cluster sample and with appropriate corrections for the Malmquist bias applied. Finally, we present the derived cluster mass functions, estimate the systematic uncertainties in this measurement, and discuss the calculation of the likelihood function. We confidently measure the evolution in the cluster comoving number density at a fixed mass threshold, e.g., by a factor of 5.0 ± 1.2 at $M_{500} = 2.5 \times 10^{14} \ h^{-1} \ M_\odot$ between $z = 0$ and 0.5. This evolution reflects the growth of density perturbations, and can be used for the cosmological constraints complementing those from the distance–redshift relation.

Key words: catalogs – galaxies: clusters: general – surveys – X-rays: galaxies

1. INTRODUCTION

This work continues a series of papers in which we present the data for a new X-ray selected sample of galaxy clusters—the 400d survey—based on the data from the *ROSAT* PSPC pointed observations. In the first paper (Burenin et al. 2007, Paper I hereafter), we presented the cluster catalog and described the survey's statistical calibration (selection function, effective area and so on). A complete high-redshift subsample of the 400d clusters, 36 objects at $z = 0.35 - 0.9$ with $\langle z \rangle = 0.5$, has been observed with *Chandra*. The goal of this program was to provide X-ray data of sufficient quality for reliable estimates of the high-redshift ($z \sim 0.5$) cluster mass function.

Chandra exposures were designed to yield at least 1500–2000 photons from each cluster. This is sufficient to measure several high-quality total mass proxies—average temperature excluding the center, integrated gas mass, and the Y_X parameter (the product of tempeature and gas mass derived from X-ray data). The resulting mass estimates are much more reliable than what was achievable in many previous studies where the only available mass indicator was the X-ray flux (Kravtsov et al. 2006). Using several mass proxies also allows us to control the systematics by checking the consistency of results obtained by different methods.

Observations of the high-redshift 400d clusters are complemented by *Chandra* archival data for a complete, flux-limited sample of nearby clusters detected in the *ROSAT* All-Sky survey (49 objects at present, expected to grow by a factor of 1.5 in the near future as the completeness of the *Chandra* archive expands to lower fluxes). *Chandra* data for nearby clusters, combined when necessary with the *ROSAT* PSPC pointings, allow us to measure the same set of total mass proxies in local and distant clusters.

The present work is a significant step forward in providing observational foundations for cosmological work with the cluster mass function. First, it uses a larger sample of high-z clusters than the previous studies. For example, the best published measurement of the evolution in the cluster temperature function (Henry 2004) was based on 25 low-z objects and 19 clusters with $\langle z \rangle = 0.43$. *Chandra* provides much higher-quality data for each high-z object than were available before. Second, we use a more advanced approach to the X-ray data analysis, partly because this is called for by the *Chandra* data and partly because of the experience learned from recent deep observations of low-z clusters (e.g., Vikhlinin et al. 2006). Last but not least, the data for high- and low-z samples were obtained with the same instrument and analyzed uniformly, minimizing the potential for systematic errors—the crucial ingredient for precise measurement of the evolution of the cluster mass function.

In this paper, we present an analysis of the *Chandra* observations of our cluster sample, describe our approach to the cluster total mass estimates, derive the evolving M–L_X relation, and describe the computation of the survey volume as a function of mass. We conclude by presenting the cluster mass functions estimated in the "concordant" ΛCDM cosmology. The cosmological modeling of the cluster mass function data is presented in an accompanying paper (Vikhlinin et al. 2009, Paper III hereafter). The prime goal of this work is to provide a full reference of the data reduction procedures and discuss the sources of systematic uncertainties in the cluster mass function estimates at low and high redshifts.

All distance-dependent quantities are computed assuming a ΛCDM cosmological model with $\Omega_M = 0.30$, and $\Omega_\Lambda = 0.70$. We also assume $h = 0.72$, unless the explicit h-scaling is given. The luminosities and fluxes are in the 0.5–2 keV energy band.
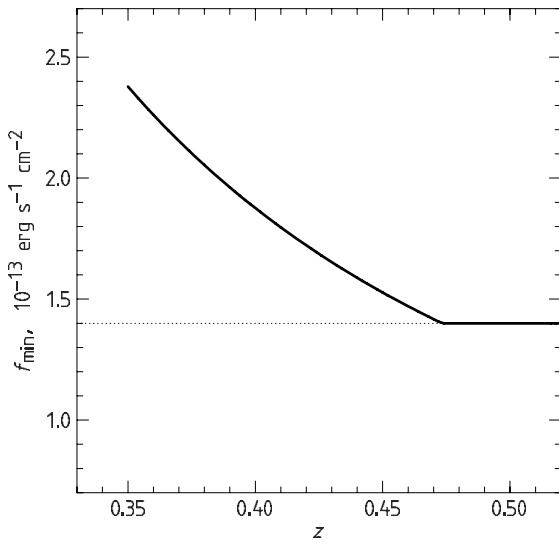
Figure 1. Limiting *ROSAT* flux for selection in the 400d *Chandra* sample, as a function of redshift. At $z > 0.473$, the limiting flux is 1.4×10^{-13}, that in the 400d catalog. At $0.35 < z < 0.473$, the flux limit corresponds to the minimum luminosity specified in Equation (1).

2. CLUSTER SAMPLES

2.1. High-Redshift Sample

Our high-redshift cluster sample is a well defined subsample of the $z > 0.35$ clusters from the 400d survey. The selection was designed to provide a quasi mass-limited sample at $z \lesssim 0.5$ by requiring that the *ROSAT*-derived luminosity was above a threshold of

$$L_{X,\text{min}} = 4.8 \times 10^{43}(1 + z)^{1.8} \text{ erg s}^{-1} \qquad (1)$$

in the default ΛCDM cosmology. This luminosity threshold approximately corresponds to a mass limit of $10^{14} M_\odot$ from the low-z L_X-M relation. The redshift factor here corresponds to an early measurement of the evolution in the $M_{\text{gas}}-L$ relation (Vikhlinin et al. 2002). The resulting selection is entirely objective and in fact is formulated as a redshift-dependent flux limit (shown in Figure 1). At $z > 0.473$, no additional selection is applied since the minimum flux of the main 400d sample satisfies the luminosity threshold in Equation (1).

Thirty-nine objects from the 400d catalog satisfy these selection criteria, and all were observed with *Chandra*. For three clusters, 0216−1747, 0521−2530, 1117+1744, the accurate total X-ray flux measured by *Chandra* was $< 10^{-13}$ erg s^{-1} cm^{-2}, significantly below the target minimum flux in the 400d catalog, 1.4×10^{-13} erg s^{-1} cm^{-2}. The existence of such deviations is expected (see, e.g., Figure 23 in Paper I) because *ROSAT* flux estimates have large statistical errors. However, the computation of the 400d selection function in this flux regime is less accurate because it depends strongly on the wings of the distribution of the flux measurement scatter (see Section 7.1 in Paper I for details). We, therefore, opted not to use these three clusters in the further analysis. The additional selection criterion, $f_{\text{true}} > 10^{-13}$ erg s^{-1} cm^{-2}, will be taken into account in the sample volume computations. The final sample of 36 high-redshift clusters we will use hereafter is presented in Table 1.

2.2. Low-Redshift Sample

The low-redshift cluster sample was selected, similarly to the procedure described in Voevodkin & Vikhlinin (2004), from several samples based on the *ROSAT* All-Sky Survey

(RASS) data (Ebeling et al. 2000: BCS; de Grandi et al. 1999; Böhringer et al. 2004: REFLEX; Reiprich & Böhringer 2002: HIGFLUGCS). Overlaps between the catalogs were removed. The objects at Galactic latitude $|b| < 20°$, as well as those around LMC, SMC, and the Virgo cluster were excluded (the exclusion regions were adopted from Reiprich & Böhringer 2002). The total area covered by these catalogs is 8.14 sr. The X-ray fluxes were remeasured (starting from a list of objects with cataloged fluxes $f > 5.3 \times 10^{-12}$ erg s^{-1} cm^{-2} in the 0.5–2 keV band), using the data from pointed *ROSAT* PSPC observations, when available. Our final sample consists of 49 clusters (Table 2) with the remeasured flux $f > 1.3 \times 10^{-11}$ erg s^{-1} cm^{-2} in the 0.5–2 keV band, well above the sensitivity limit of all initial RASS cluster catalogs, and $z > 0.025$ (the lower redshift cut was used to ensure that a large fraction of the cluster virial radius fits inside the *Chandra* field of view). All objects in this sample have archival *Chandra* observations, providing accurate X-ray spectral data.

2.3. General Characteristics of the Cluster Samples

The combined cluster sample is a unique, uniformly observed dataset. The volume coverage and effective mass limits in the low and high-redshift subsamples are similar ($M_{\text{min}} \simeq (1–2) \times 10^{14} h^{-1} M_\odot$). The median mass at all redshifts is near $M_{500} = 2.5 \times 10^{14} h^{-1} M_\odot$, which corresponds to $T = 4.5$ keV clusters at $z = 0$. Observations suggest that clusters of such and larger mass exhibit scalings between their observables and mass close to the expectations of self-similar model (Nagai et al. 2007a), which makes our sample particularly useful for cosmological applications.

Our cluster sample is selected essentially using only the X-ray flux. Cluster detection efficiency is, in principle, also depends on the object surface brightness. However, the surface brightness effects are minimal for our objects. For the low-z sample, this is achieved by selecting objects with fluxes a factor of > 5 higher than the detection threshold in the parent RASS samples. For the 400d clusters, we used a highly sensitive detection method tailored for finding extended sources. The resulting sensitivity of the detection efficiency to the cluster angular size has been extensively studied (Paper I) and found to be small. Furthermore, optical identifications also played no role in selecting the sample—essentially all X-ray candidates at both low and high redshifts were identified as galaxy clusters. Therefore, we do not miss objects because of misclassification caused by the presence of central or background active galactic nuclei (AGNs).

The redshift histograms for the low and high-redshift samples are shown in Figure 2. The depth of the low-redshift sample is $z \sim 0.15$; there are only three clusters beyond this z. Therefore, the low-redshift sample is effectively "local" and it gives us a snapshot of the cluster population at $z \approx 0$. The high-redshift sample starts at $z = 0.35$ and extends to $z = 0.9$. The median redshift of the distant sample is $\langle z \rangle = 0.5$.

Data of sufficient quality are available for utilizing three different X-ray total mass proxies for all our clusters. These observations provide us with a reliable measure of the evolution of the cluster mass function between $z \approx 0.5$ and 0, or over $\approx 37\%$ of the present age of the universe.

3. *CHANDRA* AND *ROSAT* DATA REDUCTION

Chandra observations provide the basis for our X-ray analysis of both high- and low-redshift clusters. We also make use of the

Table 1
High-redshift Cluster Sample

Name	z	L_X, (erg s^{-1})	T_X (keV)	M_Y ($10^{14} M_\odot$)	M_G ($10^{14} M_\odot$)	M_T ($10^{14} M_\odot$)	f_x, (10^{-13} cgs)	f_{ROSAT}, (10^{-13} cgs)	Merger?
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0302−0423	0.3501	5.24×10^{44}	4.78 ± 0.75	3.72 ± 0.38	3.58 ± 0.28	3.26 ± 0.77	15.34	15.9 ± 1.9	$\cdots$
1212+2733	0.3533	3.61×10^{44}	6.62 ± 0.89	6.17 ± 0.57	5.62 ± 0.37	6.16 ± 1.24	10.53	12.5 ± 1.7	$\checkmark$
0350−3801	0.3631	6.80×10^{43}	2.45 ± 0.50	1.43 ± 0.19	1.40 ± 0.18	1.34 ± 0.41	1.68	2.9 ± 0.8	$\checkmark$
0318−0302	0.3700	1.82×10^{44}	4.04 ± 0.63	2.82 ± 0.28	2.44 ± 0.21	2.86 ± 0.67	4.63	4.6 ± 0.5	$\checkmark$
0159+0030	0.3860	1.42×10^{44}	4.25 ± 0.96	2.51 ± 0.37	1.92 ± 0.22	2.67 ± 0.90	3.30	3.3 ± 0.4	$\cdots$
0958+4702	0.3900	1.04×10^{44}	3.57 ± 0.73	1.84 ± 0.25	1.34 ± 0.15	2.03 ± 0.63	2.22	2.8 ± 0.6	$\cdots$
0809+2811	0.3990	2.50×10^{44}	4.17 ± 0.73	3.69 ± 0.42	3.98 ± 0.35	2.96 ± 0.78	5.40	5.5 ± 0.8	$\checkmark$
1416+4446	0.4000	1.94×10^{44}	3.26 ± 0.46	2.52 ± 0.24	3.10 ± 0.24	1.76 ± 0.37	4.01	4.0 ± 0.5	$\cdots$
1312+3900	0.4037	1.37×10^{44}	3.72 ± 1.06	2.75 ± 0.57	2.62 ± 0.42	2.47 ± 1.06	2.71	2.6 ± 0.4	$\checkmark$
1003+3253	0.4161	1.53×10^{44}	5.44 ± 1.40	2.80 ± 0.49	1.57 ± 0.20	3.83 ± 1.47	3.04	3.5 ± 0.4	$\cdots$
0141−3034	0.4423	1.32×10^{44}	2.13 ± 0.38	1.22 ± 0.17	1.30 ± 0.24	1.03 ± 0.27	2.06	3.1 ± 0.9	$\checkmark$
1701+6414	0.4530	2.39×10^{44}	4.36 ± 0.46	3.28 ± 0.24	3.20 ± 0.20	2.66 ± 0.42	3.91	3.9 ± 0.4	$\cdots$
1641+4001	0.4640	9.46×10^{43}	3.31 ± 0.62	1.70 ± 0.20	1.34 ± 0.13	1.73 ± 0.49	1.43	2.9 ± 0.8	$\cdots$
0522−3624	0.4720	1.04×10^{44}	3.46 ± 0.48	2.18 ± 0.21	1.82 ± 0.15	2.12 ± 0.45	1.47	1.8 ± 0.3	$\checkmark$
1222+2709	0.4720	9.88×10^{43}	3.74 ± 0.61	2.09 ± 0.24	1.59 ± 0.16	2.08 ± 0.51	1.39	1.9 ± 0.4	$\cdots$
0355−3741	0.4730	1.76×10^{44}	4.61 ± 0.82	3.02 ± 0.35	2.44 ± 0.22	2.87 ± 0.76	2.48	2.9 ± 0.7	$\cdots$
0853+5759	0.4750	8.43×10^{43}	3.42 ± 0.67	2.05 ± 0.27	1.63 ± 0.17	2.09 ± 0.61	1.22	2.0 ± 0.5	$\checkmark$
0333−2456	0.4751	9.79×10^{43}	3.16 ± 0.58	1.90 ± 0.22	1.64 ± 0.17	1.85 ± 0.51	1.33	2.4 ± 0.5	$\checkmark$
0926+1242	0.4890	1.50×10^{44}	4.74 ± 0.71	3.00 ± 0.30	2.02 ± 0.16	3.42 ± 0.77	2.04	1.7 ± 0.3	$\checkmark$
0030+2618	0.5000	1.57×10^{44}	5.63 ± 1.13	3.43 ± 0.41	2.04 ± 0.19	4.41 ± 1.33	2.09	2.4 ± 0.3	$\checkmark$
1002+6858	0.5000	1.71×10^{44}	4.04 ± 0.83	2.80 ± 0.40	2.34 ± 0.27	2.65 ± 0.81	2.19	2.0 ± 0.4	$\checkmark$
1524+0957	0.5160	2.07×10^{44}	4.23 ± 0.51	3.24 ± 0.27	3.08 ± 0.21	2.82 ± 0.51	2.45	3.0 ± 0.4	$\checkmark$
1357+6232	0.5250	1.63×10^{44}	4.60 ± 0.69	2.96 ± 0.29	2.40 ± 0.18	2.78 ± 0.62	1.90	2.0 ± 0.3	$\cdots$
1354−0221	0.5460	1.40×10^{44}	3.77 ± 0.53	2.31 ± 0.23	1.69 ± 0.16	2.32 ± 0.48	1.45	1.5 ± 0.2	$\checkmark$
1120+2326	0.5620	1.79×10^{44}	3.58 ± 0.44	2.50 ± 0.21	2.32 ± 0.16	2.13 ± 0.39	1.68	2.1 ± 0.4	$\checkmark$
0956+4107	0.5870	1.85×10^{44}	4.40 ± 0.50	2.93 ± 0.22	2.44 ± 0.14	2.87 ± 0.49	1.64	1.6 ± 0.3	$\checkmark$
0328−2140	0.5901	2.30×10^{44}	5.14 ± 1.47	3.42 ± 0.66	2.92 ± 0.38	3.17 ± 1.36	2.09	2.1 ± 0.6	$\cdots$
1120+4318	0.6000	3.75×10^{44}	4.99 ± 0.30	3.92 ± 0.17	4.20 ± 0.24	3.00 ± 0.27	3.24	3.0 ± 0.3	$\cdots$
1334+5031	0.6200	2.22×10^{44}	4.31 ± 0.28	2.62 ± 0.17	1.88 ± 0.22	2.73 ± 0.27	1.76	1.8 ± 0.3	$\checkmark$
0542−4100	0.6420	2.91×10^{44}	5.45 ± 0.77	4.07 ± 0.39	3.70 ± 0.25	3.86 ± 0.82	2.21	2.2 ± 0.3	$\checkmark$
1202+5751	0.6775	2.22×10^{44}	4.08 ± 0.72	2.90 ± 0.37	2.85 ± 0.29	2.42 ± 0.64	1.34	1.5 ± 0.4	$\checkmark$
0405−4100	0.6861	2.23×10^{44}	3.98 ± 0.48	2.51 ± 0.20	2.17 ± 0.16	2.32 ± 0.42	1.33	1.5 ± 0.4	$\checkmark$
1221+4918	0.7000	3.35×10^{44}	6.63 ± 0.75	4.88 ± 0.38	4.16 ± 0.23	5.04 ± 0.86	2.06	2.1 ± 0.5	$\checkmark$
0230+1836	0.7990	2.55×10^{44}	5.50 ± 1.02	3.46 ± 0.46	2.70 ± 0.27	3.57 ± 0.99	1.09	2.2 ± 0.6	$\checkmark$
0152−1358	0.8325	5.46×10^{44}	5.40 ± 0.97	3.91 ± 0.52	3.94 ± 0.40	3.40 ± 0.91	2.24	1.8 ± 0.3	$\checkmark$
1226+3332	0.8880	8.42×10^{44}	11.08 ± 1.39	7.59 ± 0.61	5.75 ± 0.28	9.91 ± 1.86	3.27	2.9 ± 0.3	$\checkmark$

Notes. Column 2: cluster redshift; Column 3: total X-ray luminosity (0.5–2 keV band, object frame) measured from accurate *Chandra* flux; Column 4: average temperature from the spectrum integrated in the $[0.15 - 1] \, r_{500}$ annulus; Column 5: total mass estimated from Y_X parameter (Section 4.3); Column 6: M_{tot} estimated from integrated gas mass (Section 4.2); Column 7: mass estimated from the $M_{tot}-T_X$ relation (Section 4.1); Column 8: total X-ray flux measured by *Chandra* (0.5–2 keV, observer's frame); Column 9: total X-ray flux (0.5–2 keV, observer's frame) reported in the 400d catalog from *ROSAT* PSPC data; Column 10: approximate classification into mergers and relaxed clusters (Section 4.1.3).

ROSAT PSPC data for the low-z objects (pointed observations when available and All-Sky Survey data for eight objects). In low-z clusters, the statistical accuracy of the X-ray surface brightness determination at large radii is limited mostly by the *Chandra* field of view. The analysis in such cases benefits from using the *ROSAT* data that cover a much larger region although with a lower sensitivity. Below, we discuss the issues related to the initial data preparation, spectral analysis, and producing the "calibrated" X-ray images. How these data are used to derive the basic intracluster medium (ICM) parameters and the cluster M_{tot} is discussed in Sections 3.3, 3.4, and 4.

3.1. Initial Data Reduction and Calibration Uncertainties

3.1.1. Chandra

For *Chandra*, our data reduction procedure is adopted with no changes from Vikhlinin et al. (2005, V05 hereafter). This includes careful filtering for high background periods and

applying all the latest calibration corrections to the detected X-ray photons, and determination of the background intensity in each observation.

The quiescent *Chandra* background is dominated by the events induced by charged particles. This component can be subtracted exquisitely accurately (with a $\lesssim 2\%$ scatter; see Hickox & Markevitch 2006). A much smaller contribution is provided by a fraction of the cosmic X-ray background not resolved into discrete sources. This component is modeled adequately by using the "blank-sky" background datasets which include both the particle-induced and unresolved sky components. Finally, there is a non-negligible diffuse soft component attributable to the Galactic interstellar medium (ISM) emission (Markevitch et al. 2003) and, in some cases, to the geocoronal charge exchange (Wargelin et al. 2004). The soft background component is the hardest to model because its intensity depends on the pointing direction, and can even be variable in the case of charge exchange emission. Fortunately, the soft component can

Table 2
Low-redshift Sample

Name	f_x, (10^{-11} cgs)	z^a	L_X, (erg s^{-1})	T_X (keV)	M_Y ($10^{14} M_\odot$)	M_G ($10^{14} M_\odot$)	M_T ($10^{14} M_\odot$)	Merger?
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A 3571	7.42	0.0386	2.37×10^{44}	6.81 ± 0.10	5.90 ± 0.06	5.30 ± 0.07	6.61 ± 0.15	...
A 2199	6.43	0.0304	1.27×10^{44}	3.99 ± 0.10	2.77 ± 0.05	2.80 ± 0.04	2.92 ± 0.11	...
2A 0335	6.24	0.0346	1.60×10^{44}	3.43 ± 0.10	2.33 ± 0.05	2.53 ± 0.05	2.32 ± 0.11	...
A 496	5.33	0.0328	1.23×10^{44}	4.12 ± 0.07	2.96 ± 0.04	3.02 ± 0.04	3.07 ± 0.07	...
A 3667	4.64	0.0557	3.14×10^{44}	6.33 ± 0.06	7.35 ± 0.07	8.62 ± 0.15	6.74 ± 0.09	✓
A 754	4.35	0.0542	2.78×10^{44}	8.73 ± 0.00	8.47 ± 0.13	6.68 ± 0.12	11.05 ± 0.00	✓
A 85	4.30	0.0557	2.91×10^{44}	6.45 ± 0.10	5.98 ± 0.07	5.91 ± 0.10	6.03 ± 0.14	...
A 2029	4.23	0.0779	5.72×10^{44}	8.22 ± 0.16	8.64 ± 0.14	8.35 ± 0.20	8.66 ± 0.25	...
A 478	4.16	0.0881	7.24×10^{44}	7.96 ± 0.27	8.15 ± 0.17	7.82 ± 0.12	8.20 ± 0.42	...
A 1795	4.14	0.0622	3.52×10^{44}	6.14 ± 0.10	5.46 ± 0.06	5.34 ± 0.06	5.58 ± 0.14	...
A 3558	4.11	0.0469	1.96×10^{44}	4.88 ± 0.10	4.78 ± 0.07	5.43 ± 0.09	4.54 ± 0.15	✓
A 2142	3.94	0.0904	7.20×10^{44}	10.04 ± 0.26	11.96 ± 0.20	11.91 ± 0.16	11.70 ± 0.45	...
A 2256	3.61	0.0581	2.66×10^{44}	8.37 ± 0.24	7.84 ± 0.15	6.14 ± 0.09	10.33 ± 0.45	✓
A 4038	3.48	0.0288	6.18×10^{43}	2.61 ± 0.05	1.65 ± 0.02	2.03 ± 0.05	1.52 ± 0.04	...
A 2147	3.47	0.0355	9.40×10^{43}	3.83 ± 0.12	3.10 ± 0.08	3.52 ± 0.14	3.15 ± 0.15	✓
A 3266	3.39	0.0602	2.69×10^{44}	8.63 ± 0.18	9.00 ± 0.13	7.66 ± 0.12	10.82 ± 0.34	✓
A 401	3.19	0.0743	3.90×10^{44}	7.72 ± 0.30	8.63 ± 0.24	9.27 ± 0.20	7.88 ± 0.46	...
A 2052	2.93	0.0345	7.47×10^{43}	3.03 ± 0.07	1.84 ± 0.03	1.95 ± 0.04	1.91 ± 0.07	...
Hydra-A	2.91	0.0549	1.93×10^{44}	3.64 ± 0.06	2.83 ± 0.03	3.34 ± 0.04	2.51 ± 0.06	...
A 119	2.47	0.0445	1.06×10^{44}	5.72 ± 0.00	4.50 ± 0.03	3.61 ± 0.06	5.80 ± 0.00	✓
A 2063	2.39	0.0342	5.98×10^{43}	3.57 ± 0.19	2.21 ± 0.08	2.13 ± 0.07	2.46 ± 0.19	...
A 1644	2.33	0.0475	1.14×10^{44}	4.61 ± 0.14	4.21 ± 0.09	4.66 ± 0.11	4.16 ± 0.19	✓
A 3158	2.30	0.0583	1.72×10^{44}	4.67 ± 0.07	4.13 ± 0.05	4.74 ± 0.09	3.67 ± 0.09	...
MKW 3s	2.08	0.0453	9.28×10^{43}	3.03 ± 0.05	2.09 ± 0.03	2.52 ± 0.05	1.90 ± 0.05	...
A 1736	2.04	0.0449	8.94×10^{43}	2.95 ± 0.09	2.10 ± 0.06	2.10 ± 0.13	2.10 ± 0.09	✓
EXO 0422	2.01	0.0382	6.35×10^{43}	2.84 ± 0.09	1.51 ± 0.04	1.46 ± 0.06	1.73 ± 0.09	...
A 4059	2.00	0.0491	1.05×10^{44}	4.25 ± 0.08	2.81 ± 0.04	2.58 ± 0.04	3.19 ± 0.09	...
A 3395	1.95	0.0506	1.09×10^{44}	5.10 ± 0.17	6.74 ± 0.18	6.74 ± 0.20	6.74 ± 0.34	✓
A 2589	1.94	0.0411	7.09×10^{43}	3.17 ± 0.27	1.94 ± 0.11	2.01 ± 0.10	2.04 ± 0.26	...
A 3112	1.89	0.0759	2.43×10^{44}	5.19 ± 0.21	4.20 ± 0.11	4.12 ± 0.09	4.28 ± 0.26	...
A 3562	1.84	0.0489	9.58×10^{43}	4.31 ± 0.12	3.28 ± 0.07	3.48 ± 0.09	3.26 ± 0.14	...
A 1651	1.80	0.0853	2.93×10^{44}	6.41 ± 0.25	5.78 ± 0.15	5.55 ± 0.12	5.89 ± 0.35	...
A 399	1.78	0.0713	2.01×10^{44}	6.49 ± 0.17	6.18 ± 0.11	5.66 ± 0.12	6.95 ± 0.27	✓
A 2204	1.74	0.1511	9.35×10^{44}	8.55 ± 0.58	9.40 ± 0.43	9.32 ± 0.28	8.87 ± 0.90	...
A 576	1.72	0.0401	5.99×10^{43}	3.68 ± 0.11	2.34 ± 0.05	2.27 ± 0.06	2.57 ± 0.12	...
A 2657	1.62	0.0402	5.66×10^{43}	3.62 ± 0.15	2.24 ± 0.06	2.14 ± 0.05	2.51 ± 0.16	...
A 2634	1.61	0.0305	3.20×10^{43}	2.96 ± 0.09	1.74 ± 0.04	1.83 ± 0.04	1.85 ± 0.08	...
A 3391	1.58	0.0551	1.05×10^{44}	5.39 ± 0.19	4.06 ± 0.10	3.58 ± 0.09	4.58 ± 0.24	...
A 2065	1.56	0.0723	1.82×10^{44}	5.44 ± 0.09	4.98 ± 0.09	4.90 ± 0.09	5.31 ± 0.14	✓
A 1650	1.53	0.0823	2.33×10^{44}	5.29 ± 0.17	4.59 ± 0.11	4.78 ± 0.10	4.39 ± 0.21	...
A 3822	1.48	0.0760	1.91×10^{44}	5.23 ± 0.30	4.63 ± 0.18	4.50 ± 3.91	4.98 ± 0.43	✓
S 1101	1.46	0.0564	1.03×10^{44}	2.44 ± 0.08	1.57 ± 0.03	1.99 ± 0.05	1.36 ± 0.07	...
A 2163	1.38	0.2030	1.37×10^{45}	14.72 ± 0.31	21.98 ± 0.31	24.17 ± 0.34	22.83 ± 0.72	✓
Zw Cl1215	1.38	0.0767	1.80×10^{44}	6.54 ± 0.21	5.75 ± 0.12	5.32 ± 0.10	6.10 ± 0.29	...
RX J1504	1.35	0.2169	1.56×10^{45}	9.89 ± 0.53	10.07 ± 0.35	9.01 ± 0.20	10.70 ± 0.86	...
A 2597	1.35	0.0830	2.09×10^{44}	3.87 ± 0.11	2.84 ± 0.06	3.03 ± 0.06	2.72 ± 0.12	...
A 133	1.35	0.0569	9.60×10^{43}	4.01 ± 0.11	2.57 ± 0.05	2.37 ± 0.04	2.91 ± 0.12	...
A 2244	1.34	0.0989	2.98×10^{44}	5.37 ± 0.12	5.11 ± 0.08	5.80 ± 0.10	4.46 ± 0.15	...
A 3376	1.31	0.0455	5.89×10^{43}	4.37 ± 0.13	3.01 ± 0.07	2.53 ± 0.06	3.84 ± 0.17	✓

Notes. Columns 3–9 have the same meaning as in Table 1. Column 2 gives the total flux (0.5–2 keV) from the best source available (*Chandra* if cluster the cluster is at sufficiently high redshift to fit the field of view, *ROSAT* PSPC pointing, and remeasurement from the All-Sky survey data as a last resort).
a Redshifts were converted to the CMB reference frame.

still be subtracted sufficiently accurately because it is separated spectrally from the cluster emission (since it is dominated by emission lines near 0.6 keV; see Section 2.3.2 in V05).

Uncertainties in determining each of the background components were propagated in the further analysis. Their impact on the analysis of the *Chandra* cluster observations is extensively discussed in V05. Here we only note that this source of uncer-

tainty is negligible for the measurements of the average cluster temperatures dominated by the bright inner region; similarly, the gas mass measurements are based on the surface brightness profiles in the soft band where the background is lower relative to the cluster flux.

Conversion of the observed X-ray fluxes to physical quantities such as the temperature and density of the intracluster gas relies
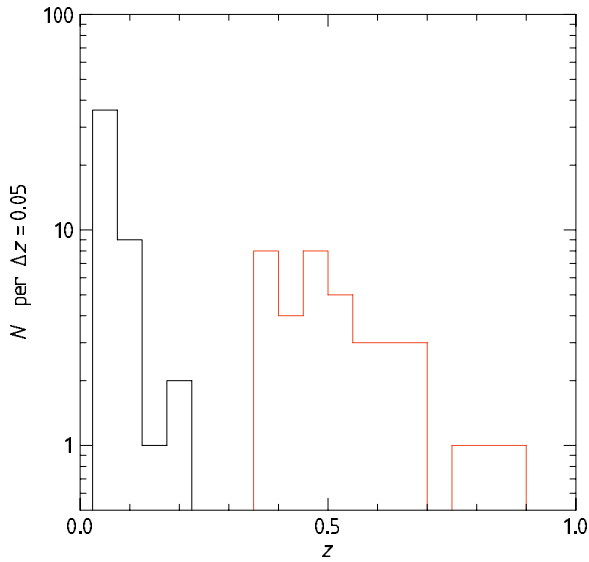
Figure 2. Redshift distribution of clusters in our low- and high-z samples.

on accurate calibration of the spectral response. An extensive preflight calibration program was designed to provide *absolute* calibration of the effective area of the mirror+ACIS system to within 2% at all locations and across the entire energy band. The in-flight performance was degraded somewhat, but by 2005–2006 the calibration accuracy was restored to near-preflight levels. Currently, the uncertainty in relative (position and time-dependent) variations of the effective area is $< 3\%$ within the energy band we use in the present work.[9] The estimated uncertainties in the absolute effective area are $\lesssim 5\%$ at all energies. The systematic effect of such uncertainties on the estimated cluster mass function is small, as summarized in Section 8.1.1. We also note that the calibration uncertainties in the measurement of the *evolution* of the mass function are nearly canceled because we use the same telescope and uniform analysis of both low- and high-z samples.

3.1.2. ROSAT

The *ROSAT* PSPC data were reduced as described in Vikhlinin et al. (1999). The reduction pipeline was based on S. Snowden's software (Snowden et al. 1994). This software eliminates periods of high particle and scattered solar backgrounds as well as those intervals when the detector may be unstable. Exposure maps in several energy bands are then created using detector maps obtained during the RASS. The exposure maps include vignetting and all detector artifacts. The unvignetted particle background is estimated and subtracted from the data, even though the PSPC particle background is low compared to the cosmic X-ray background. The scattered solar X-ray background also should be subtracted separately, because, depending on the viewing angle, it can introduce a constant background gradient across the image. Most solar X-rays were eliminated by simply excluding time intervals when this emission was high, but the remaining contribution was also modeled and subtracted. If the cluster was observed in several pointings, each pointing was reduced individually and the resulting images were merged.

The energy resolution of the *ROSAT* PSPC is insufficient to separate the soft background components spectrally, which was possible in the case of *Chandra*. However, the *ROSAT* field of view is much larger and usually we can reliably measure the uniform background level from the cluster observations themselves. Our procedure for the background determination was to fit the observed surface brightness profile at large radii, $r \gtrsim 0.7\,r_{500}$, to a power law plus constant model (as discussed in Vikhlinin et al. 1999). The additional power law component is required since at lower z, the contribution of the cluster brightness is small but non-negligible even near the edge of the *ROSAT* PSPC field of view. The tests show that this procedure provides a relative uncertainty in the background determination of $\sim 5\%$ (Vikhlinin et al. 1999). This uncertainty was propagated into the further analysis.

The limited bandpass of the *ROSAT* PSPC (limited to $E < 2$ keV) does not allow one to measure the cluster temperatures with an accuracy useful for our purposes. However, the observed count rate can be converted to a broadband flux very reliably, as confirmed directly by excellent agreement with the *Chandra*-derived flux from the same region (see Section 3.4).

3.2. Removal of Substructures and Identification of the Cluster Center

After the initial data preparation, we have flat-fielded and background-subtracted images in the 0.7–2 keV energy band.[10] These images contain only the cluster emission and other X-ray sources. Our next step is to remove all pointlike sources, as well as substructures within the cluster. The point source removal is the most straightforward step. Our detection routine is based on the wavelet decomposition technique documented in Vikhlinin et al. (1998). The point sources are identified using the small scales of the wavelet decomposition and the corresponding regions are masked out from all further analysis. The exclusion radius takes into account the variation of the point-spread function (PSF) size with the offaxis angle (this is especially important in the case of *ROSAT* PSPC pointed observations).

We also mask out any detectable, well-defined substructures within the cluster (they are included only in the total X-ray luminosity). The detection of substructures was fully automatic and based on the analysis of large scales of the wavelet decomposition process. We masked out only the regions associated with the prominent secondary maxima in the X-ray surface brightness, keeping the weaker components such as filamentary structures. Examples are shown in Figures 3 and 4. Removal of obvious substructure reduces the scatter in the relation between the total mass and X-ray proxies, although the effect is small in most cases because we exclude only a small fraction ($< 20\%$) of the total flux. We note that removal of substructures was included in the mock *Chandra* analysis (Nagai et al. 2007b) which we use to assess the uncertainties in the calibrations of the M_{tot} versus proxy relations.

The only quantity we measure without removing the large scale-substructures is the total X-ray luminosity. The luminosity determines the detectability of the cluster in shallow surveys. These surveys usually lack sensitivity and angular resolution to remove the substructure and detect the clusters on the basis of its total flux.
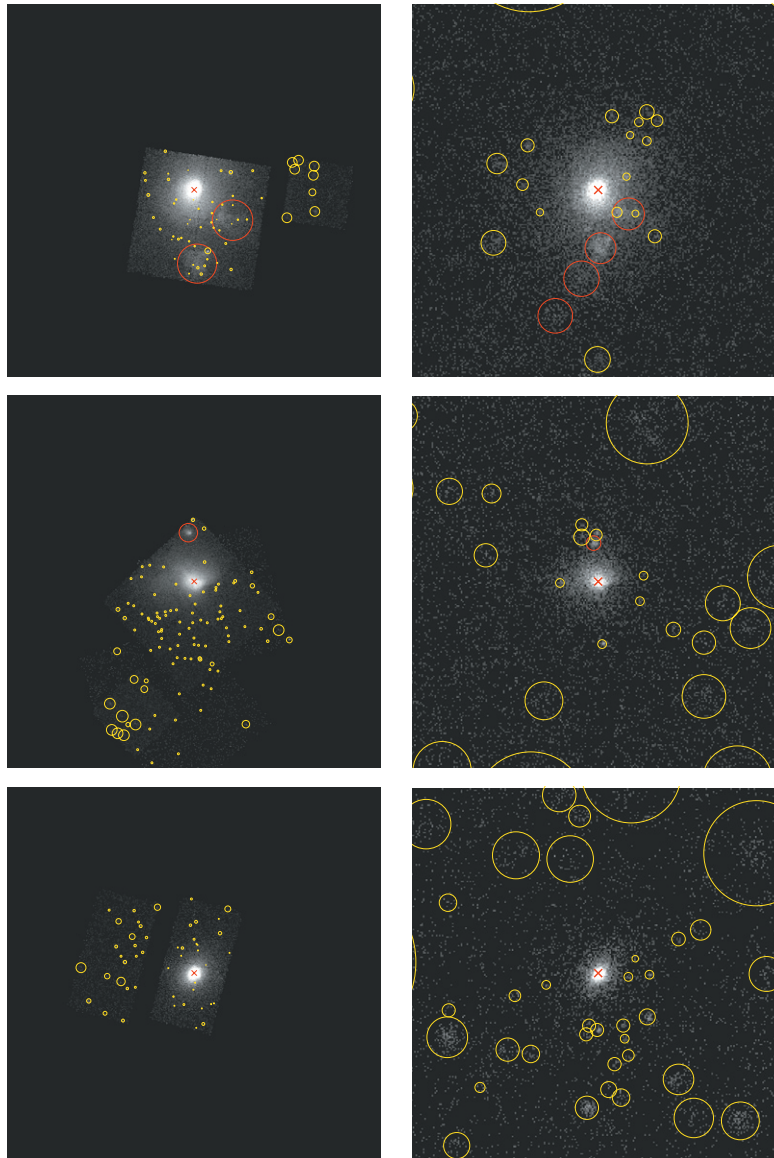

Figure 3. Typical examples of X-ray images for the low-redshift clusters (A85, A2163, and A2597 top to bottom). The left panels show the *Chandra* images (each panel is $50' \times 50'$). *ROSAT* PSPC images ($64' \times 64'$) are shown on the right. Yellow circles show detected sources unrelated to the clusters; the general increase of their radius at large off-cluster distances reflects the degradation of the telescope PSF. The red circles indicate the cluster substructures that were removed from the profile analysis (Section 3.2). The red crosses mark the location of the adopted cluster centroid (Section 3.2).
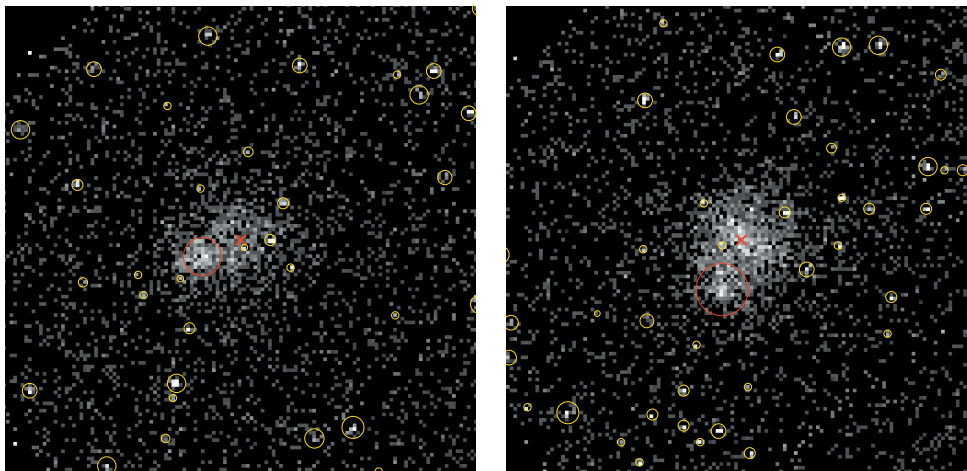
Figure 4. Typical examples of *Chandra* images for the high-redshift clusters (0230+1836, $z = 0.80$ and cl1120+2326, $z = 0.56$). Each panel is $8\rlap{.}'4 \times 8\rlap{.}'4$. The meaning of the region marks is the same as in Figure 3.

The further steps in the X-ray data reduction are based on the analysis of the azimuthally averaged profiles. For this, we need to define the cluster center in each case. In the case of relaxed clusters with cooling flows, the center is defined to be simply at the location of the X-ray peak. The situation is less straightforward for the noncooling flow clusters or those with substructure. Instead of using the maximum in the X-ray brightness map, we center the profiles at the "center of gravity" for the main cluster body. This is done by computing the mean emission-weighted coordinates using the X-ray brightness in the annulus $r = [250-500]$ kpc, and iterating this procedure 2–3 times. The selection of the centroids is illustrated in Figure 3.

3.3. Chandra *Spectral Analysis*

There is an important difference in the approach for determination of the average temperature from the *Chandra* data for high- and low-z clusters. The procedure is straightforward for the high-z objects that fall entirely inside the *Chandra* field of view. In this case, we can measure the average temperature simply by fitting a single-T model to the X-ray spectrum in the 0.6–10 keV band integrated in the radial range of interest, e.g., $r = (0.15-1)\,r_{500}$. This is a common, straightforward analysis; the interested reader can find all the details of our approach in Vikhlinin et al. (2005).

The situation is more complicated for low-z clusters where typically not all position angles fall inside the ACIS field of view at large radii (Figure 3). If we simply fit the integrated spectrum within the ACIS field of view, the contribution of the central region to the total flux will be higher than it should be in the case of complete coverage. This introduces a bias if the ICM temperature distribution is not uniform. Usually, T is overestimated because the observed $T(r)$ decreases at large radii (Markevitch et al. 1998; de Grandi & Molendi 2002; Vikhlinin et al. 2005; Pratt et al. 2007). Our solution is to measure the temperatures independently in several annuli (we use annuli of equal logarithmic width, $r_{\rm out}/r_{\rm in} = 1.5$, within which the overall gradient of $T(r)$ can be neglected) and then average the obtained temperature profile weighting each bin not with observed counts but with the total flux expected in the given annulus if it were completely covered with the field of view. The surface brightness profile needed to compute this weighting function can always be derived from *ROSAT* data that always cover the radial range of interest. In principle, this is not an exact method since this weighting function is proportional essentially to the emission measure integral, while the weighting corresponding to the spectroscopic mean is different (Mazzotta et al. 2004; Vikhlinin 2006). In practice, however, this makes a negligible difference for our clusters, as was verified using the clusters from the Vikhlinin et al. (2005) sample that have the adequate radial coverage for exact computation of $\langle T \rangle$.

The X-ray spectral model we fit to the observed *Chandra* spectra includes foreground absorption in the Galactic ISM. In most cases, the absorbing column density, N_H, was fixed at the value provided by radio surveys (Dickey & Lockman 1990) but we always checked that it is consistent with the observed spectrum. In a few cases (2A 0335, A2634, A478, A2390) the X-ray spectrum indicated a significantly higher absorption than suggested by the radio data, most likely due to the presence of molecular gas and dust along the line of sight. In these cases, N_H was derived directly from the X-ray spectrum. A cautionary note is that small variations of N_H, of order $\pm 2 \times$

10^{20} cm^{-2}, cannot be detected in the *Chandra* spectra[11] because they are indistinguishable from variations of the temperature. For the typical values $N_H = 4 \times 10^{20}$ cm^{-2} and $T = 5$ keV, the variation of N_H by $\pm 2 \times 10^{20}$ cm^{-2} changes the best-fit temperature by $\pm 7\%$, and also changes the derived gas mass by $\pm 3.5\%$, anticorrelated with T. Such variations are smaller than the scatter of these quantities for a fixed mass (Kravtsov et al. 2006) but still should be kept in mind. In this regard, we note that $Y_X = M_{\rm gas} \times T_X$ is less sensitive to variations of N_H because they have the opposite effect on T_X and $M_{\rm gas}$.

The last issue that should be discussed in relation with the X-ray spectral analysis is the treatment of the ICM metallicity. In low-z clusters, the statistical quality is sufficient to measure the metal abundance simultaneously with the temperature. This is impossible for most of our high-z clusters. In these cases, we fixed the metallicity at $Z = 0.3\,Z_\odot$, the typical value at both low and high redshifts (Mushotzky & Loewenstein 1997; Tozzi et al. 2003). We verified that variations of z in the range $0.1-0.5$ (conservative bracket) have a small effect on the derived parameters—for a $T = 5$ keV cluster at $z = 0.5$ the temperature changes by $\pm 5\%$ and $M_{\rm gas}$ changes by $\pm 2\%$, correlated with T.

The instrumental uncertainties in T_X measurement are systematic and uniform (do not introduce object-to-object scatter or any significant redshift-dependent trends). They are considered separately in Section 8.1.1.

3.4. *Gas Mass Measurements*

Two of the mass proxies we utilize for the $M_{\rm tot}$ estimates (Section 4) use the gas mass within $r < r_{500}$. Derivation of the gas mass from the X-ray imaging data is relatively straightforward, but a few points are still worth noting here. Our procedure for the $M_{\rm gas}$ measurements follows that used for a more detailed analysis of a smaller sample of low-redshift clusters described in Vikhlinin et al. (2006, V06 hereafter), and the main steps are outlined here for completeness.

The X-ray flux in the 0.7–2 keV energy band is very insensitive to the plasma temperature, as long as $T \gtrsim 2$ keV (Fabricant et al. 1980). The observed brightness gives essentially the integral of ρ_g^2 along the line of sight. This is why the ICM mass is robustly derived from the X-ray data even if the detector has almost no energy resolution and a limited bandpass. Even though the effects of the temperature and metallicity are very weak, we applied the appropriate corrections to the observed surface brightness profiles as detailed in V06; this correction also removes the effects of spatial variations of the telescope effective area. The corrected profiles are expressed in units of the projected emission measure integral, $\int n_e\, n_p\, dl$. They are deprojected to reconstruct the three-dimensional profile of $\rho_g(r)$. This is done by fitting the projected data to an analytical model,

$$n_p n_e = n_0^2 \frac{(r/r_c)^{-\alpha}}{\left(1 + r^2/r_c^2\right)^{3\beta - \alpha/2}} \frac{1}{(1 + r^\gamma/r_s{}^\gamma)^{\varepsilon/\gamma}}$$

$$+ \frac{n_{02}^2}{\left(1 + r^2/r_{c2}^2\right)^{3\beta_2}}, \qquad (2)$$

which represents all main features observed in real clusters— the β-model (Cavaliere & Fusco-Femiano 1978) profile (Jones & Forman 1984) that may steepen at large radii, and also show

---

[11] Nor in the combined *Chandra* & *ROSAT* spectrum in the 0.2–10 keV band if the nominal N_H is greater than approximately 5×10^{20} cm^{-2}. We note, however, that the *ROSAT* data were checked for consistency with the nominal N_H for all clusters in our low-redshift sample.
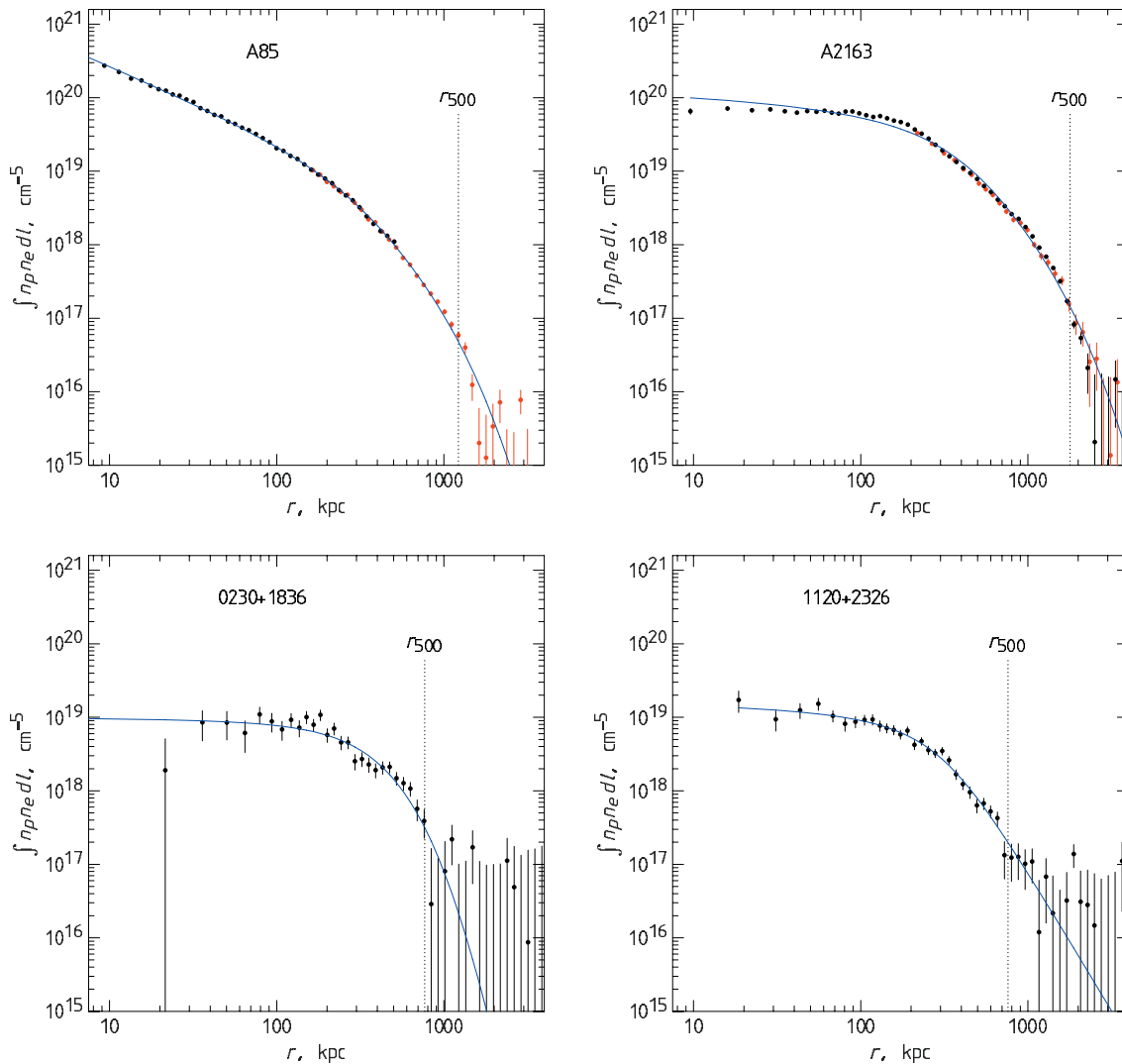
Figure 5. Examples of the surface brightness profile modeling for clusters shown in Figures 3 and 4. The observed X-ray count rates are converted to the projected emission measure integral (see Section 3.4 and V06). The black and red data points show the *Chandra* and *ROSAT* measurements, respectively. The best fit models (the projected emission measure integral for the three-dimensional distribution given by Equation (2)) are shown by solid lines. The dashed lines indicate the estimated r_{500} radii (the Y_X-based value; see Section 4) for reference. Note that in all cases, the surface brightness is traced accurately to r_{500}. In relaxed clusters such as A85, the model describes the data very accurately. In strong mergers such as A2163, we see systematic deviations from the fit. The effect of such deviations on the cluster mass proxies was studied in Nagai et al. (2007b; see also Section 3.5).

a power-law cusp and possibly a separate component in the center. These modifications of the β-model greatly enhance the functional freedom and improve the reliability of the X-ray modeling at large radii (see discussion in V06).

The parameters of the three-dimensional model in Equation (2) are obtained by numerically projecting it along the line of sight and fitting to the observed profile. The best fit directly gives us the analytic expression for the three-dimensional profile of $\rho_g(r)$ which can be integrated to determine M_{gas} in the given range of radii. Several examples of this analysis are shown in Figure 5. Note the excellent agreement between the *Chandra* and *ROSAT* measurements in the same regions indicating an accurate cross-calibration between the two instruments. The uncertainties of $\rho_g(r)$ and M_{gas} are derived via Monte Carlo simulations (see V06).

3.5. Verification by Mock Observations of Cosmological Simulations

We note that our approach to the measurements of the ICM mass and average temperature has been fully tested by the analysis of the mock *Chandra* observations of the clusters from high-resolution cosmological simulations (Nagai et al. 2007b). The cosmological cluster simulations used in this work should correctly reproduce the main aspects of the ICM structure in real clusters, including the large-scale deviations of the main cluster body from spherical symmetry and intermediate-scale nonuniformities of the ICM density and temperature. In fact, the simulations reproduce the detailed X-ray properties of the ICM in the cluster outskirts ($r \gtrsim 0.2r_{500}$) quite well (Nagai et al. 2007a) and are therefore sufficiently realistic for our purposes. In constructing the mock observations of these simulations, we carefully reproduced the essential observational effects such as the *Chandra* sensitivity to plasma of different temperatures, the background level and photon statistics found in typical observations for both low- and high-redshift clusters. The mock data were reduced by the same software that we use for the analysis of real cluster observations.

The mock data analysis thus tests the combined effect of inaccuracies in all steps of our analysis, including removal of substructures, temperature measurements, and modeling of the X-

ray brightness profile. The mock analysis shows that we recover M_{gas} and average temperatures very accurately. For example, the bias in M_{gas} within $r = r_{500}$ due to small-scale nonuniformities of the ICM is only +3%, independent of redshift. The unrelaxed clusters are not significantly different from the relaxed ones, except for a small number of outliers where the M_{gas} measurement can be biased by 10%–15%. This is significantly smaller than the biases reported in the earlier work by Mathiesen et al. (1999). The improvement can be explained by advances in the data analysis (in particular, the relaxation of the assumption that the ICM density follows the β-model) and inclusion in our mock analysis of the effect of substructure removal which was always used by observers.

To summarize, we can state that the results from the analysis of mock observations validate our analysis methods. The expected residual biases have almost no effect on the derivation of the cluster mass function.

4. TOTAL MASS ESTIMATES

The cluster mass is not a well-defined quantity and can be defined in a variety of ways (see, e.g., White 2001). We choose to define mass within the radius corresponding to a fixed mean overdensity, Δ, with respect to the critical density at the cluster redshift, $\rho_c \equiv 3H^2(z)/8\pi G$:

$$M_\Delta = M(< r_\Delta): \quad \frac{M_\Delta}{4/3\,\pi\,r_\Delta^3} = \Delta \times \rho_c. \quad (3)$$

The choice of the overdensity threshold is driven by practical considerations. The ultimate goal of these measurements is to compare the observed mass function with the theoretical predictions. The mass function models, which are calibrated by numerical simulations (e.g., Jenkins et al. 2001), are more robust for low values of Δ, where the role of numerical resolution in the simulations and nongravitational effects within clusters is minimal. On the contrary, the masses derived from X-ray data are more robust for high values of Δ, where the statistical quality is higher, hydrostatic equilibrium assumption is more accurate, etc. We need, therefore, to choose a compromise between conflicting theoretical and observational requirements. We choose $\Delta = 500$, the radius within which the clusters are relatively relaxed (Evrard et al. 1996) and good measurements of gas mass and temperature can be obtained with our *Chandra* observations (Vikhlinin et al. 2006; Nagai et al. 2007b). This is, effectively, the largest radius at which the ICM temperature can be reliably measured with *Chandra* and *XMM-Newton* (e.g., Vikhlinin et al. 2005; Pratt et al. 2007). Using significantly lower Δ dramatically increases observational uncertainties; at significantly higher values of Δ, the theoretical uncertainties start to increase while there is no crucial gain on the observational side.

The total cluster masses, M_{500}, are estimated from observed ICM parameters. We employ the three X-ray proxies for M_{tot} discussed in Kravtsov et al. (2006, KVN hereafter): the core-excised average temperature, T_X, the hot gas mass, M_{gas}, and the estimated total thermal energy, $Y_X = T_X \times M_{gas}$. We rely on the existence of low-scatter scaling relations between these parameters and M_{tot}, as predicted by self-similar theory and confirmed by high-resolution cosmological simulations.

The mass versus proxy relations are calibrated using the hydrostatic M_{tot} estimates in a sample of well observed, low-redshift, relaxed clusters, ten clusters from V06 plus seven additional objects (A2717, A3112, A1835, A1650, A2107, A4059, RXJ 1504–0248) whose deep *Chandra* observations appeared

in the archive since 2006.[12] In principle, the hydrostatic method can underestimate the total mass due to nonthermal pressure components. For example, the analysis of mock observations presented in Nagai et al. (2007b) suggests that M_{500} can be underestimated by $\sim 15\%$, and this effect can be attributed to the bulk motions of the gas at large radii. We do not correct the normalization of the mass versus proxy relation for any such effects because there are theoretical uncertainties in their magnitudes (e.g., the ICM viscosity can affect the average velocity of small-scale bulk motions). We simply account for the possible M_{tot} biases in the total systematic error budget (see Section 8). Ultimately, a reliable calibration of the mass versus proxy relation can be obtained through a stacked weak lensing analysis (e.g., Sheldon et al. 2001) of a representative sample of clusters with high-quality X-ray data. Such data are only starting to become available now (Hoekstra 2007; Mahdavi et al. 2008; Zhang et al. 2008) and we in fact use them to place limits on systematic errors in our calibration of the M_{tot} measurements (Section 4.3.1).

We do apply, however, small first-order corrections to the observed mass versus proxy relations when they are required to transfer the calibration from relaxed clusters to the entire population or to account for expected departures from self-similarity in the evolution of these relations. In doing this, we try to use only the most robust predictions from the simulations and to rely on the directly observed properties as much as possible. The corrections to each proxy are detailed below. The largest corrections are applied for the M_{tot}–T_X relation, while the M_{tot}–Y_X relation does not require any corrections (and hence is potentially the most reliable).

4.1. M_{tot}–T_X Relation

The average X-ray temperature is one of the most widely used cluster mass indicators. The M–T relation expected in self-similar theory is given by

$$M_{500} \propto T^{3/2}\,E(z)^{-1}, \quad \text{where} \quad E(z) \equiv H(z)/H_0. \quad (4)$$

This relation arises in a self-similar model simply because the ICM temperature is expected to scale with the depth of gravitational potential $T \propto M/R$ and mass and radius in our adopted definition are related ($R \propto M^{1/3}$). Relation (4) also generally describes the ICM temperatures found in the cosmological numerical simulations (Evrard et al. 1996; Mathiesen & Evrard 2001; Borgani et al. 2004, KVN).

The average cluster temperature can be defined in different ways but the most practical, from the observational point of view, is the average spectral temperature—the value derived from a single temperature fit to the total cluster spectrum integrated within a given radial range. We refer to this temperature as T_X.

4.1.1. Definition and Determination of T_X

Spatially resolved X-ray spectroscopy became available with the launch of the *ASCA* satellite, and since then many studies has indicated that the cluster scaling relations become tighter if the average temperature is measured excluding the cluster central region which is often affected by radiative cooling. This

---

[12] These data were reduced completely identically to Vikhlinin et al. (2005) and V06. All primary conclusions of these papers hold for these additional objects. The only effect is to improve the accuracy of the M_{tot} versus proxy relations.

is well illustrated by the reduction in scatter in the L_X–T relation shown in Figures 1 and 2 of Markevitch ([1998]). The temperature profiles show a large object-to-object scatter in the centers of even relaxed clusters (V06). Clearly, the central cluster region should be excluded from the measurement of T_X. Markevitch ([1998]) has used $r_{in} = 70$ kpc (this inner cutoff radius was also used in V06 and several other works). Perhaps a better motivated choice is to set r_{in} at a fixed fraction of r_{500} (Arnaud et al. [2005]). We will use, as in KVN, $r_{in} = 0.15\,r_{500}$, because approximately outside this radius the observed profiles of relaxed clusters are self-similar (V06). *Chandra*'s angular resolution is sufficient to resolve $0.15\,r_{500}$ even in the highest-redshift objects. An algorithmic complication is that the cutoff radius is expressed through M_{tot} which is itself estimated from, e.g., T_X. This is not a big problem since T_X is not very sensitive to the exact value of r_{in}, and hence the following iteration scheme converges quickly: (1) measure T_X including the central region; (2) estimate mass from M–T relation; (3) remeasure T_X using $r_{in} = 0.15\,\hat{r}_{500}$ and estimate new mass; (4) repeat step (3) until convergence is reached.

We also need to address the issue of the outer radius for integration of the X-ray spectrum. The cluster properties seem to become progressively self-similar at large radii (Nagai et al. [2007a]). Therefore, ideally, the spectrum should be integrated as far out as possible. The exact value of r_{out} is unimportant because the total X-ray flux converges quickly at $r \rightarrow \infty$. A good practical choice is to set $r_{out} = r_{500}$, because outside approximately this radius, the X-ray brightness is low compared with the background (e.g., Figure 5).

Given the arguments presented above, we determine T_X in the radial range $0.15\,r_{500} - 1\,r_{500}$. This is a straightforward measurement for our high-z clusters because this region fits completely inside the *Chandra* field of view and exposures were designed to provide a sufficient statistical accuracy. However, for a large fraction of the low-z clusters, integration to r_{500} is impossible because of the limited field of view.[13] A simple investigation shows that we can use a smaller value of r_{out} in such cases. First we note that the temperature can be integrated to $0.5\,r_{500}$ for all clusters (we refer to this value as $T_{X,2}$). For clusters that have sufficient radial coverage, we measured temperatures both in $(0.15$–$0.5)\,r_{500}$ and $(0.15$–$1)\,r_{500}$ radial ranges. The ratio of the two values is shown in Figure 6. It is consistent with a linear relation,

$$T_X / T_{X,2} = 0.9075 + 0.00625\,T_{X,2}, \qquad (5)$$

where the temperatures are in units of keV. The observed scatter around the linear fit is negligible, $\lesssim 3\%$. The ratio is also consistent with a constant value, $\simeq 0.95$, except for a few outliers at low temperatures. Since a tight correlation is observed, we can measure $T_{X,2}$ and then estimate T_X with a sufficient accuracy using Equation (5) for those clusters that are not covered by *Chandra* at large radii.

Finally, we note that even if the trend in $T_X / T_{X,2}$ is real, this does not necessarily imply deviations from self-similarity. Because T is not constant as a function of radius, we have a mixture of spectral components within any aperture. A single-temperature fit to such a spectrum gives a weighted average which is different from the mass-weighted T and weighting itself depends on the typical temperature in the spectrum (Mazzotta et al. [2004]; Vikhlinin [2006]). Therefore, we expect trends in the
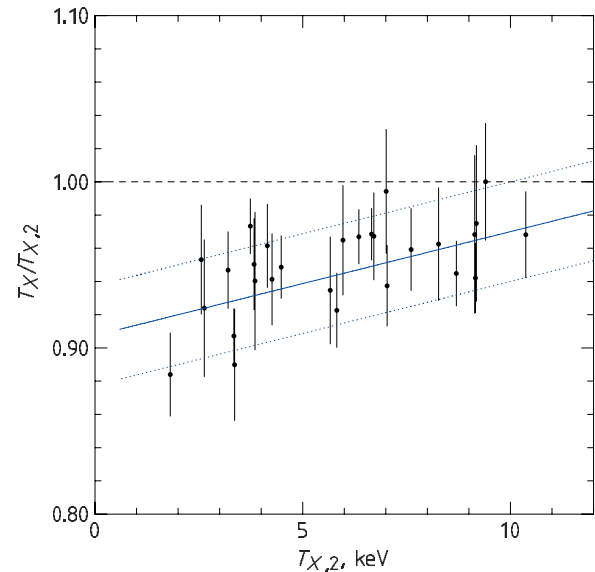
---

[13] Note that the gas mass can still be measured in these clusters out to r_{500} using the *ROSAT* PSPC data.

Figure 6. Ratio of the X-ray spectral temperatures measured in the radial ranges $(0.15 - 0.5)\,r_{500}$ ("$T_{X,2}$") and $(0.15 - 1)\,r_{500}$ ("T_X"), for clusters in the local sample that have a sufficient *Chandra* coverage. The solid line shows the linear approximation given by Equation ([5]), with a 3% level of scatter indicated by dotted lines. The ratio of the temperatures is also consistent with a constant value, $\simeq 0.95$ except for a few outliers at low T.

$T_X / T_{X,2}$ ratio even if the scaled three-dimensional temperature profiles for low- and high-T clusters are identical.

4.1.2. Calibration of M_{tot}–T_X Relation at Low Redshifts using Relaxed Clusters

For 17 low-redshift relaxed clusters, there exist very high-quality *Chandra* observations, providing temperature profiles extending sufficiently far to permit hydrostatic mass estimates at $r = r_{500}$ (see the introduction to Section [4]). These observations are a basis of our calibration of the M_{tot}–T_X relation at low redshifts. The mass and temperature measurements for these 17 clusters (Figure [7]; note that we symmetrize the error bars for simplicity) are fit to the power law,

$$M = M_0 E(z)^{-1} (T/5\,\text{keV})^{\alpha}, \qquad (6)$$

normalized at $T = 5$ keV because this is approximately the median temperature for this sample and therefore the estimates for M_5 and α should be uncorrelated. The fit is performed using the bisector modification of the Akritas & Bershady ([1996], and references therein) linear regression algorithm that allows for intrinsic scatter and nonuniform measurement errors in both variables. The uncertainties were evaluated by bootstrap resampling (e.g., Press et al. [1992]), while simultaneously adding random measurement errors to M and T. The results are shown in Figure [7] and the best-fit parameters of the power law fit are reported in Table [3]. The best-fit slope, 1.53 ± 0.08 is consistent with the expectation of the self-similar theory (Equation [4]). Fixing the power law slope at 1.5 does not significantly reduce the uncertainty in the normalization (Table [3]). The *XMM-Newton* determination of the M–T relation (Arnaud et al. [2005]) is close to our measurement.

Our procedure for hydrostatic M_{tot} estimates was fully tested using mock data from the simulations in Nagai et al. ([2007b]). This work shows that the inaccuracies introduced by the X-ray data analysis—e.g., those related to departures of the cluster body from spherical symmetry—are small. The

Table 3
Calibration of Mass-observable Relations

Relation	Form	$M_0, f_{g,0}$	α
$M_{500} - T_X$	$M_{500} = M_0(T/5\,\text{keV})^\alpha\,E(z)^{-1}$	$(3.02 \pm 0.11) \times 10^{14}\,h^{-1}\,M_\odot$	1.53 ± 0.08
$M_{500} - T_X$	$M_{500} = M_0(T/5\,\text{keV})^\alpha\,E(z)^{-1}$	$(2.95 \pm 0.10) \times 10^{14}\,h^{-1}\,M_\odot$	1.5, fixed
$M_{500} - M_{gas}$	$f_g = f_{g,0} + \alpha \log M_{15}$	$(0.0764 \pm 0.004)\,h^{-1.5}$	0.037 ± 0.006
$M_{500} - Y_X$	$M_{500} = M_0(Y_X/3 \times 10^{14}\,M_\odot\,\text{keV})^\alpha\,E(z)^{-2/5}$	$(5.77 \pm 0.20) \times 10^{14}\,h^{1/2}\,M_\odot$	0.57 ± 0.03
$M_{500} - Y_X$	$M_{500} = M_0(Y_X/3 \times 10^{14}\,M_\odot\,\text{keV})^\alpha\,E(z)^{-2/5}$	$(5.78 \pm 0.30) \times 10^{14}\,h^{1/2}\,M_\odot$	0.6, fixed

Notes. To apply the relations, measure the mass proxy for your h of choice, and scale the normalization factor in Column 2 according to the h-dependence given in Column 3. $M_{tot} - Y_X$ relation should be applied according to Equation (14). The f_g trend is used in the $M_{tot} - M_{gas}$, Equations (8) and (7); the z-dependence of this relation is discussed in Section 4.2.2.
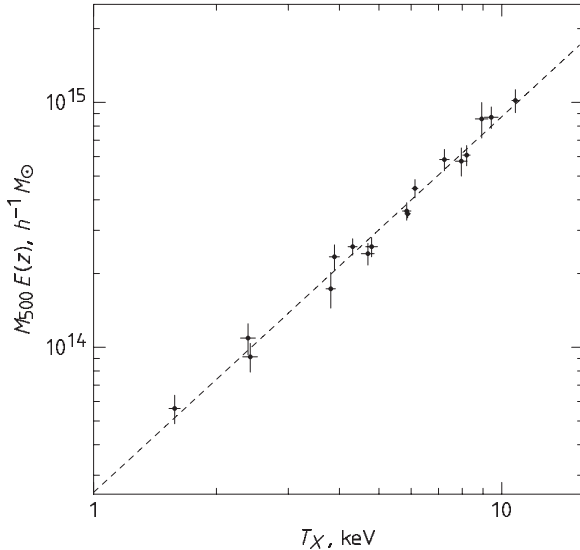
Figure 7. Calibration of the M–T relation using X-ray hydrostatic mass measurements for a sample of ten relaxed *Chandra* clusters with the temperature profile measurements extending to $r = r_{500}$. The mass measurements are taken from V06 with seven additional clusters (see Section 4), the temperatures match our definition of T_X (see Section 4.1.1). The dashed line shows the best-fit power law relation (parameters given in Table 3).
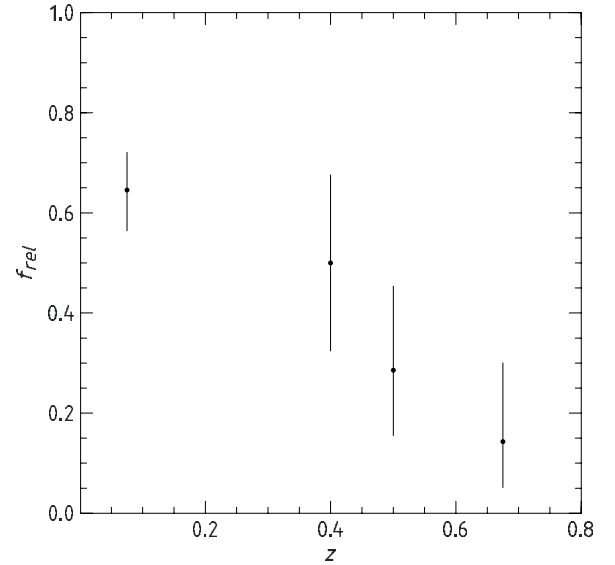

Figure 8. Fraction of clusters which could be classified as "relaxed" based on their observed X-ray morphology (presence of secondary peaks, large centroid shift etc., see Section 4.1.3), as a function of z.

dominant source of error are departures from equilibrium and non-thermal pressure components, the effect fundamentally missed by the X-ray hydrostatic mass estimates. For example, the residual random gas motions in "relaxed" clusters in the Nagai et al. (2007b) sample seem to result in a 10%–20% underestimation of M_{tot} near $r = r_{500}$. Unfortunately, direct measurements of the ICM turbulence (and other nonthermal pressure terms) presently are unavailable. We thus face a dilemma: should we use the theoretical modeling to estimate corrections to the X-ray mass estimates, or should we rely only on observations? Our choice is to follow the philosophy outlined in the introduction to Section 4 and to use the corrections suggested by simulations as an estimate of the systematic errors. A better estimate (9%) for the systematic uncertainties in the *Chandra* cluster mass scale can be obtained from comparison of X-ray and weak lensing mass measurements (see Section 4.3.1).

4.1.3. Transfer of M_{tot}–T_X Calibration to Entire Population

The simulations suggest a systematic offset in the normalization of the M_{tot}–T_X relation for relaxed and unrelaxed clusters, in the sense that the merging clusters tend to have lower temperatures for the same mass (Mathiesen & Evrard 2001; Ventimiglia et al. 2008, KVN). Since our calibration of the M_{tot}–T_X is for a subsample of relaxed clusters, we need a procedure to transfer this calibration for the entire population that contains both

relaxed and merging clusters. This can be achieved using a simple, first-order correction outlined below.

First, we note that the systematic offset in the M_{tot}–T_X relation cannot be measured directly using the X-ray data. Ultimately, it can be measured with a weak lensing analysis of a large sample. The results presented in KVN (their Table 2) suggest that the offset is $(17 \pm 5)\%$ in mass for a fixed T_X. There is no obvious trend of this offset with redshift, or the difference in the slope of the relations for relaxed and merging clusters. Most importantly for our application, this offset can lead to departures from self-similar evolution in the M_{tot}–T_X relation for the entire cluster population, because the fraction of merging clusters is expected to increase at high redshifts (e.g., Gottlöber et al. 2001; Cohn & White 2005), as we indeed observe in our sample (Figure 8).

Second, KVN and Nagai et al. (2007b) classified the simulated clusters as relaxed and unrelaxed using only the morphology of their mock X-ray images. We, therefore, can apply the equivalent classification to the observed clusters in our high- and low-z samples. "Unrelaxed" clusters are those with secondary maxima, filamentary X-ray structures, or significant isophotal centroid shifts. Nagai et al. (2007b) show examples of this classification applied to simulated data; more examples in the real data can be found in Figures 3 and 4. The T_X-based mass estimates for clusters identified as mergers should be corrected upwards by a factor of 1.17. Formal uncertainties on this correction factor are ± 0.05 (Table 2 in KVN); the average uncertainties for the entire population are further reduced because masses have to be corrected only for a fraction of clusters (see

Table 4
Summary of Main Systematic Uncertainties in the Cluster Mass Function

Source of Error	Affects	Uncertainty	ref
400d selection function	$V(M)$ for high-z sample	$\pm 3\%$	Section 5, Paper I
Scatter in $L_X - M_{tot}$	$V(M)$	$\pm 10\%$	Section 5
Evolution in $L_X - M_{tot}$	$V(M)$ for high-z sample	$\pm 22\%$ for median mass	Section 5
Chandra calibration	M_0	-9% to 0	Section 8.1.1
ICM metallicity versus z	$M_Y(z), M_G(z), M_T(z)$	$+3\%, +1\%, +6\%$, respectively	Section 8.1.2
Accuracy of X-ray hydrostatic mass estimates	M_0	$0 - 9\%$	Section 4.3.1
Departures from self-similar evolution	$M_T(z)$	$\pm 7\%$	Section 4.1.4
Departures from self-similar evolution	$M_Y(z)$	$\pm 5\%$	Section 4.3
$f_g(z)$	$M_G(z)$	$\pm 5\%$	Section 4.2.2

Notes. The volume uncertainties are quoted for the median mass in the sample. The M_{tot} uncertainties are separated into calibration of the low-z mass versus proxy relations (M_0) and uncertainties in the evolutions of these relation, $M(z)$. $M_T(z)$, $M_G(z)$, and $M_Y(z)$ stand for total masses estimated using T_X, M_{gas}, and Y_X, respectively. The evolutionary uncertainties are quoted at $z = 0.5$.

below). Applying such a correction for individual objects automatically takes into account any redshift-dependent changes in the fraction of mergers, and thus removes this source of departures from self-similar evolution. Ventimiglia et al. (2008) show that the deviations from the mean $M_{tot}-T_X$ relation in the simulated clusters are correlated with the quantitative substructure measures; using such an approach instead of our simple classification is potentially more accurate and would be warranted in samples of larger size.

We note that the underlying source of difference in the $M_{tot}-T_X$ normalization between mergers and relaxed clusters is incomplete relaxation of the intracluster gas. A fraction of energy is contained in bulk motions of the gas and it is gradually converted into heat as the cluster relaxes after a merger. This process (unlike, e.g., radiative cooling in the center) should be reliably reproduced by current simulations, and so our reliance on the simulations to derive this correction is justified. Our dichotomical classification is, of course, very approximate and a more accurate approach should take into account the cluster relaxation history. However, the current X-ray data does not allow us to quantify the cluster dynamical state with the required precision. Even rather simple substructure measures (e.g., Jeltema et al. 2008) require more photons than we have for distant objects. Moreover, the process of relaxation should be sensitive to effective viscosity in the ICM and it is unclear that the current simulations which incorporate only low, numerical viscosity can accurately predict the $M_{tot}-T_X$ relation for semirelaxed clusters. They, however, should still be reliable for the extreme cases. In nearly relaxed clusters, the turbulent motions are very weak in the inner regions (which dominate the T_X measurements) even in the zero-viscosity simulations. In postmerger clusters, most of the turbulent energy is in the large-scale flows which dissipate on long time scales even if viscosity is high (Coulomb). Furthermore, the magnitude of this correction is relatively small. We estimate that the fraction of nonrelaxed clusters in the sample changes from 35% at $z = 0$ to $\sim 80\%$ at $z = 0.6$ (Figure 8). The corresponding correction for the $M_{tot}-T_X$ normalization for the entire populations is +6% and +13% at $z = 0$ and $z = 0.6$, respectively; thus the redshift-dependent correction is only 7% in mass.

4.1.4. Summary of Mass Estimates through $M_{tot}-T_X$ Relation

This section provides a summary of how we use the X-ray temperature for the cluster M_{tot} estimates.

First, an initial value of T_X is obtained from the *Chandra* spectrum integrated within a wide aperture (not excluding the center). This T_X is used to estimate M_{500} using the power

law fit in Equation (6) and thus r_{500}. The temperature is then remeasured in the annulus $(0.15–1)\, r_{500}$ and this procedure is iterated several times until convergence is reached. If the radius r_{500} is well outside the *Chandra* field of view, we use a smaller aperture, $(0.15 - 0.5)\, r_{500}$, and apply corrections detailed in Section 4.1.1.

Our $M_{tot}-T_X$ relation is calibrated by very high quality *Chandra* observations of 17 low-redshift relaxed clusters with a wide range of masses. The statistical accuracy of this calibration is $\approx 3\%$, so the dominant source of uncertainty is systematics, mostly related to the possible presence of nonthermal pressure components in the ICM.

Next, we need to compensate for the expected systematic difference in the $M_{tot}-T_X$ relation for relaxed and unrelaxed clusters. If the X-ray morphology shows that the cluster is unrelaxed, the mass estimated from the $M_{tot}-T_X$ relation is multiplied by a factor of 1.17 (Section 4.1.3). In doing so, we assume that $M_{tot}-T_X$ relations for relaxed and unrelaxed clusters separately evolve precisely as expected in the self-similar theory (M for fixed T scales as $E(z)^{-1}$). We cannot verify this assumption independently of the background cosmology we would like to measure. Instead, we rely on the simulations to estimate the magnitude of possible departures from the self-similar scaling. Such departures must be treated as systematic errors which affect the cosmological constraints. From the results presented in Nagai et al. (2007b), we estimate this uncertainty to be equivalent to $\approx 7\%$ difference in the normalization of $M_{tot}-T_X$ relations at $z = 0$ and $z = 0.6$.

To properly compute the likelihood function for the estimated cluster mass functions, we need to know the intrinsic scatter in the T_X-based mass estimates. The simulations suggest that this scatter is $\simeq 20\%$, and we adopt this value. We later verified that reasonable variations of the scatter (in the range 15%–25%) have negligible effect on fitting the estimated mass function. This range brackets the scatter observed in the simulations separately for relaxed and unrelaxed subpopulations, as well as for low and high-redshift clusters (Table 2 in KVN). Therefore, our analysis is insensitive to realistic trends of the scatter with redshift.

4.2. $M_{tot}-M_{gas}$ Relation

Our second method of estimating the cluster total mass uses the X-ray derived hot gas mass as a proxy. The application of this proxy is extremely simple in an ideal case in which all cluster baryons are in the ICM, the ICM strictly follows the distribution of dark matter, and clusters contain exactly the cosmic mix of baryonic and nonbaryonic matter (Voevodkin &

Vikhlinin [2004](). The total mass in this case is given simply by

$$M_{\text{tot}} = f_g^{-1} M_{\text{gas}}, \tag{7}$$

where M_{gas} is provided by the X-ray data, and f_g, to the first approximation, equals Ω_b/Ω_M, the ratio which is accurately given by the CMB measurements. To estimate the mass corresponding to a given critical overdensity, we need to solve

$$\frac{M_{\text{gas}}(r) \, f_g^{-1}}{4/3 \, \pi \, r^3 \, \rho_c(z)} = \Delta \tag{8}$$

for r to find the corresponding overdensity radius, r_Δ. Equation ([7]) with M_{gas} evaluated at r_Δ is then used to find M_{tot}.

4.2.1. Corrections for Nonuniversality of Gas Fraction

In reality, the M_{gas}-based estimate is more complicated because the observed gas fraction in clusters is significantly lower than the cosmic average (e.g., Ettori [2003]; Allen et al. [2004]; LaRoque et al. [2006]; Afshordi et al. [2007]) and, moreover, there are trends of observed f_g with the cluster mass (e.g., Mohr et al.; [1999], V06; Zhang et al. [2006]). This trend can be related to the baryon cooling and galaxy formation (Kravtsov et al. [2005]), energy feedback from the central AGNs (Bode et al. [2007]), evaporation of suprathermal protons (Loeb [2007]), etc.—processes whose theoretical modeling is highly uncertain at present. The best approach is therefore to derive the trend $f_g(M)$ observationally. Once this is done, it can be straightforwardly taken into account in Equation ([8]); we just need to use $f_g(4/3 \, \pi \, r^3 \, \rho_c)$ instead of a constant.[14]

The main problem is that direct X-ray hydrostatic M_{tot} measurements near r_{500} are feasible only in a small number of clusters, insufficient to establish the functional form of the $f_g(M)$ trend. We can, however, follow the approach used in Mohr et al. ([1999])—the total mass (and hence, r_{500}) can be estimated from the average temperature (see Section [4.1] above), and then the gas mass determined from the X-ray image within that radius. Such estimates of f_g have substantial uncertainties because of the scatter in the M_{tot}–T_X relation, but this method can be applied virtually to any cluster. The results for our low-z sample are shown by gray points in Figure [9]. The histogram shows the averages of these crude estimates in several mass intervals. Clearly, the data suggest an approximately linear trend of f_g with $\log M$. The f_g values obtained from hydrostatic mass measurements closely follow the same trend (solid black points in Figure [9]). These, more accurate, values are used to determine the normalization and slope of the $f_g(M)$ trend,

$$f_g(h/0.72)^{1.5} = 0.125 + 0.037 \log M_{15}, \tag{9}$$

where M_{15} is the cluster total mass, M_{500}, in units of $10^{15} \, h^{-1} \, M_\odot$. Extrapolation of this trend to lower masses described the observed f_g for galaxy groups (Sun et al. [2008]). The uncertainties of the coefficients are such that f_g is determined to $4\% - 5\%$ across the useful mass range, 10^{14}–$10^{15} \, h^{-1} \, M_\odot$, resulting in the same systematic uncertainty in the M_{tot} estimates

---

[14] We assume that the cluster mass is the only parameter controlling systematic trends in f_g. This assumption is consistent with current observations (see the caption to Figure [9]). If there are additional parameters, their role would be to introduce systematic scatter in the $M_{\text{gas}}/M_{\text{tot}}$ ratio for fixed M_{tot}. The observed scatter is consistent with the value we adapt based on the simulations. If the scatter can be related to easily measured X-ray observables, it would be possible to improve the quality of the M_{gas} proxy still further.
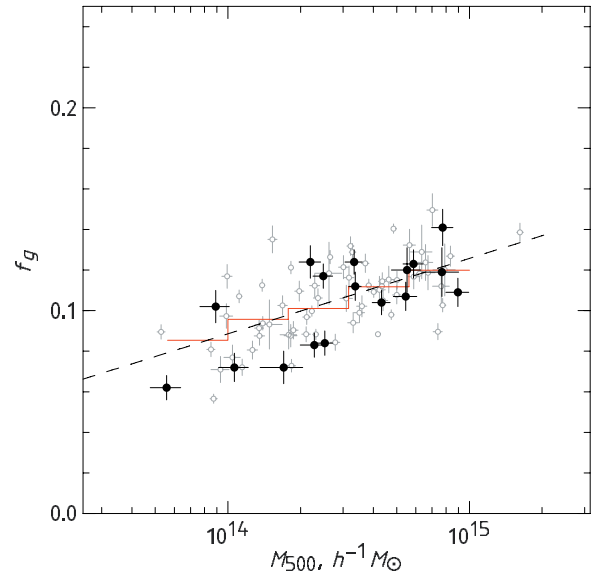

Figure 9. Trend of f_g within $r = r_{500}$ with cluster mass derived from X-ray observations. The solid black circles show the results from direct hydrostatic mass measurements (V06 with seven additional clusters, see Section [4]). Gray circles show approximate estimates using the M_{tot}–T_X correlation (see the text). The scatter is consistent with being purely due to mass measurement uncertainties, either from hydrostatic estimates (Nagai et al. [2007b]) or from M_{tot}–T_X correlation (KVN). The error bars indicate only the formal measurement uncertainties.

because of the $f_g(M)$ trend. The systematic uncertainties are, however, dominated by those of the hydrostatic mass estimates (discussed in Section 4.1.2 and 4.3.1) and so the overall calibration of the absolute mass scale with the M_{gas} method is the same as that in the M_{tot}–T_X or M–Y_X relations.

The observed f_g within for the highest-mass clusters is $\sim 25\%$ lower than the cosmic baryon fraction, $\Omega_b/\Omega_M = 0.165 \pm 0.005$ (Komatsu et al. [2009]). Partly, the remaining baryons can be in the form of stars. The observed star-to-gas ratios for massive clusters are in the range of 0.05–0.1 (Gonzalez et al. [2007]) but the stellar masses are derived from population synthesis models and thus can uncertain by factors of order 2. The tension is reduced still further if the Hubble constant value is lower than we assume. For example, for $h = 0.685$ (Komatsu et al. [2009], the lower 1σ bound for the combined constraints in), the X-ray derived f_g values are 8% higher than we quote in Equation ([9]).

4.2.2. Evolutionary Corrections

Unfortunately, we cannot observationally establish the $f_g(M)$ trend for high-z clusters independent of the underlying cosmology. Therefore, we have to rely on the theoretical models that explain the observed trend at $z = 0$ and can predict its evolution at least for the cosmologies close to the "concordance" ΛCDM. Since, unfortunately, no completely satisfactory model currently exists, this step is a major source of systematic uncertainties.

One such model can be based on the numerical simulations presented in Kravtsov et al. ([2005]). The simulated clusters show the trend in $f_g(M)$ which is very close to that observed at $z = 0$, both in terms of slope and magnitude of the deviation from the global baryon fraction, Ω_b/Ω_m (Figure [10]a). The "missing" baryon component in the Kravtsov et al. clusters is converted into stellar material, so that $(M_{\text{stars}} + M_{\text{gas}})/M_{\text{tot}}$ is within 10% of the value Ω_b/Ω_m specified in the simulation. Observational support for this model is provided by the recent
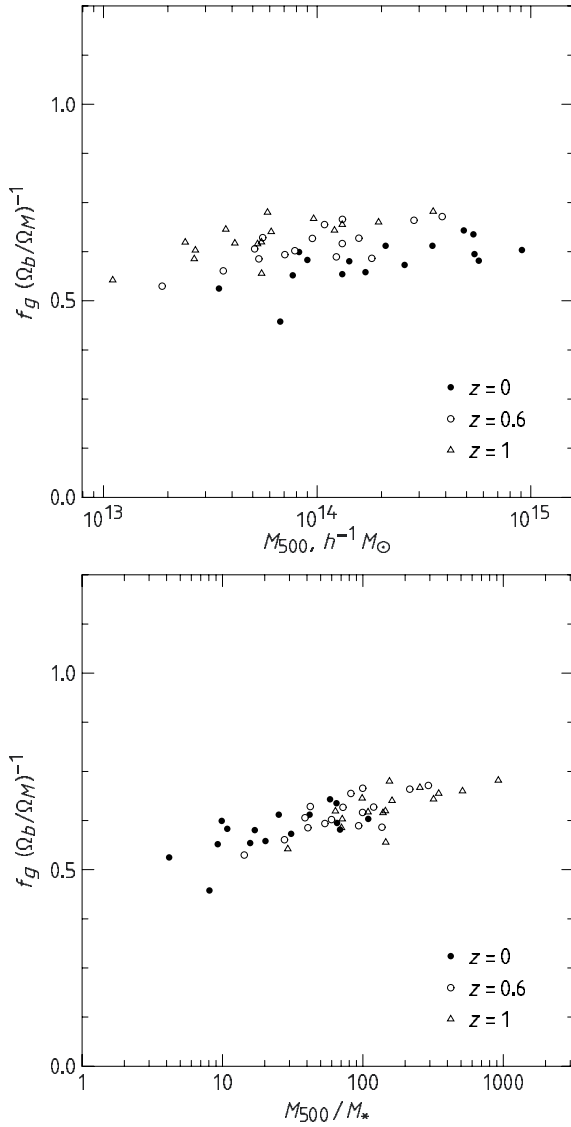
Figure 10. Dependence of f_g within r_{500} on the cluster mass observed in high-resolution cosmological simulations with cooling, star formation, and feedback (Kravtsov et al. 2005; Nagai et al. 2007b). The simulated clusters show qualitatively the same trend as that observed at $z = 0$ (Figure 9), but there is a clear evolution of f_g for the given M. The z-dependence is almost completely removed if we scale the cluster masses by the characteristic nonlinear mass scale, M_* (lower panel).

work of Gonzalez et al. (2007) who show that the trend in stellar mass fraction, $f_{stars}(M)$, is such that it roughly compensates for decreasing f_g in low-mass clusters, making $f_{stars} + f_g$ nearly constant at $M_{500} \gtrsim 10^{14} \, h^{-1} M_\odot$ clusters, although not quite bringing it to the WMAP value of Ω_b/Ω_M (see discussion at the end of Section 4.2.1).

The trend in the Kravtsov et al. simulations show a clear dependence on the redshift (Figure 10(a)) in the sense that a given value of f_g corresponds to a systematically decreasing M_{500}, although at each z, $f_g(M)$ seems to follow lines with the same slope. Empirically, we find that the dependence on the redshift is almost completely removed (Figure 10(b)), if we scale the cluster masses by M_*, the mass scale corresponding to a linear fluctuations amplitude of 1.686: $\sigma(M_*) = 1.686$, where $\sigma(M)$ is the rms fluctuation of density field smoothed with a top hat filter containing mass M. In other words, the simulations indicate that $f_g(M/M_*)$ is almost independent of redshift, at

least at $z \leqslant 1$. A simple explanation of such a scaling can be related to the mass distribution of the cluster progenitors at high redshifts.[15] We also note that qualitatively similar scaling (f_g for a fixed mass increases at high z) is expected if the gas distribution in a cluster potential well does not evolve at all (e.g., inner regions of a cluster remain in equilibrium and do not evolve significantly) and gas fraction is constant, but mass M_{500} changes simply due to evolution of the background critical density, $\rho_c(z)$, to which it is tied by the definition.

In the spirit of our general approach of using the theoretical results in the cluster mass estimates as minimally as possible, we use the observed dependence of f_g on mass for low-redshift clusters (Equation (9)), and take a suggestion from simulation that the same trend should hold at all redshifts, if masses are scaled by M_* computed for the given cosmological model. This gives us $f_g(M, z)$, necessary to estimate the total cluster mass from the observed M_{gas} (Equation (8)). Although this adopted z-dependence is motivated only qualitatively, the predicted overall correction is small. For example, for the cosmological model with $\Omega_M \approx 0.28$, $\sigma_8 \approx 0.78$ (close to the best-fit to our cluster data), $M_* = 3.6 \times 10^{12} \, M_\odot$ and $8.1 \times 10^{11} \, M_\odot$ at $z = 0.05$ and 0.5, respectively. The median masses of clusters in our sample are $4.8 \times 10^{14} \, M_\odot$ at $z = 0.05$ and $2.3 \times 10^{14} \, M_\odot$ (see below). The ratio M_{500}/M_*, therefore, varies from ~ 130 to ~ 280, corresponding to a predicted change in f_g for the median mass clusters of 11% (Equation (9)). A reasonable estimate for the systematic error is around 50% of this overall correction, or 5%–6% in terms of mass between redshifts of 0 and 0.5.

4.2.3. Summary for M_{tot}–M_{gas} Relation

To summarize, our approach to the M_{gas}-based estimates of the total cluster mass is based on using Equation (8) to find r_{500}, and hence M_{500}, for each cluster. In this equation, $M_{gas}(r)$ is the observed gas mass profile derived from the X-ray image, and f_g is the estimated gas fraction as a function of mass and redshift.[16] The dependence $f_g(M)$ is determined empirically at $z \approx 0$ (Section 4.2.1, Equation (9)). It is assumed that this trend evolves with redshift such that f_g remains constant for clusters with a fixed M/M_* (this is justified in Section 4.2.2).

The systematic uncertainties of this M_{tot} estimate are dominated by those of f_g. The latter can be factorized into two components, the uncertainties of the empirical measurements at $z \approx 0$, and the uncertainties of the assumed evolution with redshift. The low-redshift uncertainties are essentially those of the X-ray total mass estimates, discussed above in connection with the M_{tot}–T_X relation. More important for cosmological constraints is the redshift-dependent uncertainty. Within our redshift range, it can be estimated as 5%–6% (Section 4.2.2).

The object-to-object scatter in the M_{gas}-based total mass estimates can be easily derived from the analysis of mock X-ray data for simulated clusters. This was done in Kravtsov et al.

---

[15] Most of the stellar mass in the red and old galaxies of observed clusters, which contain the bulk of the cluster stellar mass, and in cosmological simulations is in place by $z \sim 1$–2, so the stellar fraction within each lower redshift cluster is effectively "preset" at high redshifts. Efficiency of star formation within each galaxy-sized dark matter halo depends on the halo mass. Therefore, the stellar fraction in clusters is probably defined by the mass function of its progenitors near the redshift of the peak star formation ($z = 2$–3). Indeed, the calculations of the progenitor mass functions using extended Press–Schechter theory (Lacey & Cole 1993) show that they are much more similar for clusters with the same M_{500}/M_* than for those with the same M_{500} at different z.

[16] Formally, we can write the gas fraction to be a function of radius, as $f_g = f_g(M/M_*) = f_g(500 \, \rho_c(z) \, 4/3 \, \pi \, r^3/M_*(z))$, and then Equation (8) becomes an implicit equation, which can be solved for r_{500} numerically.

(2006) and Nagai et al. (2007b), who find that the scatter in the $M_{tot}-M_{gas}$ relation is approximately 11% in M_{tot} for a given M_{gas}. Most of this scatter results from the X-ray analysis, as intrinsic scatter of the gas mass for a fixed total mass in simulated clusters is < 5%.

4.3. $M_{tot}-Y_X$ Relation

The final M_{tot} proxy we use is the most robust X-ray mass estimator proposed by KVN. The quantity, Y_X, is defined as

$$Y_X = T_X \times M_{gas,X}, \quad (10)$$

where T_X is the temperature derived from fitting the cluster X-ray spectrum integrated within the projected radii $0.15\,r_{500} - 1\,r_{500}$, and $M_{gas,X}$ is the hot gas mass within the sphere r_{500}, *derived from the X-ray image.*

The quantity that Y_X approximates is the total thermal energy of the ICM within r_{500}, and also the integrated low-frequency Sunyaev–Zeldovich flux (Sunyaev & Zeldovich 1972). The total thermal energy, Y, was found in the simulations to be a very good indicator of the total cluster mass (da Silva et al. 2004; Motl et al. 2005; Hallman et al. 2006; Nagai 2006). In the simplest self-similar model (Kaiser 1986, 1991), Y scales with the cluster mass as

$$M_{tot} \propto Y^{3/5}\,E(z)^{-2/5} \quad (11)$$

(e.g., KVN). This scaling is a consequence of the expected evolution in the $M_{tot} - T$ relation (Equation (4)) and the assumption of the self-similar model that f_g is independent of cluster mass. Hydrodynamic simulations show that the expected scaling Equation (11) is indeed valid, and moreover, the relation shows a smaller scatter in M for fixed Y than, e.g., the $M - T_X$ relation. The primary reason is that the total thermal energy of the ICM is not strongly disturbed by cluster mergers (Poole et al. 2007), unlike T_X or X-ray luminosity (Ricker & Sarazin 2001).

It is reassuring that the $M_{tot} - Y$ scaling also appears to be not very sensitive to the effects of gas cooling, star formation, and energy feedback (Nagai 2006); these effects do not affect the power slope or the evolution law, although change somewhat the overall normalization. The stability of Y is primarily explained by the fact that gas cooling tends to remove from the ICM the lowest entropy gas (Voit & Bryan 2001), increasing the average temperature of the remaining gas and thus affecting T_X and M_{gas} in opposite ways. Direct hydrodynamic simulations of Nagai et al. (2007a) confirm this expectation.

As discussed in KVN, the X-ray proxy, Y_X, is potentially even more stable with respect to cluster mergers than the "true" Y. In the postmerger state, for example, the temperature and thus Y is biased somewhat low because of incomplete dissipation of bulk ICM motions. The same bulk motions, however, cause the gas density fluctuations, which leads to an overestimation of M_{gas} from the X-ray analysis (Mathiesen et al. 1999). Therefore, the merger-induced deviations of the average temperature and derived M_{gas} are anticorrelated and hence partially canceled out in Y_X. Even the strongest mergers in the simulated cluster sample used in KVN do not lead to large deviations of Y_X from the mean scaling. There is also no detectable systematic offset in the normalization of the $M_{tot}-Y_X$ relations for relaxed and unrelaxed clusters. The upper limit for the difference in M_{tot} for fixed Y_X within the KVN simulated sample is 4% (see their Table 2).

Since Y_X is so insensitive to the cluster dynamical state, it is straightforward to calibrate the $M_{tot}-Y_X$ relation using the sample of *Chandra* clusters from V06, and then it is reasonable
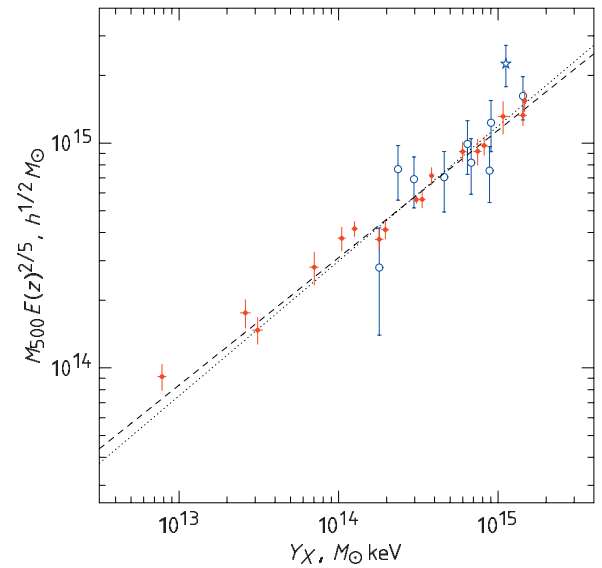

Figure 11. Calibration of the $M_{tot}-Y_X$ relation. Points with errorbars show *Chandra* results from Vikhlinin et al. (2006) with seven additional clusters (Section 4). The dashed line shows a power law fit (excluding the lowest mass cluster) with the free slope. The dotted line shows the fit with the slope fixed at the self-similar value, 3/5 (parameters for both cases are given in Table 3). Open points show weak lensing measurements from Hoekstra (2007; these data are not used in the fit); the strongest outlier is A1689 (open star), a known case of large-scale structures superposed along the line of sight.

to assume that the same relation is also valid for unrelaxed clusters. The observed $M_{tot}-Y_X$ relation does follow very closely the expected self-similar scaling of Equation (11) (Figure 11; see also Arnaud et al. 2007). The best-fit power law is

$$M\,E(z)^{2/5} \propto Y_X^{0.53\pm0.04} \quad (12)$$

when all clusters are included. The marginal deviation of the slope from a self-similar value of 3/5 is driven primarily by the lowest-temperature cluster (MKW4), for which both the total mass and Y_X measurements are most uncertain. Excluding this cluster (its M_{tot} is in any case smaller than the lower mass threshold in the cluster mass functions in our samples), the power law fit becomes

$$M\,E(z)^{2/5} \propto Y_X^{0.57\pm0.05}, \quad (13)$$

fully consistent with the self-similar relation (shown by a dashed line in Figure 11). We use the latter fit for the Y_X-based cluster mass estimates. Note that Sun et al. (2008) find a slope of 0.57 when they fit jointly their galaxy group sample with the V06 clusters, supporting the notion that the MKW4 measurement can be ignored. The normalization constant is provided in Table 3 (it is consistent with the *XMM-Newton* results of Arnaud et al. 2007). Note that the h-dependence of the normalization constant in the $M_{tot}-Y_X$ relation is $\propto h^{1/2}$, different from the usual h^{-1} in, e.g., the $M_{tot}-T_X$ relation. This is the consequence of the h-dependence of the X-ray M_{tot} and M_{gas} estimates; see KVN for details.

The overall uncertainties of the calibration of the $M_{tot}-Y_X$ are identical to those for the $M_{tot}-T_X$ relation (see Section 4.1), with the exception that we do not expect an additional source of uncertainty related to the transfer of calibration from relaxed clusters to the entire population. As for the $M_{tot}-T_X$ relation, we also have to rely on the simulations for an estimate of redshift-dependent departures from the expected self-similar scaling.

The results of KVN provide an upper limit of $< 5\%$ for the evolution of the amplitude of the relation at a fixed Y_X between $z = 0$ and 0.6. The expected level of scatter in the $M_{tot}-Y_X$ relation (7% in M_{tot}, see KVN) is below the uncertainties of *Chandra* hydrostatic mass estimates for individual clusters. Indeed, the intrinsic scatter is undetectable in the data. Since the scatter is expected to be small, its exact value is unimportant for modeling the mass function, and thus can be safely adopted from the simulations.

4.3.1. Systematic Error of Chandra Mass Measurements

Using the $M_{tot}-Y_X$ relation, we can address the question of absolute calibration of the *Chandra* mass estimates through comparison with recent weak lensing mass measurements in representative samples of clusters. Weak lensing measurements of M_{500} in individual objects still have $\sim 30\%$ uncertainties, and are expected to have a similar intrinsic uncertainty due to projection of structures along the line of sight (Metzler et al. 2001). However, as the current weak lensing samples start to include more than ten objects, the average normalization of M_{tot} versus proxy relations can be measured to better than 10%. The two useful recent studies are those of Hoekstra (2007) and Zhang et al. (2008). In Figure 11, we compare the *Chandra* $M_{tot}-Y_X$ relation with that for low-z clusters in the Hoekstra (2007) sample. The Y_X for all Hoekstra clusters in this plot were derived from *Chandra* data using the procedure applied to our cosmological samples. With the exception of a single outlier (A1689, a known case of large-scale structures superposed along the line of sight, e.g., Łokas et al. 2006), the weak lensing masses for given Y_X are in good agreement with the *Chandra* values. Fitting the ratio of normalizations of the $M_{tot}-Y_X$ relations obtained from bisector fits to the two dataset with the slope fixed at 0.57, we find $M^{(wl)}/M^{(Chandra)} = 1.01 \pm 0.11$. A similar agreement is found for the weak lensing masses in Zhang et al. (2008). The normalizations relevant for our case are presented in their Table 3. After correcting their Y_X by +7% to compensate for a systematic difference currently observed between *Chandra* versus *XMM-Newton* temperatures,[17] we find $M^{(wl)}/M^{(Chandra)} = 1.05 \pm 0.07$. The weighted average for the two samples is $M^{(wl)}/M^{(Chandra)} = 1.04 \pm 0.06$. The integrated probability within the $M^{(wl)}/M^{(Chandra)} = [0.91 - 1.09]$ interval is 0.7, thus $\pm 9\%$ is a good estimate for 1σ systematic uncertainties in the *Chandra* cluster mass scale calibration.

4.3.2. Application of the $M_{tot}-Y_X$ Relation for Real Data

In application of the Y_X-based mass estimates to the real data, we face a practical problem that Y_X should be determined within r_{500}, which is itself unknown. Moreover, $Y_X(r)$ diverges at $r \to \infty$, although less quickly than $M_{gas}(r)$. The total mass should thus be estimated with the approach similar to the M_{gas}-based method (Equation (8)); we find r_{500} and hence M_{500} by solving the following implicit equation:

$$C\,(T_X\,M_{gas}(r))^\alpha\,E(z)^{-2/5} = 500 \times 4/3\,\pi\,r^3\rho_c(z), \qquad (14)$$

where C and α are the parameters of the power law approximation to the $M_{tot}-Y_X$ relation, $M_{tot} = C\,Y_X^\alpha\,E(z)^{-2/5}$.

---

[17] Here, we are interested only in bringing all the measurements to the *Chandra* temperature scale because we use *Chandra* data. It may well be that the *XMM* temperatures are in fact correct. The temperature calibration uncertainties should be treated as an additional source of systematic errors (see Section 8.1.1 for more details). Fortunately, the estimated uncertainties are within the bounds suggested by comparison of the *Chandra* and weak lensing mass measurements.

5. SURVEY VOLUMES

We now need to turn to the next critical component of the cluster mass function derivation—determination of the effective survey volume. Our cluster samples are derived from essentially purely X-ray flux limited surveys. We can then straightforwardly compute the sample volumes as a function of X-ray luminosity,

$$V(L) = \int_{z_1}^{z_2} A(f_x, z)\,\frac{dV}{dz}\,dz, \qquad (15)$$

where f is the X-ray flux corresponding to the object with luminosity L_X at redshift z, dV/dz is the cosmological volume–redshift relation, and $A(f_x, z)$ is the effective survey area for such objects. A relation between cluster luminosity and flux,

$$f = \frac{L}{4\pi\,d_L(z)^2}\,K(z), \qquad (16)$$

depends on the cosmological background through the bolometric distance $d_L(z)$ and the K-correction factor (see, e.g., Jones et al. 1998, specifically for the case of the cluster X-ray spectra). The K-correction depends on the assumed cluster temperature but this dependence is *very* weak if both fluxes and luminosities are measured in the soft energy band (0.5–2 keV as we use here). In practice, a sufficient level of accuracy is achieved by estimating T from the nonevolving L_X-T relation accurately measured for low-z clusters (Markevitch 1998; Fukazawa et al. 1998).

Because the objects in our low-redshift sample are all well above the *RASS* detection threshold, their survey area, $A(f, z)$, is adequately approximated by a constant value, 8.14 sr, equal to the geometric area of the sky regions covered (see Section 2.2 and Reiprich & Böhringer 2002). The situation is more complex for our high-z clusters drawn from the 400d survey. Sky coverage there is a function of flux because our distant clusters are generally not much brighter than the detection thresholds in individual *ROSAT* pointings and because the detection thresholds also vary widely depending on the exposure time of each pointing. Formally, the sky coverage is also a function of redshift because detection efficiency is somewhat sensitive to the cluster angular size. A detailed discussion of these effects in application to the 400d survey as well as a careful calibration of $A(f, z)$ for the full 400d sample was presented in Paper I (see their Section 7).

An additional complication arises because we use only a brighter subsample of the 400d sample at $0.35 < z < 0.473$ (see Section 2.1 and Figure 1). We need, therefore, to recompute $A(f, z)$ using Equations (2) and (3) from Paper I with f_{min} in their Equation (2) set to the actual selection fluxes used in our subsample. This is a straightforward calculation but the results cannot be conveniently presented in a paper. We provide machine-readable tables for $A(f, z)$ at the 400d survey Web site.[18]

The stability of the 400d survey area calculations was extensively discussed in Paper I. The general conclusion is that the uncertainties in $A(f, z)$ do not exceed 3%, and therefore they make a negligible contribution to our overall error budget. A dominant source of uncertainty in determining the volume as a function of mass is the details of the L_X-M relation.

5.1. L_X-M_{tot} relations

To fit mass function models to the data, we need to know the survey volume as a function of mass, not luminosity. The two

---

[18] http://hea-www.harvard.edu/400d/CCCP

are trivially related if there is a well defined relation between the cluster mass and its luminosity:

$$\frac{dV(M)}{dz} = \int_L \frac{dV(L)}{dz} P(L|M, z) \, dL, \qquad (17)$$

where $P(L_X|M, z)$ is the probability for a cluster with mass M to have a luminosity L_X at redshift z. The volume in the given redshift interval is obtained by integrating this equation,

$$V(M) = \int_{z_1}^{z_2} dz \int_L A(f_x, z) \frac{dV}{dz} P(L|M, z) \, dL, \qquad (18)$$

where dV/dz is the cosmological volume–redshift relation and $A(f_x, z)$ is the survey area coverage (see Equation (15)).

The simplest model that seems to adequately describe the observed $M_{tot} - L_X$ relations can be represented as a power law with approximately log-normal intrinsic scatter around the mean which is independent of mass and redshift, and the redshift evolution that changes the normalization but keeps constant the slope of the power law,

$$P(\ln L|M) \propto \exp\left(-\frac{(\ln L - \ln L_0)^2}{2\,\sigma^2}\right), \qquad (19)$$

where

$$L_0 = A(z)\, M^\alpha. \qquad (20)$$

The evolution factor is sometimes approximated as a power law of $(1 + z)$ (the simplest model) and sometimes as a power law of $E(z)$ (self-similar evolution inspired models, see e.g., Bryan & Norman 1998):

$$A(z) = A_0 (1 + z)^\gamma \quad \text{or} \quad A(z) = A_0 E(z)^\gamma \qquad (21)$$

A recent study by Maughan (2007, consistent with our results below) indicated that the evolution factor, $E(z)^\gamma$, is in fact close to that expected in the self-similar model for the "concordant" cosmological model. However, the general consensus has been (e.g., Borgani et al. 2001) that we should not rely on the simplest theory for the evolution in the L_X–M_{tot} relation and instead should determine it empirically for each background cosmology. We take this approach in the present study.

5.1.1. Fitting Procedure and Treatment of Malmquist Bias

In the model specified above, the L_X–M relation is characterized by four parameters, A_0, α, γ, and σ. These parameters can be determined using mass estimates for clusters from our sample. The large size of our sample even allows us to test the basic assumptions of the model, e.g., that the scatter is log-normal. A major complication, however, is Malmquist bias. In a flux limited sample, the average luminosity of selected clusters is higher than that in the parent population. The effect becomes strong if the scatter in L_X for fixed M is large (Stanek et al. 2006; Nord et al. 2008), as observed. Nord et al. address the question in which regimes the bias is unimportant. Stanek et al. describe how to estimate the magnitude of the Malmquist bias by simulating mock catalogs. Stanek et al. also discuss how to derive the mean L_X–M relation and scatter from the cluster X-ray luminosity function if the cosmological model (including σ_8) is assumed known. A similar approach was used by Ikebe et al. (2002) who consistently modeled the L–T relation together with the cosmological fit to the cluster temperature function. However, as we show in Appendix A, estimation of

the Malmquist bias can be separated from the cosmological fit to the mass function, which leads to simpler algorithms than those used by Stanek et al. and Ikebe et al. An approach similar to ours was independently developed by Pacaud et al. (2007).

Our algorithm is fully described in Appendix A. Here, we briefly outline the main results and modeling steps and then proceed to presenting the results. A typical situation in low-z surveys with a high flux limit is that the search volume is a power law function of the object luminosity (e.g., $V \propto L^{3/2}$ in Euclidean space and no low-z cutoff), and that the evolution can be neglected within the survey's effective redshift depth. In this case (Appendix A.1), the Malmquist bias leads to a constant offset in the normalization of the observed L_X–M relation; the observed rms scatter in $\ln L$ for fixed M equals the standard deviation of the log-normal distribution (σ in Equation (19)). The true relation is therefore very simple to recover for typical low-z surveys. For the observed scatter, $\sigma = 0.39$ (see below), the bias is $\Delta \ln L \approx 3/2\,\sigma^2 = 0.23$ (Equation (A4)), or L_X for fixed M is overestimated by $\approx 26\%$ (much smaller than the factor of ~ 2 bias advocated by Stanek et al. 2006 but consistent with the limit from Reiprich 2006). This bias is independent of the actual flux limit of the low-z survey. If uncorrected for, it leads to overestimation of the volume for fixed mass by $\approx 40\%$ (because $V(M) \propto L_0^{3/2}$ at low z; see Equations (18) and (19)).

The treatment of the Malmquist bias is more complicated if the evolution in the L_X–M relation cannot be neglected. However, in this case it is still possible to derive a likelihood function which can be computed numerically given the survey selection functions and which implicitly depends on the parameters of the evolving L_X–M relation, A_0, α, γ, and σ (Equations (A10) and (A11)). One can also compute the average bias for each cluster given A_0, α, γ, and σ (Appendix A.2); using these corrections we can easily check if the basic model assumptions (e.g., that the scatter is log-normal and independent of both M and z) are sufficiently accurate. The tests of our fitting procedure using the mock catalogs show that it recovers the true parameters of the L_X–M relation without significant biases.

5.1.2. Results for $L_X - M_Y$ Relation

We independently derive the L_X–M relation for each of our mass proxies. In this section, we summarize the results obtained with the Y_X proxy (hence the relation is called $L_X - M_Y$); the results for the T_X and M_{gas} proxies are very similar. The best fit (obtained with our maximum likelihood method) to the evolving relation in the form Equations (20)–(21) is

$$\begin{aligned}
\ln L_X = {} & (47.392 \pm 0.085) + (1.61 \pm 0.14) \ln M_{500} \\
& + (1.850 \pm 0.42) \ln E(z) - 0.39 \ln(h/0.72) \\
& \pm (0.396 \pm 0.039)
\end{aligned} \qquad (22)$$

where the last term on the right hand side indicates the observed scatter in L_X for fixed M. The uncertainties for each parameter are obtained from the mock catalog simulations described in Appendix A.3. For the median mass in our sample, the best-fit normalization agrees very well with that from (Reiprich & Böhringer 2002, their Table 10, after converting the luminosities to the 0.5–2 keV band), even though we do expect some difference due to corrections for the Malmquist bias applied in our analysis (Appendix A.1). In this regard, we note that our more elaborate procedure for hydrostatic mass estimates should lead to systematically different results than a simple isothermal β-model analysis used in Reiprich &
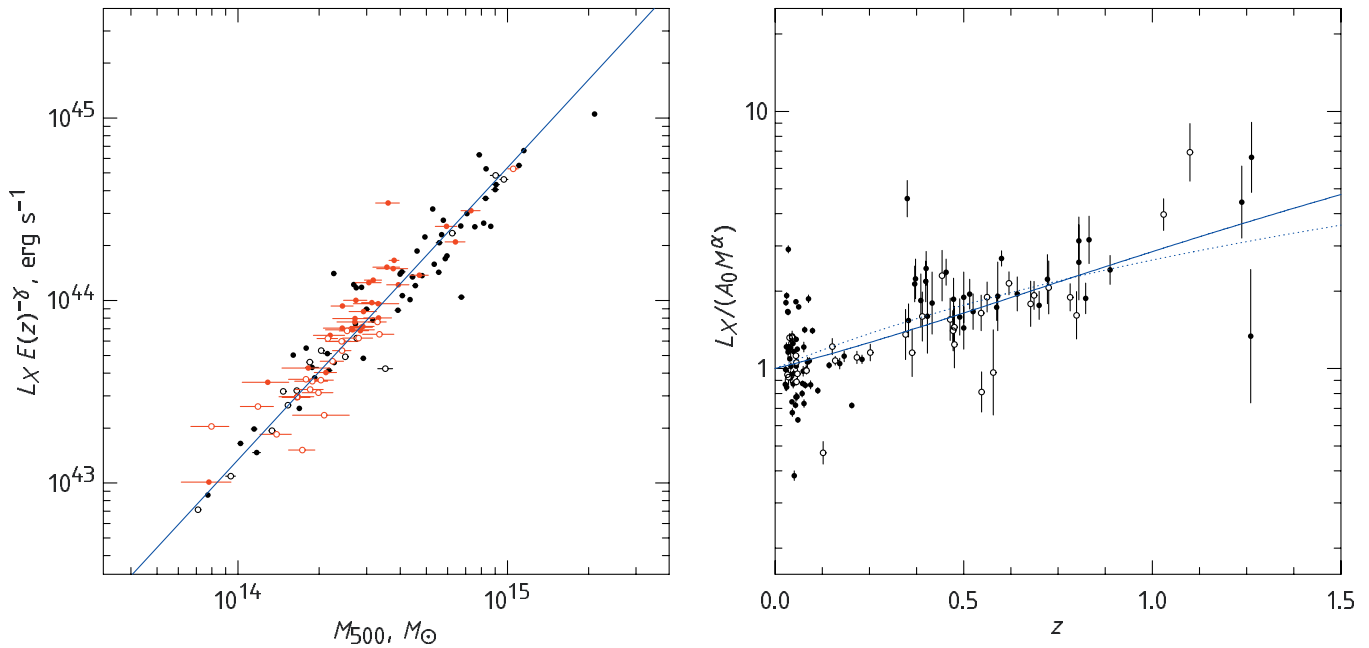
Figure 12. Results for the mass–luminosity relation with M_{tot} estimated from Y_X. Left: the correlation for low-redshift clusters (black points) with the best-fit power law relation. The red points show the data for high-z clusters with the luminosities corrected for the evolution $[E(z)^\gamma]$. All luminosities are corrected for the expected Malmqiust bias (see Appendix A.2). Right: evolution in the normalization of the L_X–M relation. Individual measurements have been corrected for Malmquist bias and divided by the best-fit low-z relation. Solid and dotted lines show the best fit in the form $E(z)^\gamma$ and $(1 + z)^\gamma$, respectively. In both panels, the clusters with large correction ($\Delta \ln L > 0.5$) are shown with open symbols. The lack of a systematic offset between clusters with the estimated strong and weak Malmquist bias proves that the correction has been applied correctly. The $z > 1$ clusters in this panel are from the RDCS survey (Tozzi et al. 2003); they were not used in the fit and are shown only to demonstrate that the extrapolation of our best-fit $E(z)^\gamma$ evolution to higher redshifts still produces reasonable results.

Böhringer (2002); the net effect of updated M_{tot} measurements and corrections for the Malmquist bias appears to lead to very small revisions of the normalization of the L_X–M_{tot} relation.

The left panel in Figure 12 shows that, indeed, the low-z data are adequately described by a single power law relation. The high-z clusters also follow the same relation with approximately the same scatter, after correction for the evolution in the overall normalization $[E(z)^{1.85}]$. The observed evolution in the normalization (the right panel of Figure 12) is consistent with the $E(z)^\gamma$ scaling, but also with a $(1 + z)^\gamma$ law. The exact form of the evolution law is not crucial for our purposes since we use the L_X–M relation only to estimate the survey coverage at each redshift and not to estimate the cluster masses. The effect of the choice of the parametrization on the derived $V(M)$ is discussed below.

The observed deviations from the mean relation at low redshifts (Figure 13) are consistent with the log-normal distribution with a scatter of $\sigma_{\ln L} = 0.396$ (or $\approx \pm 48\%$) in L_X for fixed M. The contribution of the measurement uncertainties to this scatter is negligible for low-z objects. The expected scatter in the M_{tot} estimates using Y_X is also significantly lower. Therefore, it is reasonable to expect that the observed scatter is a good representation of that in the relation between L_X and true mass.[19] The current data quality is insufficient to characterize the shape of the scatter distribution precisely. For example, we cannot check if the tails of the distribution are consistent with the log-normal model. The knowledge of tails in the $P(L_X|M)$ distribution is crucial if one uses L_X as a proxy for cluster mass (Lima & Hu
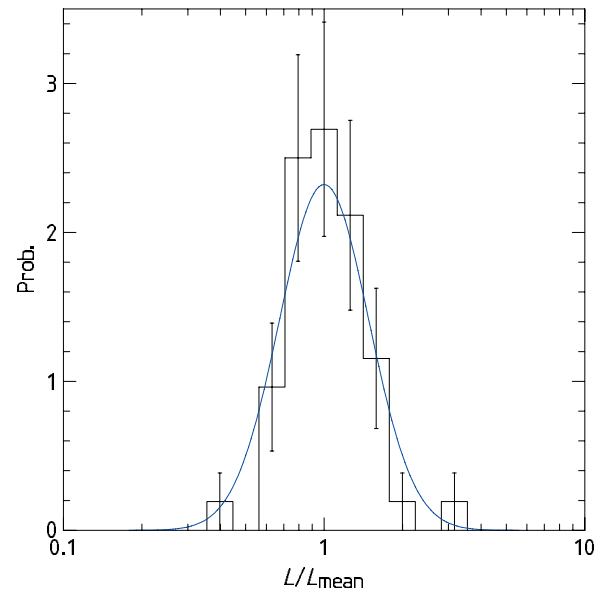
---

[19] Note that we are forced to use the total luminosities, including centers and substructures, for reasons given in Section 3.2. If these components are excluded from the flux measurements, the scatter can easily be made lower, see, e.g., Maughan (2007).
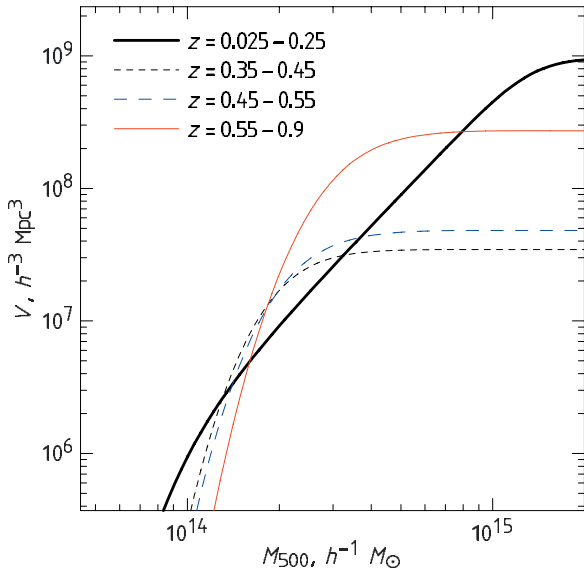

Figure 13. Distribution of the deviations from the mean L_X–M relation for the low-z sample (where the contribution of measurement uncertainties is negligible). The solid line shows the best-fit log-normal distribution with the scatter $\sigma_{\ln L} = 0.396$.

2005). In our case, however, the L_X–M relation is used only for the survey volume calculations, where the effects of the $P(L_X|M)$ are minor (see Section 5.1.3).

The observed 48% scatter in the L_X–M relation implies that Malmquist bias effects are very significant. For example, in a purely flux-limited low-z sample, the average bias in the luminosity for fixed M is $\Delta \ln L = 0.235$ or 26% (see Equation (A4) in Appendix A.1). This is qualitatively similar to the conclusions

Figure 14. Volumes (comoving) as a function of cluster mass for our low redshift sample ($z = 0.025 - 0.25$) and three high-redshift subsamples of the 400d survey.

Figure 15. Sensitivity of the volume calculations to various variations of the L_X-M relation model.

of Stanek et al. (2006), although our predicted bias is lower because Stanek et al. have assumed a larger scatter in the L_X-M relation than that observed in our data.

5.1.3. Results for $V(M)$

With the model for the L_X-M relation at hand, we can now compute the search volumes as a function of cluster mass (Equations (15)–(17)). The results for our local sample and the three redshift bins in the 400d sample are shown in Figure 14. The volume for the local sample follows a power law function of M in a broad range of masses, as expected for a flux-limited sample. A sharp decline of the volume at $M \lesssim 1.5 \times 10^{14} \, M_\odot$ is due to a combination of the flux threshold and a lower redshift cutoff of the sample ($z > 0.025$). The sample becomes volume-limited for high masses because we imposed an upper cutoff ($z < 0.25$) in the volume calculation. For the three high-z subsamples shown in Figure 14, the dynamic range in z is smaller and the transition from the volume-limited to strongly incomplete regimes is much sharper.

We should now discuss how sensitive the survey volume computation is to the assumptions in the L_X-M relation model. The largest uncertainty in the volume computation is related to the measurement errors of the luminosity scale for fixed M. The effect is strongest for the high-z data because the normalization of the L_X-M relation is derived using a smaller number of clusters, with larger measurement uncertainties, and spanning a range of redshifts (see Figure 12). Overall, the uncertainty in the high-z relation corresponds to $\pm 10.5\%$ in the L_X scale at $z = 0.55$ (Appendix A.3). This is equivalent to varying γ by ± 0.33 assuming that the low-z normalization is fixed; note that the range $\Delta \gamma = \pm 0.33$ is smaller than that quoted in Equation (22) because the latter also includes uncertainties in the low-redshift normalization. The long-dashed line in Figure 15 shows how the volume calculation for our high redshift sample, $0.35 < z < 0.9$, is affected by changing γ by +0.33. Reassuringly, the relative change of sample volume is large only for low-mass clusters where it becomes comparable to the Poisson uncertainty of the derived mass function (see Section 8.2.1).
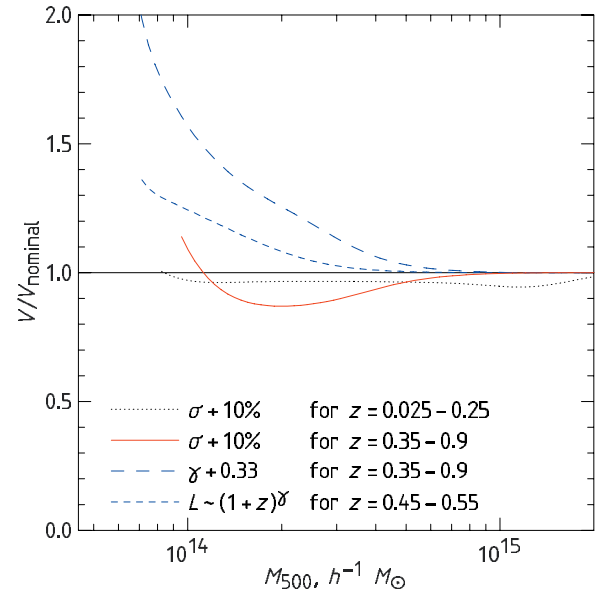
By comparison, the sensitivity of $V(M)$ to a particular choice of the L_X-M evolution model is relatively minor. For example, if we use the $(1+z)^\gamma$ scaling instead of $E(z)^\gamma$ (Equation 21), the largest difference in the best-fit L_X-M relations is near $z = 0.5$ (the right panel in Figure 12). The corresponding relative change of $V(M)$ at $z = 0.45 - 0.55$ (the dashed line in Figure 15) is much smaller than the Poisson uncertainties of the mass function in the same redshift bin (see below). Therefore, the uncertainties related to the parametrization of the L_X-M evolution can be neglected for our purposes.

The effects of the scatter uncertainties on the $V(M)$ computations are comparably small. Note that the situation is crucially different if one uses L_X to estimate the cluster masses. Consider for example, a case of volume-limited survey. The $V(M)$ function is unchanged in this case by variations of σ, while the estimate of the cluster mass function is still very strongly affected (see Lima & Hu 2005). Variations of σ affect the $V(M)$ computations in two ways. First, there is a positive correlation of V and σ because of the scatter term in Equation (17) (assuming that $V(L)$ increases with L). However, σ also implicitly enters the determination of the L_X-M normalization because we need to correct for Malmquist bias; the larger the σ, the lower the L_X for fixed M inferred from the same data (e.g., Appendix A), and hence the smaller $V(M)$. We need to include both these effects to test properly the effect of the σ uncertainties on $V(M)$. This was achieved by fixing the value of σ at the boundaries of its measurement uncertainties ($\pm 10\%$ of the best-fit value, see Equation (22)), refitting all other parameters of the L_X-M relation, and computing $V(M)$ for these new fits. The results are shown in Figure 15 by the solid and dotted lines for the high- and low-z bins, respectively. The variation of volume is negligible for the low-z sample, but is more substantial for the high-z clusters. It is, however, much smaller than the effect of uncertainties in the value of γ considered above. Note that increasing the scatter reduces the volume, indicating that the effect of extra Malmquist bias correction on the L_X-M normalization outweighs the boost in volume due to an increased scattering kernel in Equation (17).
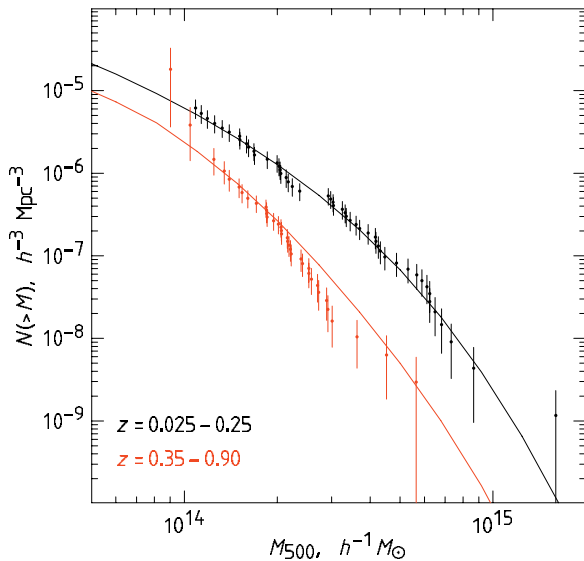
Figure 16. Cluster mass functions for our low- and high-z samples. The masses were estimated by the Y_X method. The errorbars show the Poisson uncertainties. Solid lines show the model predictions for the adapted cosmological model $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.72$, with only σ_8 fit to the cluster data (see the text). The evolution of the mass function is non-negligible within either redshift range. To take this into account, the model number densities for each mass were weighted with $dV(M)/dz$ (Equation (17)) within the redshift bin.

6. CLUSTER MASS FUNCTIONS IN THE CONCORDANT ΛCDM COSMOLOGY

With the survey volume in hand, we can finally compute the mass functions. Figure 16 shows the mass function in the cumulative representation computed as

$$N(> M) = \sum_{M_i > M} V(M_i)^{-1}. \qquad (23)$$

Our samples span similar mass at low and high redshifts, which is very important for the robustness of the derived cosmological constraints. A strong and highly significant decrease in the comoving cluster number density at a fixed mass is observed between $z = 0$ and $z \simeq 0.5$, by a factor of 5.0 ± 1.2 at $M_{500} = 2.5 \times 10^{14}\ h^{-1}\ M_\odot$. This reflects the growth of cosmic structure between these redshifts. Indeed, the observed evolution of the cluster mass function is in good agreement for the "concordance" cosmological model with the power spectrum normalization $\sigma_8 = 0.746$ (solid lines in Figure 16; we use the mass function model from Tinker et al. 2008 and our approach to the model fitting is discussed in Section 7). The strongest observed deviation of the data from the model is a marginal deficit of clusters in the distant sample near $M_{500} = 3 \times 10^{14}\ h^{-1}\ M_\odot$ – we observe four clusters where 9.5 are expected, a 2σ deviation. The cumulative function fully recovers by $M_{500} = 2 \times 10^{14}\ h^{-1}\ M_\odot$, approximately the median mass in the distant sample. The differential representation of the mass function (Figure 17) also shows that this deficit is consistent with the Poisson noise expected in the data.

Our high-z sample can be split into several redshift bins to check if the observed evolution within the sample is still consistent with the model. Figure 18 shows the results for the three bins, $z = 0.35$–0.45, 0.45–0.55, and 0.55–0.9, approximately 14 clusters in each. The data are still in good agreement with the model predictions. The strongest deviation is a marginal ($\simeq 1\sigma$) deficit of clusters at $z = 0.35$–0.45.
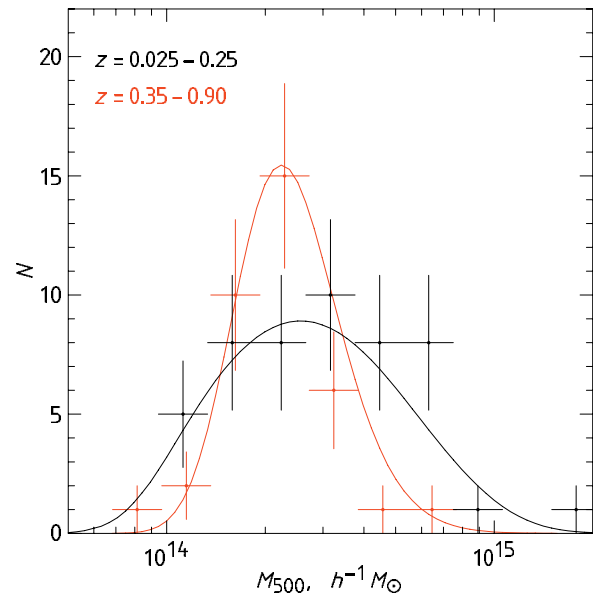

Figure 17. Differential representation of the mass functions shown in Figure 16. The error bars in representation are uncorrelated (unlike Figure 16), so statistical significance of the observed deviations from the best-fit model can be easily assessed.
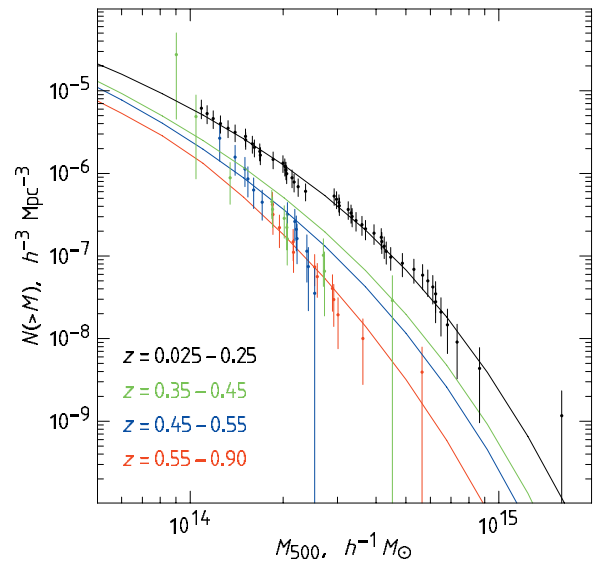

Figure 18. Same as Figure 16 but the high-z sample is split into three redshift bins.

6.1. Sample Variance

In addition to the Poisson cluster counting uncertainties, there is sample variance in the number of clusters in a survey of limited volume due to large-scale clustering. Depending on the mass scale, the sample variance can be comparable to, or larger than, the Poisson errors (Hu & Kravtsov 2003). We follow the formalism of Hu & Kravtsov to assess the importance of sample errors in the error budget in our case.

We calculate the sample variance for the two geometries. For the local sample we assume all-sky coverage with an exclusion zone of $\pm 20°$ from the Galactic plane; the variance for this geometry is given by Equation (A7) of Hu & Kravtsov (2003). The second is a pencil-beam volume with a small circular footprint on the sky, which is appropriate for the individual *ROSAT* fields included in the 400d survey; the variance for this

geometry can be computed using the flat-sky approximation (Equation (7) in Hu & Kravtsov 2003). The variance calculations are done for our reference cosmology with the power spectrum normalization $\sigma_8 = 0.8$, resulting in a slightly higher variance than what would be predicted for our best-fit cosmological model with slightly lower σ_8. The halo mass function model is from Jenkins et al. (2001, their Equation (B3)) and the cluster bias model is from Sheth & Tormen (1999). These mass function models use cluster masses defined within the aperture enclosing an overdensity of 180 with respect to the *mean* density, and so we need to relate it to the mass definition adopted here ($\Delta = 500$ with respect to critical density). We assumed a simple relation, $M_{500} \approx 0.55 M_{180}$, appropriate for typical concentrations of clusters in our mass range.

The relative importance of the sample variance increases for low-mass clusters (Hu & Kravtsov 2003). Fortunately, it is still sufficiently small near our mass limit. For example, at the limiting $M_{500} = 1.5 \times 10^{14} \ h^{-1} \ M_\odot$, the local survey volume corresponds to the effective redshift depth $z_{max} = 0.043$; the sample variance calculated for this mass limit and redshift range $z = 0.025 - 0.043$ is $\sigma_{var} \equiv (\langle n^2 \rangle / \langle n \rangle^2 - 1)^{1/2} \approx 0.16$. This should be compared with the Poisson errors at this mass limit, $\sigma_{shot} = 0.24$ (Figure 16). Combining these variances in quadrature, we find that the total uncertainty $\sigma_{tot} = (\sigma_{var}^2 + \sigma_{shot}^2)^{1/2}$ is only 17% larger than the Poisson value. The contribution of sample variance quickly becomes small for higher masses. For example, at $M_{500} = 3 \times 10^{14} \ h^{-1} \ M_\odot$ (the median mass for the low-z sample), the total uncertainty is only 7% larger than the Poisson error; the contribution becomes negligible at higher masses. Going to lower masses, we predict that σ_{var} becomes comparable to σ_{shot} for $M_{500} \approx 7 \times 10^{13} \ h^{-1} \ M_\odot$, below our mass limit.

The high-z sample consists of 1600 widely separated (and therefore independent) pencil-beam pointings. For a single pointing of circular radius of $17'.5$ and redshift range $z = 0.35 - 0.45$ (the variances for the higher redshift bins are similar but somewhat smaller), the sample variance is $\sigma_{var,1} \approx 0.65 - 1.65$ for the samples with M_{500} thresholds between 10^{14} and $10^{15} \ M_\odot$. Assuming that the individual pointings are uncorrelated (a good assumption for the widely separated pointings of the 400d survey), the total sample variance is $\sigma_{var} \approx \sigma_{var,1} \ N^{-1/2} \approx 0.02 - 0.05$, where $N = 1600$ is the number of 400d survey pointings, much smaller than the Poisson uncertainties. The sample variance can therefore be safely neglected for our high-z sample.

In principle, sample variance can be included in the calculation of the likelihood functions for the low-z sample (Holder 2006; Hu & Cohn 2006). The procedure, however, would be quite cumbersome in our case and is not worth the effort because the variance increases the measurement errors by only 17% in the worst case, and by 7% or less for the median sample mass. This is considerably smaller than expected systematic effects and we will therefore neglect the sample variance hereafter.

7. LIKELIHOOD FUNCTION

Let us now consider the expression for the likelihood function appropriate for our sample and for our method of deriving the mass functions. The basics of the likelihood function are very standard and used in large number of other works. We follow the derivation presented in Cash (1979) for the case of purely Poisson statistics. We split the mass intervals into narrow bins, ΔM, so that the probability to observe a cluster with an estimated mass in this bin is small, $p(M^{est}, z) \Delta M \ll 1$, and we have at

most one cluster per bin. The likelihood function in this case can be written as (see Cash 1979)

$$\ln L = \sum_i \ln \left(p\left(M_i^{est}, z_i \right) \Delta M_i \right)$$
$$- \iint_{M,z} p(M^{est}, z) \, dM^{est} \, dz, \qquad (24)$$

where summation is over the clusters in the sample and integration is over preselected $z_{min} - z_{max}$ and $M_{min} - M_{max}$ intervals. Usually, the ΔM terms can be dropped because they are independent of the model parameters and thus simply add a constant to the likelihood function. In our case, however, the estimated masses are also a function of the background cosmology. When M_i^{est} is changed because of the variation in the cosmological parameters, we should correspondingly stretch the mass interval, $\Delta M = \Delta M^{(0)} M / M^{(0)}$, where $M^{(0)}$ and $\Delta M^{(0)}$ are the estimated mass and width of the interval for some fixed reference cosmological model.[20] Taking the logarithm of this expression and dropping constant terms ($M^{(0)}$ and $\Delta M^{(0)}$), we obtain the likelihood function in the form

$$\ln L = \sum_i \ln p\left(M_i^{est}, z_i \right) + \sum_i \ln M_i^{est}$$
$$- \iint_{M,z} p(M^{est}, z) \, dM^{est} \, dz. \qquad (25)$$

The calculation of individual terms in this expression is discussed in Appendix B. The likelihood function implicitly depends on the cosmological parameters through the model of cluster mass function (reflecting the growth, normalization, and shape of the density perturbation power spectrum), through the cosmological volume–redshift relation which determines the survey volume, and through the distance–redshift and $E(z)$ relations which affect our cluster mass estimates. The best fit parameters are obtained by maximizing the likelihood function in Equation (25). We also can use standard methods (e.g., Cash 1979) to estimate uncertainties of the model parameters. The advantage of this approach is that we do not use any binning in either mass or redshift.

In addition to the best fit parameters and confidence intervals, it is also useful to be able to characterize the goodness of fit. Even though the likelihood function cannot be used for this purpose directly, we can utilize it to obtain an effective χ^2 for every cosmological model. First, we note that essentially all the cosmological information provided by the cluster mass function is the normalization and slope of the linear perturbations power spectrum at ~ 10 Mpc scales. Statistical quality of our sample is sufficient to fit σ_8 independently in 4 redshift bins (the local sample and the high-z subsamples $z = 0.35 - 0.45, 0.45 - 0.55$, and 0.55–0.9) and tilt to the entire sample. Individual best-fit values of σ_8 should be consistent within the errors if the background cosmology is "correct;" similarly, tilt (relative to the best-fit slope constrained by CMB data) should be consistent with zero if we trust the CDM transfer function models (see Eisenstein & Hu 1998, and references therein). To characterize how close the tilt is to 0 and individual σ_8's to a constant value, we can take advantage of the fact that the deviation of the quantity $C = -2 \ln L$ from the minimum has statistical

---

[20] This is equivalent to the rule of transformation of the probability density function under change of variables, $p(y) \, dy = p(x) \, dx$.

properties of the χ^2 distribution (Cash 1979). The effective "tilt" component of the total χ^2 can be computed as

$$\chi_t^2(\boldsymbol{\theta}) = \min_{\sigma_8} C(0, \sigma_8, \boldsymbol{\theta}) - \min_{t,\sigma_8} C(t, \sigma_8, \boldsymbol{\theta}), \qquad (26)$$

where $\boldsymbol{\theta}$ is the vector of cosmological parameters other than tilt and normalization of the power spectrum. Similarly, the effective χ^2 component for the evolution in the normalization of the mass function is

$$\chi_{\rm evol}^2(\boldsymbol{\theta}) = \sum_j C_j(0, \tilde{\sigma}_8, \boldsymbol{\theta}) - \min_{\sigma_8} C_j(0, \sigma_8, \boldsymbol{\theta}), \qquad (27)$$

where summation is over several sufficiently wide redshift bins and $\tilde{\sigma}_8$ is the best-fit value to the entire sample. Adding these terms, we obtain the total effective χ^2 for the cluster mass function data,

$$\chi_{\rm clus}^2(\boldsymbol{\theta}) = \chi_{\rm evol}^2(\boldsymbol{\theta}) + \chi_t^2(\boldsymbol{\theta}). \qquad (28)$$

This effective χ^2 can be used to check how consistent are the model and observed mass functions in terms of general shape and evolution in the normalization.

A detailed discussion of fitting cosmological parameters to our cluster data will be presented in Paper III. Here, we quote only the results of fitting the power spectrum normalization for our reference cosmology, $\sigma_8 = 0.746 \pm 0.009$ (purely statistical uncertainties). The best fit models are shown by solid lines in Figures 16–18.

8. SYSTEMATIC ERROR BUDGET

We conclude the analysis with a summary of the systematic error budget in the mass function measurements (see Table 4). We begin with a discussion of several sources of observational uncertainties (those affecting measurements of the basic cluster parameters), and then summarize the modeling uncertainties—those related to the mass versus proxy relations and determination of the survey volumes.

8.1. Observational Uncertainties

8.1.1. Calibration Uncertainties

The accuracy of the basic X-ray observables—average T and soft X-ray flux—is limited by absolute calibration of the *Chandra* effective area. The calibration parameters most relevant for our study is the absolute value of the soft-band effective area and the relative hard- to soft-band calibration. Thanks to the great effort put into calibration of the *Chandra* telescopes, both on the ground and in flight, the associated uncertainties are small, but they still need to be discussed for the sake of completeness.

The absolute soft-band (~ 0.5–2 keV) effective area affects the measured cluster luminosities and gas masses, $L_X \propto A^{-1}$, $M_{\rm gas} \propto A^{-1/2}$. The largest source of uncertainty in $A_{\rm soft}$ is in-flight contamination of the ACIS optical blocking filters by a hydrocarbon compound. Fortunately, in our energy band of interest, this contamination can be accurately measured as a function of time and position using the onboard calibration source, and so the softband effective area can be brought to its absolute preflight calibration, which is accurate to $\approx 3\%$ (Edgar & Vikhlinin 2004). The effect of such uncertainties on the derived mass function (through the mass proxies and the L_X–M relation) is negligible. The validity of the softband calibration is

indirectly confirmed by the excellent agreement in the *Chandra* and *ROSAT* flux measurements.

The relative hard-to-soft area calibration affects temperatures and hence hydrostatic total mass measurements. We will characterize this effect approximately by the relative change of measured temperatures $\delta T_{\rm cal} = \Delta T/T$ for the 5 keV clusters. The hydrostatic $M_{\rm tot}$ measurements are affected as $\Delta M_{500}/M_{500} = 3/2 \, \delta T_{\rm cal}$ (see, e.g., Appendix A in Vikhlinin et al. 2006). This uncertainty is transferred to our mass function determinations because all mass versus proxy relations are calibrated using hydrostatic $M_{\rm tot}$ measurements.

Calibration uncertainties for the cluster temperatures cannot be characterized exactly. Approximate estimates can be made from comparison of the values derived by different telescopes calibrated independently or by looking at the effect of the most relevant "fudge" factors for *Chandra*. The systematic difference between *XMM-Newton* and *Chandra* temperatures is approximately 7% (V05). The largest remaining *Chandra* calibration uncertainty is, as of this writing, related to the effect of the 10–20 Å hydrocarbon overlayer on the X-ray mirrors. Experimenting with variations of the overlayer model, we find that the range of possible temperature variations is $-6\% < \delta T_{\rm cal} < 0$, and $\delta T_{\rm cal}$ is nearly independent of the cluster temperature and redshift. This would be equivalent to up to -9%, z-independent shift in the mass scale.

8.1.2. Astronomical Uncertainties

In addition to calibration uncertainties, we checked a number of "astronomical" effects which also could affect the measurement of basic cluster properties. The effects that we checked and determined to be negligibly small include uncertainties in the Galactic interstellar absorption measurements (based on neutral hydrogen 21 cm maps Dickey & Lockman 1990); absorption by ionized (warm) ISM in the Galaxy (Reynolds 1993); difference between plasma spectral codes; possible variations of the He abundance around the cosmic average (our conclusion is based on the analysis of Peng & Nagai 2008).

The only effect which is marginally significant is the possible evolution of the ICM metallicity. Since the statistics in the data for our high-z sample are insufficient for ICM metallicity measurements, we assumed in each object that the metallicity is equal to 0.3 Solar, approximately the mean value for the low-z population. If in fact there is an evolution in the heavy element metal abundance, our derived values for T_X and $M_{\rm gas}$ are affected slightly. For example, if the mean abundance for high-z clusters is 0.15 Solar, the derived T_X and $M_{\rm gas}$ will be higher by $\approx +4\%$, and $+1\%$, respectively. Such a trend (which is probably outside the range allowed by the data, see Tozzi et al. 2003; Maughan et al. 2008), will be equivalent to changing the mass scale for our high-z sample by $+3\%$, $+1\%$, and $+6\%$ if masses are estimated through the $M_{\rm tot}$–Y_X, $M_{\rm tot}$–$M_{\rm gas}$, and $M_{\rm tot}$–T_X relations, respectively.

8.2. Modeling Uncertainties

Modeling uncertainties in the mass function measurements can be separated into two components: (1) how accurately we can predict the survey volume for clusters of a given mass, and (2) how accurately we can derive cluster masses from the data. The first component mainly depends on the accuracy of the L_X–$M_{\rm tot}$ relation, and the second, on the $M_{\rm tot}$ versus proxy relation. All these uncertainties were discussed in detail above and so we provide only a summary here.

8.2.1. Uncertainties in $V(M)$

Uncertainties in the survey volume mainly depend on how accurately we can recover the L_X–M_{tot} relation from the data, assuming that masses are accurately reconstructed from the Y_X, M_{gas}, or T_X proxies. The effects of these uncertainties on $V(M)$ are considered in Section 5. The largest error is related to the measurement of the evolutionary factor (Equation (21)) and amounts to $\pm22\%$ in volume for the median mass in our high-z sample, $\sim 2.1 \times 10^{14}~h^{-1}~M_\odot$ (Figure 15). This source of error is statistical in nature (related to measurement uncertainties in the L_X–M_{tot} parameters). It can therefore be added in quadrature to the purely Poisson errors ($\pm26\%$ in the cumulative mass function for the same mass threshold), resulting in a moderate increase in the statistical errorbars. Although it is possible to include these uncertainties approximately in the cluster likelihood function or effective χ^2, a more accurate estimate of their effect on the final results can be obtained by repeating the entire analysis procedure with the evolution factor varied within its measurement errors. We take this approach and we will quote the associated parameter uncertainties in Paper III.

Other sources of uncertainty from the L_X–M_{tot} modeling, such as the exact scatter in the relation, functional form of the evolution term, etc., are comparably small. The accuracy of statistical calibration of the 400d survey selection function also makes a negligible contribution, $\pm3\%$, to the volume error (Paper I).

8.2.2. Uncertainties in derived M_{tot}

Separate sources of uncertainties are related to potential biases of the M_{tot} estimates from X-ray proxies. Note that these biases will have little effect on the volume computations for the given cluster (if we change the estimated M_{tot}s, we need to refit the L_X–M_{tot} relation and the net effect will be that the volume for the given cluster is almost unchanged). In a sense, the $V(M)$ systematics move the cumulative mass functions in Figure 16 up and down, while the potential M_{tot} biases shift the mass function along the M axis.

The mass biases can be naturally separated into two components. The first is related to calibration of the M_{tot} versus proxy relations for low-z clusters. Assuming that the evolution in the relation is nominal, such biases will shift the low and high-z mass function by the same amount, or, equivalently, will affect the overall normalization, but not the evolution in the comoving number density. As we discussed above (Section 4.3.1), comparison of X-ray and weak lensing masses provides a good estimate, $\pm9\%$ in mass, for such biases.

The second source is departures of the evolution in the M_{tot} versus proxy relation from the assumed forms. Since evolution is negligible within the low-z sample, such biases are important only for the high-z mass function and thus will affect the derived evolution in the cluster number density, but not the overall normalization of the mass functions. We estimate that by $z = 0.5$, the evolutionary M_{tot} biases can be up to $\pm7\%$ in the M_{tot}–T_X relation, and $\pm5\%$ for the M_{tot}–M_{gas} and M_{tot}–Y_X relations (Sections 4.1.3, 4.2.2, and 4.3, respectively). In the case of the M_{tot}–T_X we also need to add a $\pm6\%$ uncertainty related to the potential evolution in the ICM metallicity (Section 8.1.2).

The estimated uncertainties in the M_{tot} calibration cannot be easily included in the likelihood function. Instead, we check (Paper III) how they affect the cosmological fit by repeating the entire analysis procedure with the parameters M_{tot} versus proxy relations within the bounds specified above. Note also that the use of three different mass proxies, each with its own bias, provides a good consistency check, because results obtained with different proxies can be compared to each other to check for biases.

9. SUMMARY

We presented a report on data analysis procedures leading to a measurement of the galaxy cluster mass functions using *Chandra* observations of statistically complete samples of low and high-z clusters originally selected in the X-ray data from *ROSAT*. This measurement relies on a careful selection of the parent samples, rather detailed *Chandra* observations of selected objects, and using several robust X-ray proxies for the total cluster mass (Y_X, M_{gas}, T_X).

The scaling relations between proxies we use and M_{tot} mostly follow the predictions of the self-similar theory, a very basic and hence reliable model. We used advanced high-resolution numerical simulations to test the predictions of this theory with regard to our proxies; these simulations indicate that only small corrections are necessary, which we use cautiously. At low redshifts, the M_{tot} versus proxy relations were calibrated by detailed *Chandra* observations of a sample of relaxed clusters spanning a wide range of mass; our *Chandra* results were cross-checked against recent weak lensing measurements.

As a part of this project, we derive a relation between cluster mass and total X-ray luminosity, using large statistically complete samples and properly taking into account the Malmquist bias. The relation is adequately described by a single power law, substantial log-normal scatter, and evolution of the power law normalization following $E(z)^{1.85}$, assuming that the evolution in the M_{tot}–Y_X relation is exactly self-similar as we use Y_X to estimate cluster masses.

We present the cluster mass functions estimated assuming a "concordant" ΛCDM cosmology. These data show a significant evolution in the cluster comoving number density at a fixed mass threshold, by a factor of ≈ 5 at $M_{500} = 2.5 \times 10^{14}~h^{-1}~M_\odot$ between $z = 0$ and 0.5.

Finally, we provide a summary of estimated systematic uncertainties in our mass function measurement. Most source of systematics lead to corrections which are smaller than the Poisson errors in our data. The main exception is uncertainties in calibration of the absolute mass scale at low redshifts but it has little impact on the measurement of evolution in the cluster number density.

The evolution in the cluster mass function reflects the growth of density perturbations and can be used for the cosmological constraints complementing those from the distance–redshift relation. The cosmological modeling of these data will be discussed in a future paper.

APPENDIX A

CORRECTION OF THE LUMINOSITY–MASS RELATION FOR THE MALMQUIST BIAS

Stanek et al. (2006) recently pointed out that observational determinations of the mass–luminosity relation can be significantly affected by Malmquist bias because of the flux-limited nature of most of the available cluster samples. Stanek et al. were interested in the relation where L_X is the independent variable; such a relation is useful for M_{tot} estimates using L_X as a proxy. Corrections for Malmquist bias in this case lead to complicated computations involving the cluster mass function model. We, instead, are interested in computing the survey volume for objects of given mass, for which we need to treat M_{tot} as independent variable and compute luminosity for a given mass (Equation (18)). The calculations of the Malmquist bias are then much simpler and can be done independently of the mass function modeling.

We assume that the scatter in L for fixed M has a log-normal distribution,

$$p(\ln L) = \frac{dN}{d \ln L} \propto \exp\left(-\frac{(\ln L - \ln L_0)^2}{2 \sigma^2}\right), \qquad (A1)$$

where L_0 is the average luminosity for the given mass. Typically, L_0 is a power law of mass, $L_0 \propto M^\alpha$, but we do not make this assumption in the calculations below.

A.1. Corrections for Malmquist Bias in Nonevolving L_X–M Relation

Calculations of the Malmquist bias are particularly simple if the evolution in the L_X–M can be neglected, e.g., $L_0(M)$ is the same at all redshifts within the sample. This situation is applicable for the analysis of the low-z samples and in our case, for establishing the low-z reference relations.

Let us assume that the survey volume as a function of the object luminosity can be approximated as a power law

$$V(L) \propto L^\delta. \qquad (A2)$$

For example, in the case of Euclidean space and a pure flux-limited survey, $V(L) \propto L^{3/2}$ exactly. If there is a lower redshift cutoff in the survey, $\delta \neq 3/2$ at the low-L end, even in Euclidean space. Likewise, if there is a higher-redshift cutoff, $\delta \to 0$ in the high-L end (sample becomes volume-limited).

The L_X–M relation is usually fit in the $\ln M - \ln L$ coordinates, so we need to compute the bias in $\ln L$ for given M:

$$\langle \ln L \rangle - \ln L_0 = \frac{\int_{-\infty}^{\infty} (\ln L - \ln L_0)\, p(\ln L)\, V(\ln L)\, d \ln L}{\int_{-\infty}^{\infty} p(\ln L)\, V(\ln L)\, d \ln L}$$

$$= \frac{\int_{-\infty}^{\infty} x\, \exp(-x^2/2\sigma^2)\, \exp(x\delta)\, dx}{\int_{-\infty}^{\infty} \exp(-x^2/2\sigma^2)\, \exp(x\delta)\, dx}, \qquad (A3)$$

where we used the substitution $x = \ln L - \ln L_0$. The integrals can be worked out analytically,

$$\langle \ln L \rangle - \ln L_0 = \frac{\delta \sqrt{2\pi}\, \sigma^3\, \exp(\delta^2 \sigma^2/2)}{\sqrt{2\pi}\, \sigma\, \exp(\delta^2 \sigma^2/2)} = \delta \sigma^2. \qquad (A4)$$

The log-normal scatter in the relation, σ, is usually unknown a priori and thus it should be estimated from the rms scatter around the best-fit relation in the $\ln M - \ln L$ plane. Fortunately, the flux-limited survey does not introduce bias in the scatter, i.e., $\sigma_{obs} = \sigma$, as we now demonstrate.

$$\sigma_{obs}^2 = \langle (\ln L - \ln L_0)^2 \rangle - (\ln L_0 - \langle \ln L \rangle)^2$$
$$= \langle (\ln L - \ln L_0)^2 \rangle - \delta^2 \sigma^4 \qquad (A5)$$

$$\langle (\ln L - \ln L_0)^2 \rangle = \frac{\int_{-\infty}^{\infty} x^2\, \exp(-x^2/2\sigma^2)\, \exp(x\delta)\, dx}{\int_{-\infty}^{\infty} \exp(-x^2/2\sigma^2)\, \exp(x\delta)\, dx}$$
$$= \sigma^2(1 + \delta^2 \sigma^2) \qquad (A6)$$

(cf. Equations (A3) and (A4)), and so $\sigma_{obs}^2 = \sigma^2$.

A.2. Correction for Individual Clusters

The bias computations from the previous section cannot be applied if we aim to model also the evolution in the L_X–M relation, because in this case we need to compute the bias in a fixed narrow interval of z where $V(L)$ cannot in general be represented with a power law (e.g., for an ideal flux-limited survey, $V(L)$ in a narrow interval of z is close to a step-function). An alternative (approximate) approach is to compute the expected biases in L for individual clusters, as considered below. A better approach is to model all effects of selection through the likelihood function, as discussed in the next section.

Let us assume that the survey has a single flux threshold, f_{min} (i.e., the cluster is always detected if $f > f_{min}$ and not detected if $f < f_{min}$). The average luminosity bias of *detected* clusters with a given mass is

$$\langle \ln L - \ln L_0 \rangle = \langle \ln f - \ln f_0 \rangle = \frac{\int_{x_{min}}^{\infty} x\, \exp(-x^2/2\sigma^2)\, dx}{\int_{x_{min}}^{\infty} \exp(-x^2/2\sigma^2)\, dx}$$

$$= \frac{\exp\left(-x_{min}^2/2\sigma^2\right)}{\sqrt{\pi/2}\, \mathrm{erfc}(x_{min}/(\sigma\sqrt{2}))}\, \sigma, \qquad (A7)$$

where f_0 is the flux corresponding to the nominal luminosity, L_0, given by the L_X–M relation, and $x_{min} = \ln f_{min} - \ln f_0$. For $f_0 \gg f_{min}$ (very massive clusters), the bias is 0 as expected. For very low mass clusters ($L_0 \to 0$, $f_0 \to 0$, $x_{min} \to \infty$), Equation (A7) gives $\langle \ln f - \ln f_0 \rangle \simeq \ln f_{min} - \ln f_0$ (i.e., all detected clusters have fluxes just above the survey threshold).

Equation (A7) is easily generalized for the case when the survey selection probability is a smooth function of flux (as is the case for the 400d sample):

$$\langle \ln L - \ln L_0 \rangle = \langle \ln f - \ln f_0 \rangle$$
$$= \frac{\int_{-\infty}^{\infty} x\, P_{sel}(x + \ln f_0)\, \exp(-x^2/2\sigma^2)\, dx}{\int_{-\infty}^{\infty} P_{sel}(x + \ln f_0)\, \exp(-x^2/2\sigma^2)\, dx}. \qquad (A8)$$

A.3. Likelihood Function and Fitting Procedure

The best way to treat the Malmquist bias in modeling the relation is through a proper definition of the likelihood function. Let $P_{sel}(\ln f)$ be the survey selection efficiency as a function of flux. The average luminosity–mass relation gives a "nominal" luminosity for clusters of given M, which corresponds to a "nominal" flux f_0. The probability density function for the cluster to have flux f is

$$\frac{dP}{d \ln f} = C\, \exp\left(-\frac{(\ln f - \ln f_0)^2}{2\sigma^2}\right) P_{sel}(\ln f), \qquad (A9)$$
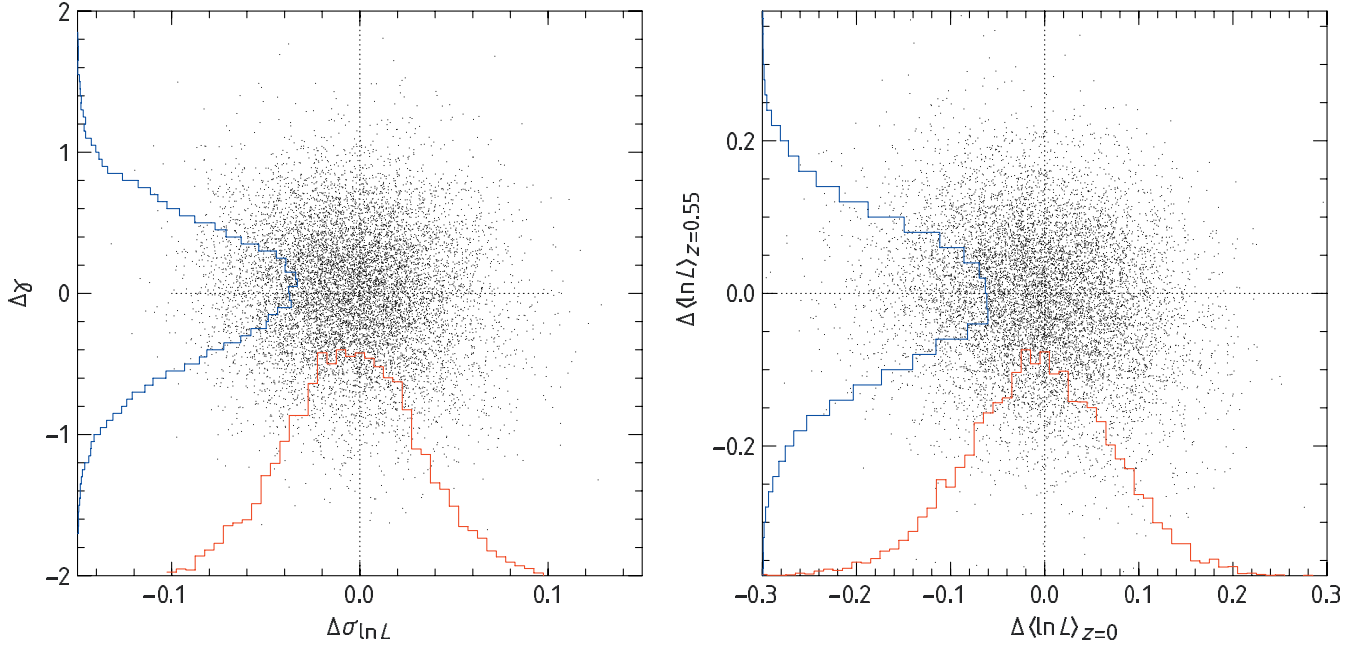
Figure 19. Distribution of the L_X−M parameters recovered from mock catalogs (see the text for details). The points show the deviations of the best-fit parameters in each realization from the input values, and the histograms show the probability density distribution for each parameter. The right panel shows the deviations of best-fit luminosities at $z = 0$ and 0.55 for the median mass in the samples (the $z = 0$ results are equivalent to those for the overall normalization, A_0) for fits with σ fixed at the nominal value (see the text).

where C is the normalization coefficient defined so that the total probability is 1,

$$C^{-1} = \int_{\infty}^{\infty} \exp\left(-\frac{(\ln f - \ln f_0)^2}{2\sigma^2}\right) P_{\text{sel}}(\ln f)\, d\ln f. \quad (A10)$$

For a survey with a single sharp flux limit [i.e., those with $P_{\text{sel}}(f) = \theta(f - f_{\min})$], Equation (A10) becomes $C = (\frac{1}{2}\,\text{erfc}[\ln(f_{\min}/f_0)/(\sigma\sqrt{2})])^{-1}$.

The total likelihood function, $\mathcal{L}$, is the product of $dP/d\ln f$ for individual clusters. The quantity $-2\ln\mathcal{L}$ can be used in place of the usual χ^2 for finding the best fit and confidence regions (Cash 1979). From Equation (A9), we have

$$-2\ln\mathcal{L} = \sum_i \frac{(\ln f_i - \ln f_{0,i})^2}{\sigma^2} - 2\ln C_i - 2\ln P_{\text{sel}}(\ln f_i). \quad (A11)$$

The first term on the right-hand side of Equation (A11) is the usual unweighted χ^2 and the extra two terms are corrections for the Malmquist bias. Masses of individual clusters and parameters of the L_X−M relation enter the likelihood function implicitly, through calculating the "nominal" luminosities (e.g., $\ln L_0 = A\ln\hat{M} + B + \text{evol}(z)$) which are then converted to f_0's.

Parameters of the L_X−M relation can be obtained from finding the global maximum of the likelihood function in Equation (A11). In practice, we use a multi-step procedure to fit the L_X−M parameters. The scatter, overall normalization, and power law slope are determined from the low-z data where the measurement uncertainties can be neglected relative to intrinsic scatter. The best-fit scatter is corrected by a factor of $(N/(N-1))^{1/2}$, the expected bias of the maximum likelihood estimate, where N is the number of clusters in the low redshift sample. Then, with A_0, α, and σ fixed, the evolutionary term (γ) is determined from the fit to the high-z data. The procedure is iterated several times until convergence.

To assess how well our fitting procedure recovers the parameters of the L_X−M relation, we applied it to mock cluster samples. The mock samples were designed to mimic closely our actual low- and high-z samples. The cluster masses and redshifts were drawn from the mass function model computed in the $\Omega_M = 0.28$, $\Lambda = 0.72$, $\sigma_8 = 0.79$ cosmology. The luminosities were then simulated assuming a mass-luminosity relation with parameters $(\ln A_0, \alpha, \gamma, \sigma) = (47.4, 1.6, 1.8, 0.4)$ (see our best-fits parameters in Equation 22), the observed fluxes computed for this background cosmology, and finally, the selections appropriate for the *ROSAT* All-Sky and 400d surveys were applied. The simulated lists and the real sample have approximately the same number of clusters.

The distribution of the deviations of best-fit parameters from their nominal input values is shown in Figure 19. We are able to recover all parameters, normalization, scatter, evolution term γ, and the slope, α, (not shown in the figure), without significant biases. The widths of the distributions, $\Delta\ln A_0 = 0.085$, $\Delta\alpha = 0.14$, $\Delta\gamma = 0.42$, and $\Delta\sigma = 0.039$, correspond to the expected measurement uncertainties for each parameter. Note that the uncertainties for individual parameters are correlated. For example, the low-z normalization is obviously anticorrelated with the evolution parameter, γ. The scatter is anticorrelated with both the low-z normalization and evolution because the Malmquis bias corrections are $\propto \sigma^2$. These correlations have to be kept in mind when we estimate the uncertainties in the survey volume computations associated with the measurement errors of the L_X−M relation. In particular, the most important parameters for $V(M)$ are the average L_X's for the median mass of our low- and high-z samples. For nearby clusters, this corresponds simply to the uncertainties in A_0, but for high-z clusters, this is a complex combination of uncertainties in A_0, α, and γ. The results for the average normalizations are shown in the right panel of Figure 19. We are able to recover the true average luminosities without a significant bias and with uncertainties of $\approx$ 8.0% and 10.5% at low- and high-z, respectively.

APPENDIX B

LIKELIHOOD FUNCTION CALCULATIONS

B.1. Calculation of $p(M^{\mathrm{est}}, z)$

Generally, the probability density distribution of the observed masses is given by convolution of the model distribution of the true masses and the scatter between M^{est} and M^{true}. The former is simply the product of the theoretical mass function dn/dM^{true} and survey volume at this redshift, $dV(M^{\mathrm{true}}, z)/dz$ (the calculation of $dV(M)/dz$ is discussed in Section 5), and so we have

$$p(M^{\mathrm{est}}, z) = \left(\frac{dn}{dM^{\mathrm{true}}} \frac{dV(M^{\mathrm{true}}, z)}{dz} \right) \otimes \mathrm{scatter}(M^{\mathrm{est}}, M^{\mathrm{true}}). \tag{B1}$$

A log-normal distribution is a good approximation for the scatter in the mass estimates, and so the convolution in Equation (B1) can be written as

$$p(M^{\mathrm{est}}, z) = \frac{1}{M^{\mathrm{est}}} \frac{1}{\sqrt{2\pi}\,\sigma^{\mathrm{est}}} \int_{-\infty}^{\infty} \frac{dn}{d\ln M^{\mathrm{true}}} \frac{dV(M^{\mathrm{true}}, z)}{dz}$$
$$\times \exp\left(-\frac{(\ln M^{\mathrm{est}} - \ln M^{\mathrm{true}})^2}{2(\sigma^{\mathrm{est}})^2} \right) d\ln M^{\mathrm{true}}. \tag{B2}$$

The function $p(M^{\mathrm{est}}, z)$ enters the expression for likelihood in summation over observed clusters (the first term in Equation (25)) and in the integral over the observed range (the second term in the same equation). Equation (B2) should be evaluated numerically, but the calculation of all the terms is straightforward. The term $dn/d\ln M^{\mathrm{true}}$ is the differential cluster mass function at the given redshift. Cosmological parameters enter the calculation of $dV(M^{\mathrm{true}}, z)$ through the volume–redshift relation and the evolving cluster L_X–M relation which is derived (Section 5.1) using L_X and M_{tot} estimated in this cosmology.

B.2. Integration of $p(M^{\mathrm{est}}, z)$

We now need to evaluate the second term in Equation (25),

$$I = \int_{M_{\min}}^{M_{\max}} \int_{z_{\min}}^{z_{\max}} p(M^{\mathrm{est}}, z)\, dM^{\mathrm{est}}\, dz. \tag{B3}$$

Using Equation (B2), we have

$$I = \frac{1}{\sqrt{2\pi}} \int_{M_{\min}}^{M_{\max}} dM^{\mathrm{est}} \int_{z_{\min}}^{z_{\max}} dz \int_{-\infty}^{\infty} \frac{1}{M^{\mathrm{est}}} \frac{1}{\sigma^{\mathrm{est}}} \frac{dn}{d\ln M^{\mathrm{true}}}$$
$$\times \frac{dV(M^{\mathrm{true}}, z)}{dz} \exp\left(-\frac{(\ln M^{\mathrm{est}} - \ln M^{\mathrm{true}})^2}{2(\sigma^{\mathrm{est}})^2} \right) d\ln M^{\mathrm{true}}. \tag{B4}$$

Changing the order of integration, we have

$$I = \int_{z_{\min}}^{z_{\max}} dz \int_{-\infty}^{\infty} d\ln M^{\mathrm{true}} \frac{dn}{d\ln M^{\mathrm{true}}}$$
$$\times \frac{dV(M^{\mathrm{true}}, z)}{dz} \int_{\ln M_{\min}}^{\ln M_{\max}} \frac{1}{\sqrt{2\pi}\,\sigma^{\mathrm{est}}}$$
$$\times \exp\left(-\frac{(\ln M^{\mathrm{est}} - \ln M^{\mathrm{true}})^2}{2(\sigma^{\mathrm{est}})^2} \right) d\ln M^{\mathrm{est}}. \tag{B5}$$

The last term in this equation is the integral of the normal distribution (can be computed numerically using the library error function):

$$\int_{\ln M_{\min}}^{\ln M_{\max}} \frac{1}{\sqrt{2\pi}\,\sigma^{\mathrm{est}}} \exp\left(-\frac{(\ln M^{\mathrm{est}} - \ln M^{\mathrm{true}})^2}{2(\sigma^{\mathrm{est}})^2} \right) d\ln M^{\mathrm{est}}$$
$$= \mathfrak{N}\left(\frac{\ln M_{\min} - \ln M^{\mathrm{true}}}{\sigma^{\mathrm{est}}}, \frac{\ln M_{\max} - \ln M^{\mathrm{true}}}{\sigma^{\mathrm{est}}} \right), \tag{B6}$$

where

$$\mathfrak{N}(x_1, x_2) \equiv \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} \exp(-x^2)\, dx, \tag{B7}$$

so finally,

$$I = \int_{z_{\min}}^{z_{\max}} dz \int_{-\infty}^{\infty} \frac{dn}{d\ln M^{\mathrm{true}}} \frac{dV(M^{\mathrm{true}}, z)}{dz} \mathfrak{N}$$
$$\times \left(\frac{\ln M_{\min} - \ln M^{\mathrm{true}}}{\sigma^{\mathrm{est}}}, \frac{\ln M_{\max} - \ln M^{\mathrm{true}}}{\sigma^{\mathrm{est}}} \right) d\ln M^{\mathrm{true}}. \tag{B8}$$

The quantities σ^{est} are the total uncertainties of the mass estimates, including intrinsic scatter and measurement errors,

$$\sigma_i^{\mathrm{est}} = \left(\sigma_{\mathrm{intr}}^2 + \sigma_{\mathrm{meas},i}^2 \right)^{1/2}. \tag{B9}$$

In practice, σ_i^{est} are not the same because at least $\sigma_{\mathrm{meas},i}$ varies from cluster to cluster. A reasonable strategy to include these variations is to replace $\mathfrak{N}(\ldots, \ldots)$ with an average over all sample members,

$$I = \int_{z_{\min}}^{z_{\max}} dz \int_{-\infty}^{\infty} \frac{dn}{d\ln M^{\mathrm{true}}} \frac{dV(M^{\mathrm{true}}, z)}{dz} \left[\frac{1}{N} \sum_{i=1}^{N} \mathfrak{N} \right.$$
$$\left. \times \left(\frac{\ln M_{\min} - \ln M^{\mathrm{true}}}{\sigma_i^{\mathrm{est}}}, \frac{\ln M_{\max} - \ln M^{\mathrm{true}}}{\sigma_i^{\mathrm{est}}} \right) \right] d\ln M^{\mathrm{true}}. \tag{B10}$$

We use this equation to evaluate the second term in the expression for likelihood function (Equation (25)).

REFERENCES

Afshordi, N., Lin, Y.-T., Nagai, D., & Sanderson, A. J. R. 2007, MNRAS, 378, 293
Akritas, M. G., & Bershady, M. A. 1996, ApJ, 470, 706
Allen, S. W., Schmidt, R. W., Ebeling, H., Fabian, A. C., & van Speybroeck, L. 2004, MNRAS, 353, 457
Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2005, A&A, 441, 893
Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2007, A&A, 474, L37
Bode, P., Ostriker, J. P., Weller, J., & Shaw, L. 2007, ApJ, 663, 139
Böhringer, H., et al. 2004, A&A, 425, 367
Borgani, S., et al. 2001, ApJ, 561, 13
Borgani, S., et al. 2004, MNRAS, 348, 1078
Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
Burenin, R. A., Vikhlinin, A., Hornstrup, A., Ebeling, H., Quintana, H., & Mescheryakov, A. 2007, ApJS, 172, 561 (Paper I)
Cash, W. 1979, ApJ, 228, 939
Cavaliere, A., & Fusco-Femiano, R. 1978, A&A, 70, 677
Cohn, J. D., & White, M. 2005, Astropart. Phys., 24, 316
da Silva, A. C., Kay, S. T., Liddle, A. R., & Thomas, P. A. 2004, MNRAS, 348, 1401
de Grandi, S., & Molendi, S. 2002, ApJ, 567, 163
de Grandi, S., et al. 1999, ApJ, 514, 148
Dickey, J. M., & Lockman, F. J. 1990, ARA&A, 28, 215

Ebeling, H., Edge, A. C., Allen, S. W., Crawford, C. S., Fabian, A. C., & Huchra, J. P. 2000, MNRAS, 318, 333

Edgar, R. J., & Vikhlinin, A. 2004, *Chandra* calibration memo http://cxc.harvard.edu/cal/Acis/Cal_prods/qe/qe_memo.ps

Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605

Ettori, S. 2003, MNRAS, 344, L13

Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494

Fabricant, D., Lecar, M., & Gorenstein, P. 1980, ApJ, 241, 552

Fukazawa, Y., Makishima, K., Tamura, T., Ezawa, H., Xu, H., Ikebe, Y., Kikuchi, K., & Ohashi, T. 1998, PASJ, 50, 187

Gonzalez, A. H., Zaritsky, D., & Zabludoff, A. I. 2007, ApJ, 666, 147

Gottlöber, S., Klypin, A., & Kravtsov, A. V. 2001, ApJ, 546, 223

Hallman, E. J., Motl, P. M., Burns, J. O., & Norman, M. L. 2006, ApJ, 648, 852

Henry, J. P. 2004, ApJ, 609, 603

Hickox, R. C., & Markevitch, M. 2006, ApJ, 645, 95

Hoekstra, H. 2007, MNRAS, 379, 317

Holder, G. 2006, arXiv:astro-ph/0602251

Hu, W., & Cohn, J. D. 2006, Phys. Rev. D, 73, 067301

Hu, W., & Kravtsov, A. V. 2003, ApJ, 584, 702

Ikebe, Y., Reiprich, T. H., Böhringer, H., Tanaka, Y., & Kitayama, T. 2002, A&A, 383, 773

Jeltema, T. E., Hallman, E. J., Burns, J. O., & Motl, P. M. 2008, ApJ, 681, 167

Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, MNRAS, 321, 372

Jones, C., & Forman, W. 1984, ApJ, 276, 38

Jones, L. R., Scharf, C., Ebeling, H., Perlman, E., Wegner, G., Malkan, M., & Horner, D. 1998, ApJ, 495, 100

Kaiser, N. 1986, MNRAS, 222, 323

Kaiser, N. 1991, ApJ, 383, 104

Komatsu, E., et al. 2009, ApJS, 180, 330

Kravtsov, A. V., Nagai, D., & Vikhlinin, A. A. 2005, ApJ, 625, 588

Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128 (KVN)

Lacey, C., & Cole, S. 1993, MNRAS, 262, 627

LaRoque, S. J., Bonamente, M., Carlstrom, J. E., Joy, M. K., Nagai, D., Reese, E. D., & Dawson, K. S. 2006, ApJ, 652, 917

Lima, M., & Hu, W. 2005, Phys. Rev. D, 72, 043006

Loeb, A. 2007, J. Cosmol. Astropart. Phys., 3, 1

Łokas, E. L., Prada, F., Wojtak, R., Moles, M., & Gottlöber, S. 2006, MNRAS, 366, L26

Mahdavi, A., Hoekstra, H., Babul, A., & Henry, J. P. 2008, MNRAS, 384, 1567

Markevitch, M. 1998, ApJ, 504, 27

Markevitch, M., Forman, W. R., Sarazin, C. L., & Vikhlinin, A. 1998, ApJ, 503, 77

Markevitch, M., et al. 2003, ApJ, 583, 70

Mathiesen, B., Evrard, A. E., & Mohr, J. J. 1999, ApJ, 520, L21

Mathiesen, B. F., & Evrard, A. E. 2001, ApJ, 546, 100

Maughan, B. J. 2007, ApJ, 668, 772

Maughan, B. J., Jones, C., Forman, W., & Van Speybroeck, L. 2008, ApJS, 174, 117

Mazzotta, P., Rasia, E., Moscardini, L., & Tormen, G. 2004, MNRAS, 354, 10

Metzler, C. A., White, M., & Loken, C. 2001, ApJ, 547, 560

Mohr, J. J., Mathiesen, B., & Evrard, A. E. 1999, ApJ, 517, 627

Motl, P. M., Hallman, E. J., Burns, J. O., & Norman, M. L. 2005, ApJ, 623, L63

Mushotzky, R. F., & Loewenstein, M. 1997, ApJ, 481, L63

Nagai, D. 2006, ApJ, 650, 538

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

Nagai, D., Vikhlinin, A., & Kravtsov, A. V. 2007b, ApJ, 655, 98

Nord, B., Stanek, R., Rasia, E., & Evrard, A. E. 2008, MNRAS, 383, L10

Pacaud, F., et al. 2007, MNRAS, 382, 1289

Peng, F., & Nagai, D. 2008, ApJ, in press (arXiv:0808.3769)

Poole, G. B., Babul, A., Mc Carthy, I. G., Fardal, M. A., Bildfell, C. J., Quinn, T., & Mahdavi, A. 2007, MNRAS, 380, 437

Pratt, G. W., Böhringer, H., Croston, J. H., Arnaud, M., Borgani, S., Finoguenov, A., & Temple, R. F. 2007, A&A, 461, 71

Press, W. H., Teukolsky, S. A., Vettering, W. T., & Flannery, B. P. 1992, Numerical Recipes (Cambridge: Cambridge Univ. Press)

Reiprich, T. H. 2006, A&A, 453, L39

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

Reynolds, R. J. 1993, in ASP Conf. Ser. 35, Massive Stars: Their Lives in the Interstellar Medium, ed. J. P. Cassinelli, & E. B. Churchwell (San Francisco, CA: ASP), 338

Ricker, P. M., & Sarazin, C. L. 2001, ApJ, 561, 621

Sheldon, E. S., et al. 2001, ApJ, 554, 881

Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119

Snowden, S. L., McCammon, D., Burrows, D. N., & Mendenhall, J. A. 1994, ApJ, 424, 714

Stanek, R., Evrard, A. E., Böhringer, H., Schuecker, P., & Nord, B. 2006, ApJ, 648, 956

Sun, M., Donahue, M., Voit, M., Forman, W. R., & Jones, C. 2008, ApJ, submitted (arXiv:0805.2320)

Sunyaev, R. A., & Zeldovich, Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173

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

Tozzi, P., Rosati, P., Ettori, S., Borgani, S., Mainieri, V., & Norman, C. 2003, ApJ, 593, 705

Ventimiglia, D. A., Voit, G. M., Donahue, M., & Ameglio, S. 2008, ApJ, 685, 118

Vikhlinin, A. 2006, ApJ, 640, 710

Vikhlinin, A., Forman, W., & Jones, C. 1999, ApJ, 525, 47

Vikhlinin, A., Kravtsov, A., Forman, W., Jones, C., Markevitch, M., Murray, S. S., & Van Speybroeck, L. 2006, ApJ, 640, 691 (V06)

Vikhlinin, A., Markevitch, M., Murray, S. S., Jones, C., Forman, W., & Van Speybroeck, L. 2005, ApJ, 628, 655

Vikhlinin, A., McNamara, B. R., Forman, W., Jones, C., Quintana, H., & Hornstrup, A. 1998, ApJ, 502, 558

Vikhlinin, A., VanSpeybroeck, L., Markevitch, M., Forman, W. R., & Grego, L. 2002, ApJ, 578, L107

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

Voevodkin, A., & Vikhlinin, A. 2004, ApJ, 601, 610

Voit, G. M., & Bryan, G. L. 2001, Nature, 414, 425

Wargelin, B. J., Markevitch, M., Juda, M., Kharchenko, V., Edgar, R., & Dalgarno, A. 2004, ApJ, 607, 596

White, M. 2001, A&A, 367, 27

Zhang, Y.-Y., Böhringer, H., Finoguenov, A., Ikebe, Y., Matsushita, K., Schuecker, P., Guzzo, L., & Collins, C. A. 2006, A&A, 456, 55

Zhang, Y.-Y., Finoguenov, A., Böhringer, H., Kneib, J.-P., Smith, G. P., Kneissl, R., Okabe, N., & Dahle, H. 2008, A&A, 482, 451