A NEW ROBUST LOW-SCATTER X-RAY MASS INDICATOR FOR CLUSTERS OF GALAXIES

Andrey V. Kravtsov1,2, Alexey Vikhlinin3,4, Daisuke Nagai1

The Astrophysical Journal, submitted

ABSTRACT

We present comparison of X-ray proxies for the total cluster mass, including the spectral temperature ($T_X$), gas mass measured within $r_{500}$ ($M_g$), and the new proxy, $Y_X$, which is a simple product of $T_X$ and $M_g$ and is related to the total thermal energy of the ICM. We use mock Chandra images constructed for a sample of clusters simulated with the eulerian N-body+gasdynamics adaptive mesh refinement ART code in the concordance $\Lambda$CDM cosmology. The simulations achieve high spatial and mass resolution and include radiative cooling, star formation, and other processes accompanying galaxy formation. Our analysis shows that simulated clusters exhibit a high degree of regularity and tight correlations between the considered observables and total mass. The normalizations of the $M - T_X$, $M_g - T_X$, and $M - Y_X$ relations agree to better than $\approx 10 - 15\%$ with the current observational measurements of these relations. Our results show that $Y_X$ is the best mass proxy with a remarkably low scatter of only $\approx 5 - 7\%$ in $M_{500}$ for a fixed $Y_X$, both at low and high redshifts and regardless of whether clusters are relaxed or not. In addition, we show that redshift evolution of the $Y_X - M_{500}$ relation is close to the self-similar prediction, which makes $Y_X$ a very attractive mass indicator for measurements of the cluster mass function from X-ray selected samples.

Subject headings: cosmology: theory - galaxies: evolution - galaxies: clusters - clusters: formation - methods: numerical

1. INTRODUCTION

The evolution of the cluster abundance is one of the most sensitive probes of cosmology, which can constrain the power spectrum normalization, matter content, and the equation of state of the dark energy. The potential and importance of these constraints have motivated efforts to construct several large surveys of high-redshift clusters during the next several years. However, in order to realize the full statistical power of the upcoming cluster surveys, it is paramount that the relation between cluster mass and observables and any potential biases are well known.

Several cluster observables based on the galaxy velocities, optical light, X-ray observables such as luminosity, temperature, mass of the intracluster medium (ICM), and Sunyaev-Zel’dovich (SZ) flux have been proposed and used in the literature as proxies of the total cluster mass (see Voit 2005 for a recent comprehensive review). In this study we focus on the mass indicators derived from cluster X-ray observables, which provide a handle on the properties of the hot ICM component. X-ray luminosity, $L_X$, computed using the flux integrated within a certain radius or a range of radii, is expected to correlate with cluster mass (e.g., Kaiser 1986) and is the most straightforward mass indicator to measure observationally. $L_X$ has been used for cosmological fits to the cluster samples from ROSAT All-Sky Survey (Reiprich & Böhringer 2002; Allen et al. 2003) and Deep Cluster Survey (Borgani et al. 2001). However, $L_X$ is also the least accurate (internally) of all proposed X-ray proxies for $M_{tot}$. $L_X$ is dominated by the cluster cores and thus is particularly susceptible to non-gravitational processes in the ICM. Given the large scatter in the $L_X - T_X$ relation (e.g., David et al. 1993; Markevich 1998; Ikebe et al. 2002), the $L_X - M$ relation for real clusters probably also has significant scatter (Stanek et al. 2000). The slope of the $L_X - M$ relation deviates from the self-similar prediction (e.g., Allen et al. 2003). In addition, X-ray luminosity is notoriously difficult to reliably model in cosmological simulations (e.g., Anninos & Norman 1998; Lewis et al. 2000), a significant disadvantage given that simulations are often used to get a handle on the expected evolution of the mass vs. proxy relations. These problems could potentially be alleviated with sufficient angular resolution by exciting the flux from cluster cores, responsible for most of the scatter (Markevich 1998).

The most common choice of mass proxy used to measure the cluster number density and constrain cosmological parameters is the X-ray temperature of the intracluster plasma (e.g., Henry & Arnaud 1991; Oukbir & Blanchard 1992 Markevich 1998; Henry 2000; Seljak 2002; Ikebe et al. 2002; Pierpaoli et al. 2002). Until recently, there was a large apparent systematic uncertainty in the normalization of the $M - T_X$ relation, as evidenced, for example, by a $\approx 30 - 50\%$ discrepancy between observational measurements and cosmological simulations (e.g., Finoguenov et al. 2001; Pierpaoli et al. 2003). Over the last several years the $M - T_X$ normalization was revised both in simulations and observations due to (1) inclusion of more realistic physics in cosmological simulations (e.g., radiative cooling and star formation, Davé et al. 2003; Muanwong et al. 2002), (2) improved analyses of observed clusters using more realistic gas density profiles (e.g., Borgani et al. 2004; Vikhlinin et al. 2006), (3) more reliable measurements of the cluster temperature profiles (Markevich et al. 1998, Nevalainen et al. 2000, Arnaud et al. 2005; Vikhlinin et al. 2006), and (4) better understanding of the meaning of the mean spectral X-ray temperature, $T_X$, and the use of uniform definition of $T_X$ in observations and in simulations analyses (Mazzotta et al. 2004; Rasia et al. 2005; Vikhlinin 2006). The current agreement between models and observations is $\approx 10\%$ (see below).
The scatter in the $M - T_X$ relation is significantly reduced compared to that in the $L_X - M$ relation (the upper limit from observations is $\approx 15\%$ in $M$ for fixed $T$ for relaxed clusters; Vikhlinin et al. 2006). In general, existence of a tight relation such as $M - T_X$ indicates that clusters are remarkably regular population of objects with their global properties tightly related to total mass (e.g., Mohr et al. 1999), and scatter caused by secondary effects such as substructure in the ICM, non-gravitational processes, and mergers (O’Hara et al. 2006).

More recently, gas mass was used as a proxy for $M_{\text{tot}}$ (Vikhlinin et al. 2003; Voevodkin & Vikhlinin 2004). The practical advantage of gas mass over temperature is that it can be measured robustly from the X-ray imaging alone. Also, the $M_g - M$ relation in principle permits external calibration from the CMB measurements of the global baryon-to-dark matter ratio. Finally, $M_g$ can be expected to be less sensitive to mergers which should translate into smaller scatter in the $M_g - M$ relation. The caveat is that trend of gas mass with cluster mass and evolution with redshift are not yet fully understood.

The use of clusters as efficient probes for precision cosmology puts stringent requirements on observable cluster mass proxies: 1) tight, low-scatter correlation between the proxy and mass, 2) with the scatter insensitive to mergers, as the frequency of mergers is expected to increase with redshift, 3) simple power-law relation and evolution which can be described by a small number of parameters and be as close as possible to the prediction of the self-similar model. The point (3) is crucial to ensure that the self-calibration strategies for analyses of large cluster surveys (Levine et al. 2002; Hu 2003; Majumdar & Mohl 2003, 2004; Lima & Hu 2004, 2005; Wang et al. 2004) are successful. It is particularly important that the scatter in the observable-mass relation is low and well-behaved (Lima & Hu 2005).

In general, a mass proxy does not have to be a single cluster property, such as $L_X$, $T_X$ or $M_g$. Any physically-motivated combination of these variables that is expected to be tightly related to cluster mass can be used to construct a valid mass indicator. A hint for a better X-ray mass proxy is provided by recent studies based on cosmological simulations of cluster formation (Motl et al. 2005; Nagai 2006), which show that integrated SZ flux, $Y_{SZ}$, proportional to the product of gas mass and temperature, is a good, robust mass indicator with low scatter in the $Y_{SZ} - M$ relation, regardless of the dynamical state of the cluster. In addition, the $Y_{SZ} - M$ relation exhibits a simple, nearly self-similar evolution with redshift (da Silva et al. 2004; Nagai 2006). The physical reason for the robustness of the SZ flux is straightforward: $Y_{SZ}$ is directly related to the total thermal energy of the ICM and thus to the depth of the cluster potential well (see eq. (3) in § 2 below).

In this study, we show that a similar robust, low-scatter mass indicator can be constructed using X-ray observables. The indicator, which is simply the product of the total ICM mass and X-ray spectroscopic temperature, $Y_X = M_g T_X$, correlates strongly with cluster mass with only $\approx 5 - 8\%$ intrinsic scatter. The scatter is robust to mergers, in the sense that even for disturbed unrelaxed systems it gives unbiased estimates of mass with the statistical uncertainty similar to that for relaxed systems. Thus, the scatter of the $Y_X - M$ relation at higher redshift is similar to the scatter at $z = 0$. In addition, we show that evolution of the slope and normalization of the $Y_X - M$ relation is nearly self-similar. These properties make $Y_X$ particularly useful for measurements of cluster mass function using X-ray surveys.

### 2. MASS PROXIES

Physical properties of virialized systems, such as clusters, are expected to correlate with their total mass. For example, in the self-similar model (Kaiser 1986, 1991), the cluster gas mass is expected to be simply proportional to the total mass:

$$M_{\Delta} = C_{\Delta M} M_{\text{gas}},$$

where masses are determined within a radius enclosing a certain overdensity $\Delta$, with respect to the critical density of the universe at the epoch of observation, $\rho_{\text{crit}}(z)$, and $C_{\Delta M}$ is a constant independent of cluster mass and redshift. The self-similar relation between cluster mass and temperature is:

$$E(z) = C T^{3/2},$$

Here the function $E(z) \equiv H(z)/H_0$ for a flat cosmology with the cosmological constant assumed throughout this study is given by (e.g., Peebles 1993):

$$E(z) = \sqrt{\Omega_M(1+z)^3 + \Omega_{\Lambda}},$$

where $\Omega_M$ and $\Omega_{\Lambda}$ are the present-day density parameters for matter and cosmological constant.

The SZ flux integrated within a certain radius, $Y_{SZ}$, is proportional to the total thermal energy of the ICM gas and thus to the overall cluster potential, which makes it relatively insensitive to the details of the ICM physics and merging:

$$Y_{SZ} = \left( \frac{k_B \sigma_T}{m_e c^2} \right) \int V n_e T_e dV \propto M_{\text{gas}} T_m,$$

where $k_B$, $\sigma_T$, $m_e$, and $c$ have their usual meaning, $n_e$ are the electron number density and temperature of the gas, and $T_m$ is the gas mass-weighted mean temperature of the ICM. Combinations of equations (1, 2, 4) gives self-similar prediction for the $Y_{SZ} - M$ relation,

$$E(z)^{3/2} M_{\Delta} = C_{Y_{SZ}} Y_{SZ}^{3/5}$$

Cosmological simulations show that $Y_{SZ}$ is a good, low-scatter cluster mass proxy and that $Y_{SZ} - M$ relation form and evolution are close to the self-similar prediction (da Silva et al. 2004; Motl et al. 2005; Hallman et al. 2006; Nagai 2006). Given the good qualities of $Y_{SZ}$ as a mass proxy, it is interesting whether a similar indicator can be constructed from the X-ray observables, which could be used in studies of the X-ray cluster abundances. The simplest X-ray analog of $Y_{SZ}$ is

$$Y_X = M_g T_X,$$

where $M_g$ is the gas mass derived from the X-ray imaging data (it is measured within a radius enclosing overdensity $\Delta$), and $T_X$ is the mean X-ray spectral temperature (Mazzotta et al. 2004; Vikhlinin 2006). As we describe below in § 3, it is advantageous to measure $T_X$ excluding the central cluster region, which can be achieved with moderate angular resolution X-ray telescopes ($\lesssim 15''$ FWHM). To excise the central regions is desirable because the observed cluster temperature profiles show a greater degree of similarity outside the core (Vikhlinin et al. 2006), and also because this makes the spectral temperature closer to the gas mass averaged $T_m$ which ideally should be used in equation (5).

### 3. MOCK CHANDRA IMAGES AND ANALYSES OF SIMULATED CLUSTERS

A detailed account of the simulations, mock image generation, and analysis will be presented elsewhere (Nagai, Vikhlinin & Kravtsov 2006, in preparation). Here, we give a brief overview of the procedure and define how the observables used in this study are derived.
3.1. Simulated Cluster Sample

In this study, we use high-resolution cosmological simulations of sixteen cluster-sized systems in the “concordance” flat ΛCDM model: Ω_m = 1 − Ω_Λ = 0.3, Ω_m = 0.04286, h = 0.7 and σ_8 = 0.9, where the Hubble constant is defined as 100h km s^{-1} Mpc^{-1}, and σ_8 is the power spectrum normalization on 8h^{-1} Mpc scale. The simulations were done with the Adaptive Refinement Tree (ART) N-body+gasdynamics code (Kravtsov1998, Kravtsov et al.2003), a Eulerian code that uses adaptive refinement in space and time, and (non-adaptive) refinement in mass to reach the high dynamic range required to resolve cores of halos formed in self-consistent cosmological simulations.

To set up initial conditions we first ran low resolution simulations of two 80 h^{-1} Mpc boxes and seven 120 h^{-1} Mpc boxes, from which we selected sixteen clusters with the virial masses ranging from $M_{vir} \approx 7 \times 10^{13}$ to $2 \times 10^{15} M_{\odot}$ for re-simulation at higher resolution. High-resolution simulations were run using 128^3 uniform grid and 8 levels of mesh refinement in the computational boxes of 120 h^{-1} Mpc for CL101–107 and 80 h^{-1} Mpc for CL3–24. These simulations achieve the dynamic range of 128 × 2^8 = 32768 and peak formal resolution of $\approx 3.66 h^{-1}$ kpc and $2.44 h^{-1}$ kpc, corresponding to the actual resolution of $\approx 7 h^{-1}$ kpc $5 h^{-1}$ kpc for 120 and 80 h^{-1} Mpc boxes, respectively. Only the region of $\sim (3 - 10) h^{-1}$ Mpc around the cluster was adaptively refined, the rest of the volume was folded onto the uniform 128^3 grid. The mass resolution corresponds to the effective 512^3 particles in the entire box, or the Nyquist wavelength $\lambda_{Ny} = 0.469 h^{-1}$ and $0.312 h^{-1}$ comoving megaparsec for CL101–107 and CL3–24, respectively, or $0.018 h^{-1}$ and $0.006 h^{-1}$ Mpc in the physical units at the initial redshift of the simulations. The dark matter particle mass in the region around the cluster was $9.1 \times 10^6 h^{-1} M_{\odot}$ for CL101–107 and 2.7 $\times 10^7 h^{-1} M_{\odot}$ for CL3–24, while other regions were simulated with lower mass resolution.

The cluster simulations used in this analysis include dissipationless dynamics of dark matter, gasdynamics, star formation, metal enrichment and thermal feedback due to the supernovae type II and type Ia, self-consistent advection of metals, metallicity dependent radiative cooling and UV heating due to cosmological ionizing background (Haardt & Madau1996). The cooling and heating rates take into account Compton heating and cooling of plasma, UV heating, atomic and molecular cooling and are tabulated for the temperature range $10^2 < T < 10^9$ K and a grid of metallicities, and UV intensities using the CLOUDY code (ver. 96b4, Ferland et al.1998). The CLOUDY cooling and heating rates take into account metallicity of the gas, which is calculated self-consistently in the simulation, so that the local cooling rates depend on the local metallicity of the gas.

Star formation in these simulations was implemented using the observationally-motivated recipe (e.g., Kennicutt1998): $\rho_{*} = \rho_{gas}^\alpha r^{-1}$, with $t_\star = 4 \times 10^8$ yrs. Stars are allowed to form in regions with temperature $T < 2 \times 10^4$ K and gas density $n > 0.1$ cm^{-3}. No other criteria (like the collapse condition $\nabla \cdot \mathbf{v} < 0$) are used. Comparison of the runs with different choices of the threshold for star formation, $n = 10, 1, 0.1, 0.01$ cm^{-3}, shows that the threshold affects gas fractions at small radii, $r/r_{vir} < 0.1$, but its effect is negligible at the radii we consider in this study. Further details about this simulated cluster sample will be presented in Nagai, Vikhlinin & Kravtsov (2006, in preparation).

The properties of the simulated cluster sample are summarized in Table1. We list the total and hot ICM masses defined within a radius $r_{500}$ enclosing the overdensity of 500 with respect to the critical density at $z = 0$. We also list the mean spectral temperature measured using mock Chandra spectra from the radial region of $(0.15 - 1)r_{500}$, as described in §3.3. Although the simulated cluster sample is small, the objects cover more than an order of magnitude in mass and are simulated with very high resolution which allows us to take into account effects of galaxy formation on the ICM.

In what follows, we use cluster total mass and observables measured within $r_{500} \approx 0.5 - 0.6 r_{vir}$, where $r_{vir}$ is a traditional definition of cluster virial radius using the virial overdensity $\Delta_{vir} (\approx 337 at z = 0$ for the cosmology adopted in our simulations) with respect to the mean density at $z = 0$. This choice of the outer radius is mainly motivated by the fact that clusters are more relaxed within $r_{500}$ compared to the outer regions (Evrard et al.1996, Lau et al.2006). Also, the analysis of real clusters is often limited to a similar radius because of the limited field of view.

3.2. Mock Chandra Images

We first create the X-ray flux maps of the simulated clusters viewed along three orthogonal projections. The flux map is computed by projecting the X-ray emission arising from the computational cells within 3 $R_{vir}$ of the cluster along a given line-of-sight, taking into account the actual gas density, temperature, and metallicity of each cell in a simulation output. We compute the X-ray plasma emission, $\Delta E(T, Z, z)$, using the MEKAL code with the relative elemental abundance from Anders & Grevesse (1989), and assuming interstellar absorption with a hydrogen column density of $n_H = 2.0 \times 10^{20}$ cm^{-2}. This provides expected emission spectra in the 0.1–10 keV energy range on a spatial grid 1024 × 1024 pixels with the size $4.88 h^{-1}$ kpc for CL101–107 and 2.44 $h^{-1}$ kpc for CL3–24. The entire map is therefore 5.0 h^{-1} Mpc and 2.5 $h^{-1}$ Mpc, respectively. In generating the spectral maps we assume the redshift of cluster observation is $z_{obs} = 0.06$ for the simulation output at $z = 0$ and $z_{obs} = 0.6$ for the $z = 0.6$ sample.

Next, we simulate the photon map by convolving the spectrum from each image pixel with the on-axis response of the Chandra ACIS-I camera and drawing the number of photons in each spectral channel from the Poisson distribution. We simulate two sets of the mock Chandra photon maps. In the
first set we assume a 100 ksec Chandra exposure, which is fairly typical for the real deep observations. From this set, we generate cluster images in the 0.7–2 keV band. We then add the uniform Poisson background with the intensity typical for ACIS-I data (Markovitch et al. 2003). These images are used to identify and mask out from the further analysis all detectable small-scale clumps, as routinely done by observers. These clumps contain a small fraction of gas mass and do not bias gas mass estimates. They can, however, bias temperature measurement significantly. Our clump detection is fully automatic and based on the wavelet decomposition algorithm described in Vikhlinin et al. (1998).

All further analysis is performed using the second set of photon maps generated for very long exposures, 10^6 sec for the z = 0 sample and 10^5 sec exposure for the z = 0.6 sample. Here, the exposure time is chosen to ensure ≈10^6 – 10^7 photons outside the cluster core region for all simulated clusters. This second set of data is used in our analysis to derive gas mass and mean spectroscopic temperature of the ICM. The exposures are artificially large by design as we are interested in the intrinsic scatter of the X-ray observable mass relation, not the statistical errors due photon noise in a particular choice of short exposure. Also, we ignore further complications present in reduction of real Chandra data, including background subtraction and spatial variations of the effective area (i.e., we assume that accurate corrections for these effects can be applied to the real data and any associated uncertainties are included in the reported measurement errors).

3.3. Analysis of Mock Chandra Data

We analyze the mock data for the simulated clusters employing the techniques and tools used in analyses of the real Chandra cluster observations, as described in detail by Vikhlinin et al. (2006). After masking out detectable substructures (see above), we fit the X-ray spectra in concentric annuli to measure the projected temperature and metallicity profiles. Next, we measure the X-ray surface brightness profile in the 0.7–2 keV band (used in the real data analysis because it provides the optimal signal-to-noise ratio). Using the effective area as a function of energy and observed projected temperature and metallicity at each radius, we convert the observed brightness from units of count rate to the projected emission measure, \( EM = \int n_e n_p d l \). The derived projected emission measure profile is fit to a three-dimensional model of \( \rho_{\text{gas}}(r) \) that has great functional freedom and can independently describe the gas density slopes at both \( r \gtrsim 500 \), and at smaller radii and in the very inner region. The best fit model directly provides the gas mass profile, \( M_{\text{gas}}(r) \).

We also measure the average X-ray spectral temperature, \( T_X \), from a single-temperature fit to the spectrum integrated within \( r_{500} \), excluding the central region strongly affected by radiative cooling. The inner cut is set at a fixed fraction \( r_{500}, r_{\text{in}} = 0.15 r_{500} \). Note that we choose to cut out central region defined using a fixed fraction of \( r_{500} \) rather than a fixed metric radius of 70 kpc as in Vikhlinin et al. (2006). This new definition for \( T_X \) results in only a small correction (~3% on average) of the \( T_X \) values reported in Vikhlinin et al. (2006), who used notation of \( T_{\text{pec}} \) instead of \( T_X \). We choose the specific value of \( r_{\text{in}} = 0.15 r_{500} \) because beyond this radius the observed temperature profiles of clusters are approximately self-similar, while at smaller radii the profiles show a large scatter. The choice of \( r_{\text{in}} \) thus should maximize the self-similarity of the relation between mass and \( T_X \).

In our analysis below we distinguish the unrelaxed clusters from the relaxed systems to test for sensitivity of the mass proxies and substructure. Specifically, as is usually done to classify observed clusters, we visually examine mock 100 ksec Chandra images for \( x, y, \) and \( z \) projections of each cluster and classify as “relaxed” clusters which have regular X-ray morphology and no secondary maxima and minimal deviations from the elliptical symmetry. “Unrelaxed” clusters are those with secondary maxima, filamentary X-ray structures, or significant isophotal centroid shifts.

4. COMPARISON OF MASS INDICATORS

In this section we compare three indicators of cluster mass: X-ray spectral temperature \( T_X \), gas mass \( M_g \), integrated SZ flux \( Y_{\text{SZ}} \), and its X-ray analog \( Y_X \equiv M_g T_X \).

Figures 1–3 show the \( M_{500} – M_g \), \( M_{500} – T_X \), and \( M_{500} – Y_g \) relations, respectively. We present correlations at redshifts \( z = 0.0 \) and \( z = 0.6 \), classifying clusters into relaxed and unrelaxed based on their X-ray image morphology, as described in § 2.3. We performed power law fits to these relations and present best fit values of parameters, as well as the amount of scatter around the best fit relation for different subsets of clusters in Table 2. The upper half of the table shows fits for best fit power law normalization and scatter for the values of slope fixed to the value predicted by self-similar model. The lower half shows best fit values for the case when both normalization and slopes were allowed to vary.

For comparison, we also include fits to the \( Y_{\text{SZ}} – M \) relation, where \( Y_{\text{SZ}} \) is the three-dimensional integrated SZ flux measured within \( r_{500} \) (Nagai 2006). The power law fits were performed for each projection (i.e., \( x, y, \) and \( z \)) separately. We estimate the uncertainties in the best fit parameters by generating 10000 bootstrap samples (Press et al. 1992) and calculating the dispersion of the best fit power law normalization and slope among the samples.

Figure 1 and Table 2 show that the slope and evolution of the \( M_{500} – T_X \) relation are quite close to the self-similar model. There is a ~ 20% scatter in \( M \) around the mean relation and much of the scatter is due to unrelaxed clusters. Note also that the normalization of the \( M_{500} – T_X \) relation for unrelaxed systems is somewhat biased with respect to that for the unrelaxed subsample. Unrelaxed clusters have somewhat lower temperatures for a given mass. This may seem counter-intuitive at first, given that one can expect that shocks can boost the ICM temperature during mergers. However, in practice the effect of shocks is relatively small (e.g., Hara et al. 2006). The main source of the bias is that during advanced mergers the mass of the system already increased but only a fraction of the kinetic energy of merging systems is converted into the thermal energy of the ICM (see, e.g., Mathiesen & Evrard 2001).

The \( M_g – M_{500} \) relation (Fig. 2) has a somewhat smaller scatter (≈ 10 – 12%) around the best fit power law relation than the \( M – T_X \), but its slope is significantly different from the self-similar prediction — we find \( M_{500} \propto M_g^{0.88±0.02} \) compared to the expected \( M_{500} \propto M_g \). This is due to the trend of gas fraction with cluster mass, \( f_{\text{gas}} \equiv M_g / M_{500} \propto M_{500}^{1.0±0.2} \) present for both the simulated clusters in our sample (see Kravtsov et al. 2005) and for the observed clusters (Vikhlinin et al. 2006). The normalization of the \( Y_X – M_{500} \) relation, on the other hand, evolves only weakly between \( z = 0.6 \) and \( z = 0 \) (yet, statistically significant evolution is present; see Table 2, reflecting slow evolution of the gas fraction with time (Kravtsov et al. 2005).

The \( Y_X – M_{500} \) relation (Fig. 3) has the smallest scatter, only
Fig. 1.— Relation between the X-ray spectral temperature, $T_X$, and total mass, $M_{500}$. $T_X$ is measured within the radial range $[0.15 - 1]r_{500}$. Separate symbols indicate relaxed and unrelaxed clusters, and also $z = 0$ and $z = 0.6$ samples. The dashed line shows the power law relation with the self-similar slope fit to the entire sample, and the dotted lines indicate 20% scatter.

Note that this value of scatter includes clusters at both low and high redshifts and both relaxed and unrelaxed systems. In fact the scatter in $Y_X - M_{500}$ for relaxed and unrelaxed systems is indistinguishable within the errors (see Table 2). Note also that the figures include points corresponding to the three projections of each cluster. Figure 2 shows that the dispersion in the projected values of $Y_X$ for each given cluster is very small, which means that $Y_X$ is not very sensitive to the asphericity of clusters. Remarkably, the scatter of $Y_X - M_{500}$ is as small as that in the three-dimensional $Y_{SZ} - M_{500}$ relation (see Table 2). This implies that $Y_X$ is in fact a better mass indicator than SZ flux, when additional possible sources of scatter for the real measurements of the SZ flux, such as projecting structures outside $r_{500}$, are taken into account.

The comparison of the mass proxies, clearly shows that $Y_X$, the product of gas mass and X-ray spectral temperature, is by far the most robust and most self-similar mass indicator. Why is the product better than its parts? Figure 3 shows that the answer lies in the anti-correlation of the residuals of temperature and gas mass from their respective relations with total cluster mass. We plot residuals from the best fit power laws with the slope value fixed to the self-similar value (i.e., the upper portion of Table 2) to illustrate both random scatter and systematic deviations from self-similar behavior.

Figure 4 also shows that the clusters with temperatures lower than the mean temperature for a given total mass tend to have gas mass higher than the mean, and vice versa. Note also that there is some redshift evolution between $z = 0$ and $z = 0.6$ — more clusters have negative deviations of temperature and positive deviations of measured gas mass at $z = 0.6$ compared to $z = 0$. This redshift evolution is thus in the opposite direction for the gas mass and temperature deviation. The measured $M_g$ systematically increases at higher $z$ for a fixed total mass, because the clusters become less relaxed on average. For unrelaxed clusters, the ICM density distribution is non-uniform which results in overestimation of $M_g$ from the X-ray data (Mathiesen et al. 1999). Some of the decrease of $M_g$ at lower $z$ may be due to continuing cooling of the ICM which decreases the mass of hot, X-ray emitting gas.

The anti-correlation of residuals and opposite evolution with redshift for gas mass and temperature is the reason why the behavior of their product, on average, has smaller scatter and is closer to the self-similar expectation in both the slope and evolution. We discuss the origin of this behavior further in §6 below.

5. PRACTICAL ALGORITHM FOR ESTIMATING CLUSTER MASS USING $Y_X$

Suppose we have $Y_X - M_{500}$ relation pre-calibrated by some external means: $M_{500} = C E(z) Y_X^{\eta}$, and we would like to use it to estimate $M_{500}$. The pre-calibration can be done using a
well-observed sample of relaxed clusters or simulations. Note that the definition of $Y_X$ includes spectral temperature and gas mass within $r_{500}$. In practice, however, when we want to use $Y_X$ to estimate $M_{500}$ we do not know the value of $r_{500}$ a priori. This is not a problem for measurements of $T_X$, because the integrated spectral temperature is dominated by the signal from the inner, brighter regions of the clusters and is thus not sensitive to the exact choice of the outer radius. For example, we checked that for clusters in the Vikhlinin et al. (2006) sample, a fairly drastic change of the outer radius to 0.5$r_{500}$ from the fiducial choice of $r_{500}$ adopted in our analysis, results in a systematic increase of the mean spectral temperature of only $\approx 5\%$. The sensitivity to the exact choice of the inner region is also comparably small, as was discussed above in § 5.5.

The measurement of gas mass, however, will depend quite sensitively on the adopted outer radius, because the profile $M_g(r)$ does not converge at large $r$. In practice, therefore, we have to solve for $r_{500}$ simultaneously with estimating $M_{500}$, which can be done with the following iterative algorithm. An initial value for $T_X$ can be obtained by using the integrated temperature (including the center) to estimate an approximate value of $r_{500}$ through a crude mass-temperature relation, and then re-measuring the spectrum in the radial range $(0.15 - 1)r_{500}$. Then, using identity $M_{500} = 4/3 \pi 500 \rho_c(z) r_{500}^2$, we can re-write the $Y_X - M$ relation as

$$4/3 \pi 500 \rho_c(z) r_{500}^2 = C E(z)^{1/2} [Y_X(r_{500})]^{\alpha},$$

(7)

where $Y_X(r) \equiv T_X M_g(r)$ and the constant $C$ is from pre-calibrated $Y_X - M_{500}$ relation. Equation (7) is used to solve for $r_{500}$. A second iteration can then be done using $T_X$ re-measured in the newly estimated radial range $(0.15 - 1)r_{500}$, continuing the procedure until convergence. Once $r_{500}$ is determined, $M_{500} = 4/3 \pi 500 \rho_c(z) r_{500}^2$.

It is useful to know how the observational calibration of the normalization constant $C$ in equation (7) scales with the assumed value of the Hubble constant. Cluster masses determined dynamically or through gravitational lensing scale as $M \propto h^{-5/2}$. This implies that the angular size that corresponds to $r_{500}$ is $h$-independent. The gas mass derived from X-ray data for a fixed angular radius scales as $M_g \propto h^{-1}$. Therefore, the observed normalization of the $Y_X - M$ relation with the slope close to self-similar scales as $C \propto h^{1/2}$.
of which is due to unrelaxed clusters. The unrelaxed clusters of gas, due to incomplete relaxation.

The energy of merging systems is converted into the thermal energy system already increased but only a fraction of the kinetic energy, because during advanced, ongoing mergers, the mass of the system has temperatures biased low for a given mass. This is likely due to self-similarity in its slope, which is due to the dependence of gas fraction within \( r_{500} \) on the cluster mass \( M_{\text{500}} \). A similar dependence exists for the observed clusters \( \text{(Vikhlinin et al. 2006)} \) and we can thus expect similar trends in the \( M_{g} - M_{500} \) relation for real clusters.

Generally, all the observable–mass relations we tested demonstrate a remarkable degree of regularity of galaxy clusters as a population. \( T_X, M_g, \) and \( Y_X \) all exhibit correlations with \( M_{500} \) which are close to the expectation of the self-similar model, both in their slope and evolution with time, within the uncertainties provided by our sample. The only exception is the slope of the \( M_{g} - M_{500} \) relation.

Given that our analysis relies on cosmological simulations, it is reasonable to ask whether the simulated clusters are realistic. Although simulations certainly do not reproduce all of the observed properties of clusters, especially in their core regions, the ICM properties outside the core in simulations and observations agree quite well. We illustrate this in Fig. 6 which shows that the \( M_{500} - T_X \) relations for simulated and observed clusters (the sample of relaxed clusters at \( z \approx 0 \) from \text{Vikhlinin et al. 2006}). For this comparison we use only those clusters from simulations that appear regular and relaxed in their X-ray surface brightness image. Clearly, both simulated and observed clusters exhibit tight correlations between \( M_{500} \) and \( T_X \) (see also \text{Mohr et al. 1999, Voevodkin et al. 2002}) which agree remarkably in their slope \( (M_{g} \propto T^{1.75}) \) and normalization\(^6\). The normalizations derived from simulated and real clusters agree to \( \approx 10\% \), while slopes are indistinguishable and both deviate significantly from the expected self-similar value of 1.5. This is a consequence of significant trends in the gas fraction with cluster mass, \( M_{g}/M_{500} \approx M_{T}^{0.2+0.25} \) for both simulated \( \text{(Kravtsov et al. 2005)} \) and observed clusters \( \text{(Vikhlinin et al. 2006)} \). The deviations from the self-similar model also manifest themselves in the absence of any noticeable evolution with redshift\(^7\). Interestingly, the real clusters show a similarly weak evolution in the \( M_{g} - T \) relation \( \text{(Vikhlinin et al. 2002)} \).

Figure 6 shows that the likely explanation is that the clusters at \( z = 0.6 \) tend to be colder for the fixed \( M_{\text{tot}} \) but have higher estimated \( M_{g,500} \) than their counterparts at \( z = 0 \) because they are less relaxed.

A similar level of agreement between the simulations and the latest Chandra measurements exists also for the total mass vs. temperature relation, \( M_{500} - T_X \). In fact, the normalization for our simulated sample (Table 2) agrees with the observational results of \text{Vikhlinin et al. 2006)} which are close to \( \approx 10\% \). This is a considerable improvement given that significant disagreements existed just several years ago (see § 1). The residual systematic 10% difference in the normalization is likely caused by non-thermal pressure support from bulk gas motions \( \text{(Faltenbacher et al. 2003, Rasia et al. 2006, Lau et al. 2006)} \), which is unaccounted for by the X-ray hydrostatic mass estimates.

In Figure 6 we compare the \( Y_X - M_{500} \) relation for the simulated clusters and for the Chandra sample of \text{Vikhlinin et al. 2006}). The observed clusters show a tight correlation with the slope close to the self-similar value. There is \( \approx 15\% \) difference in normalization, likely explained also by neglecting the turbulent pressure support in the Chandra hydrostatic mass estimates.

\(^6\) For this comparison, the gas masses of simulated clusters are rescaled by a factor of \( 0.17/0.143 = 1.19 \) to reflect the difference between the universal baryon fractions adopted in the simulation and the value measured by the WMAP \( \text{Spergel et al. 2003, Tegmark et al. 2004)} \).

\(^7\) Note that \( M_{g} \) in Fig. 6 is not multiplied by the \( E(z) \) factor unlike the total mass in Fig. 2.
The excellent agreement of simulations and observations in terms of the relation between the two X-ray observables used to compute $Y_X (M_500 - T_X)$ and a relatively good agreement in the $T_X - M_500$ and $Y_X - M_500$ relations, gives us confidence that the results presented in this paper are sufficiently realistic. One can ask whether the real clusters show the same trend of decreasing scatter when $M_g$ and $T_X$ are used as mass indicators instead of $T_X$. Unfortunately, the existing data cannot answer this question because the mass measurement uncertainties for individual clusters are of order or larger than the expected scatter for relaxed clusters.

Our results show that $Y_X$ is clearly the most robust and most self-similar X-ray cluster mass indicator. The biases existing in mass estimates based on $M_g$ and $T_X$ anti-correlate both for a given redshift and in terms of evolutionary trends (see Figure 5). This explains why their product, $Y_X$, is a better mass indicator than $T_X$ and $M_g$ individually. The quality of $Y_X$ compares well to that for the actual three-dimension integral of the ICM thermal energy (proportional to $Y_{SZ}$) in terms of its low scatter and self-similarity (see Table 2). $Y_X$ may prove to be an even better mass proxy than $Y_{SZ}$, given that we use ideal 3D measurement of the latter while reproducing the actual data analysis for the former. Note also that $Y_{SZ}$ is more sensitive to the outskirts of clusters, because it involves gas mass-weighted temperature (as opposed to the spectral temperature more sensitive to the inner regions), and thus should be more prone to projection effects.

Note that $Y_X$ is also an attractive mass proxy from the data analysis point of view. First, it reduces statistical noise by combining the two independently measured quantities, $M_g$ and $T_X$, into a single quantity. Consider for example how mass estimates are affected by $T_X$, the parameter that is the most difficult to measure and which is most affected by the dynamical state of cluster. A 10% observational uncertainty or a real deviation due to a merger in $T_X$ translates into a $\sim 15\%$ mass uncertainty through the $M - T_X$ relation and only $6\%$ uncertainty through the $Y_X - M$ relation. $Y_X$ is also less sensitive to any errors in the absolute calibration of the X-ray telescope. For example, a mis-calibration of the low energy effective area typically translates into $T_X$ and $M_g$ measurement errors of the opposite sign, $\delta T/T \approx -2 \delta M_g/M_g$. These errors partially cancel in $Y_X, \delta Y/Y \approx 0.5 \delta T/T$, and are further reduced in the mass estimate: $\delta M/M = 0.6 \delta Y/Y \approx 0.3 \delta T/T$. The error in mass is larger when other proxies are used, $\delta M/M = \delta M_g/M_g \approx 0.5 \delta T/T$ for the $M_g - M$ relation and $\delta M/M = 1.5 \delta T/T$ for the $M - T$ relation.

The robustness and low scatter make $Y_X$ an excellent mass indicator for observational measurements of cluster mass function at both $z = 0$ and higher redshifts. The necessary data — an X-ray brightness profile and a wide-beam spectrum excluding the core — are easily obtained with sufficiently deep observations with Chandra, XMM-Newton, and Suzaku telescopes. The small scatter and simple, nearly self-similar evolution of the $Y_X - M$ relation hold promise for the self-calibration strategies for future large X-ray cluster surveys.

This project was supported by the National Science Foundation (NSF) under grants No. AST-0206216 and AST-0239759, by NASA through grant NAG5-13274, and by the Kavli Institute for Cosmological Physics at the University of Chicago. AV is supported by the NASA grant NAG5-9217 and contract NAS8-39073. The cosmological simulations used in this study were performed on the IBM RS/6000 SP4 system (copper) at the National Center for Supercomputing Applications (NCSA). D.N. is supported by a Sherman Fairchild postdoctoral fellowship at Caltech. We have made extensive use of the NASA Astrophysics Data System and arXiv.org preprint server.
