THE ASTROPHYSICAL JOURNAL, 617:303–314, 2004 December 10

EXTRACTING THE DARK MATTER PROFILE OF A RELAXED GALAXY CLUSTER

J. S. ARABADJIS,[1] M. W. BAUTZ,[1] AND G. ARABADJIS[2]
Received 2003 May 28; accepted 2004 August 17

ABSTRACT

Knowledge of the structure of galaxy clusters is essential for an understanding of large-scale structure in the universe and may provide important clues to the nature of dark matter. Moreover, the shape of the dark matter distribution in the cluster core may offer insight into the structure formation process. Unfortunately, cluster cores also tend to be the site of complicated astrophysics. X-ray imaging spectroscopy of relaxed clusters, a standard technique for mapping their dark matter distributions, is often complicated by the presence of cool components in cluster cores, and the dark matter profile one derives for a cluster is sensitive to assumptions made about the distribution of this component. In addition, fluctuations in the temperature measurements resulting from normal statistical variance can produce results that are unphysical. We present here a procedure for extracting the dark matter profile of a spherically symmetric, relaxed galaxy cluster that deals with both of these complications. We apply this technique to a sample of galaxy clusters observed with the *Chandra X-Ray Observatory* and comment on the resulting mass profiles. For some of the clusters we compare their masses with those derived from weak and strong gravitational measurements.

Subject headings: dark matter — galaxies: clusters: general — X-rays: galaxies: clusters

1. INTRODUCTION

The cold dark matter (CDM) paradigm of modern cosmology has enjoyed spectacular success in describing the formation of large-scale structure in the universe (Navarro et al. 1997; Moore et al. 1999b; Lahav et al. 2002; Peacock et al. 2001). There are, however, several nagging inconsistencies between the results of numerical CDM experiments and observations. On small scales, the dark matter halos in dwarf and low surface brightness galaxies are much less cuspy than in CDM simulations (Burkert 1995; McGaugh & de Blok 1998; Moore et al. 1999b). Disk galaxies produced in simulations tend to have inadequate masses and angular momenta (Navarro & Steinmetz 2000). The number of Milky Way satellites appears to be at least an order of magnitude lower than CDM predictions (Kauffman et al. 1993; Moore et al. 1999a; Klypin et al. 1999). On larger scales, some studies (Tyson et al. 1998; Smail et al. 1995) report galaxy clusters with central density profiles that are flatter than CDM predictions, although these are somewhat controversial (Broadhurst et al. 2000; Shapiro & Iliev 2000).

The density profile of bound structures that form through the hierarchical assembly of smaller structures is usually parameterized as a power law at small scales and a separate power law on large scales (e.g., Jing & Suto 2002):

$$\rho(r) = \frac{\rho_0}{(r/r_s)^{\alpha}(1 + r/r_s)^{\gamma-\alpha}}. \qquad (1)$$

The four parameters in this description are the density ρ_0 at some fiducial radius, the inner power-law index α, the outer power-law index γ, and the scale radius r_s setting the break between the two power laws. While it is generally agreed that $\gamma = 3$, the value of α has generated considerable debate. Simulations predict a value between 1.0 (Navarro et al. 1996, 1997) and 1.5 (Moore et al. 1999b; Fukushige & Makino 2001), roughly independent of halo mass and formation epoch. In nature, however, α shows a larger variation and is likely a function of halo mass. Hα rotation curves of low surface brightness galaxies indicate density profiles that are significantly flatter ($\alpha \sim 0.5$) than CDM predictions (Swaters et al. 2000, 2003; Dalcanton & Bernstein 2000; Borriello et al. 2003). X-ray observations of galaxy clusters generally show steeper profiles, however, with $\alpha \sim 1.2$ (Lewis et al. 2003) to 1.9 (Arabadjis et al. 2002).

These discrepancies are often ascribed to limitations of the astrophysics or the physics included in the simulations. Baryon physics, if included, may be tacked on at the conclusion of a simulation according to a set of semianalytic and/or empirical prescriptions. It is likely that baryon physics will play a significant role in the evolution of the central halo. Reports of a halo "entropy floor" (Ponman et al. 1999; Lloyd-Davies et al. 2000) suggest nongravitational sources of heating and feedback either prior to or during halo formation (Balogh et al. 1999; Loewenstein 2000; Wu et al. 2000) that are probably baryonic in origin (see, however, Mushotzky et al. 2003). The question then becomes one of determining where baryon physics *ceases* to be important. While the inclusion of baryon astrophysics in sufficient detail may remedy these problems, its effects will require a great deal of effort to disentangle (Frenk 2002).

It could be, however, that the missing ingredients in the simulations are not all astrophysical. One possibility is that the initial power spectrum of the primordial fluctuations is not scale invariant. If the primordial spectral index of density perturbations is not precisely 1 (as is normally assumed by appealing to standard inflationary cosmology), the formation epoch of halos may be delayed sufficiently to ameliorate the central density problem (Alam et al. 2002; Zentner & Bullock 2002). Another possibility is that important dark matter particle physics is being overlooked and that the assumption of no nongravitational self-interactions is faulty. Proposed modifications of CDM include, although are not limited to, self-interacting dark matter (Spergel & Steinhardt 2000; Firmani et al. 2000), warm dark matter (Hogan & Dalcanton 2000), annihilating dark

[1] Center for Space Research, Massachusetts Institute of Technology, 70 Vassar Street, Building 37, Cambridge, MA 02139; jsa@space.mit.edu, mwb@space.mit.edu.
[2] Mitre Corporation, 202 Burlington Road, Bedford, MA 01730; gus@mitre.org.

matter (Kaplinghat et al. 2000), scalar field dark matter (Hu & Peebles 2000; Goodman 2000), and mirror matter (Mohapatra et al. 2002), each of which is invoked to soften the core density profile. Many of these modifications will soften the core profile of galaxy clusters as well, in conflict with many X-ray determinations of mass profiles, although other astrophysical processes such as the adiabatic contraction of core baryons (Hennawi & Ostriker 2002) may mitigate this effect.

In an effort to discriminate between CDM modifications and other astrophysical influences, we are mapping the dark matter profiles of a large sample of galaxy clusters. Specifically, we use imaging spectroscopy from the *Chandra X-Ray Observatory* (Weisskopf et al. 2002) to determine the deprojected temperature and density profiles of the baryonic content of each galaxy cluster, which we then use to derive its dark matter profile. In this paper we describe our method and apply it to a sample of low- and moderate-redshift clusters. We describe our spectral deprojection technique in § 2. We discuss the problems involved in converting these results to a mass profile and our solution in § 3. We examine the effects that cooling flow model assumptions have on our profiles and present a statistical analysis of the models and a prescription for choosing among them using Markov Chain Monte Carlo sampling in § 4. In § 5 we apply these techniques to a sample of *Chandra* clusters and discuss our results. Finally, we summarize our findings in § 6. In a subsequent paper we will examine these profiles for their implications for large-scale structure formation and dark matter particle properties (J. S. Arabadjis & M. W. Bautz 2004, in preparation, hereafter AB04).

2. SPECTRAL DEPROJECTION

We begin with a *Chandra* imaging spectroscopic observation of a galaxy cluster using the ACIS detector, with either the S3 chip or the I array. We start with a level 2 data set that has been processed in the usual way, filtered for periods of high background using the procedures described in the CIAO Science Threads,[3] with point sources removed (for details see Arabadjis et al. 2002). After locating the center of the projected emissivity, we lay down a series of adjacent, concentric annuli centered on the emission peak. The annular dimensions are set to include enough source photons (1000–2000+) to reliably determine the plasma temperature. RMF and ARF response matrices are constructed, as is a background spectrum from a region external to the outermost annulus. The spectra are recorded in PI format and grouped such that there is a minimum of 20 counts per channel.

Our spectral deprojection has been presented elsewhere (Arabadjis et al. 2002), and so we just briefly summarize here. To derive spherical radial profiles, we construct a model consisting of N concentric spherical shells whose inner and outer radii correspond to the inner and outer cylindrical radii of the projected annuli in the data set. The volume intersection matrix $\boldsymbol{V}$, whose elements V_{ij} contain the volume of spherical shell j intersected by a cylindrical shell formed by the projection of annulus i, is used to set the linear relations between each of the normalizations as specified by the binning geometry. Each shell i on $[1, N]$ (1 is the innermost shell—a sphere—and N is the outermost shell) contains an optically thin thermal plasma whose emission characteristics are determined by the MEKAL model (Mewe et al. 1985, 1986; Kaastra 1992; Liehdal et al. 1995) within XSPEC (Arnaud 1996) using two free parameters, the temperature T and the normalization K.

[3] See http://asc.harvard.edu/ciao/threads/all.html.

In some models we allow the innermost N_c shells to contain a second emission component at a (lower) temperature T_c as a first-order treatment of the cooler plasma. This cool component is assumed to be in pressure equilibrium with the hot component; this means that the cool and hot components cannot both be in hydrostatic equilibrium. We *assume* that the hot component is in hydrostatic equilibrium. We adopt this form for the cooling flow model for two reasons: (1) there is no evidence that the plasma in cooling flow cluster cores cools below about 1 keV (Peterson et al. 2003) as in the isobaric cooling flow model of Mushotzky & Szymkowiak (1988), and (2) it is arguably the simplest adjustment that can be made to the uniphase model. The cool component could be arranged in droplets or filaments that are replenished as they migrate to the unspecified sink at $r = 0$, effecting a hydrodynamical equilibrium. The details of the geometry of the cool component are unimportant since they are below our spatial resolution limit; it is only our assumption that it is in pressure equilibrium with the hot component that has observational consequences. In this study we use the pressure gradient in the core to measure the radial dependence of the enclosed gravitating mass.

The number of parameters in this model is rather large. Each MEKAL component contains six parameters, of which two (T and K) are allowed to vary. Thus, the N annuli in the data set are modeled using $N + N_c$ emission components. Including a Galactic absorption column yields a model with $6(N + N_c) + 1$ parameters (although only a subset of these actually vary). Each of the N annuli independently constrains between 1 and N of the model shells in the fitting, and so the complete XSPEC model contains $N[6(N + N_c) + 1]$ components, $2(N + N_c) + 1$ of them variable. The current version of XSPEC admits models with 1000 parameters, any 100 of which can vary. This limits our model to $N = 12$ annuli for $N_c = 0$ (876 parameters, 25 variable) and $N = 10$ annuli for $N_c = 2$ (730 parameters, 25 variable). We have written a program that reads in the data annuli dimensions and writes out an XSPEC script that handles all of the data manipulation. The script initializes the model parameters, performs the χ^2 minimization, and calculates parameter uncertainties. This software is available to the public through requests to the authors.

Most spectral deprojection schemes rely on an "onion peeling" approach (Fabian et al. 1981; Allen & Fabian 1997; David et al. 2001; Lewis et al. 2002; Sun et al. 2003): the outermost annulus is modeled using the outermost spherical shell; its model parameters are then frozen and its emission is subtracted from all annuli interior to it. The next most outer shell is then modeled, the resulting model again frozen and subtracted from the interior, and so forth, until the entire cluster has been modeled. The virtue of this technique is that the number of model components scales as N instead of N^2, allowing for greater spatial detail. However, because the parameters of each model shell are frozen and the model is subtracted from interior shells as if it contained zero uncertainty, the technique does not find the global "best fit" of the model parameters. In many cases, the errors quoted in onion peeling analyses are the uncertainties associated with a single layer of the onion (David et al. 2001). Error estimates derived from Monte Carlo simulations are more reliable, limited primarily by the number of simulations used to estimate the uncertainties (Lewis et al. 2002). In our study all of the model parameters are fitted simultaneously; the subsequent determination of error in subsets of interesting variables is a true expression of the parameter uncertainties in the model. We note here that although our model appears to have many more parameters than the other techniques, in actuality the numbers are equivalent because of the web of linear dependencies among the component normalizations.

3. MASS PROFILES

Many deprojection methods rely on analytic formulae for the radial run of temperature, surface brightness, or mass, either during or after the fitting process (Allen et al. 2001; David et al. 2001; Hicks et al. 2002; Pizzolato et al. 2003). The greatest advantage of a parametric treatment is numerical stability. In addition, it is common practice to smooth noisy profiles before using them in subsequent calculations. The latter technique is especially useful when deriving gravitating mass profiles since a numerical derivative must be computed. A serious drawback of these approaches is that it is difficult to quantify the effect of the parameterization, or the smoothing, on the results. Additionally, it is often difficult or impossible to propagate errors through to the results.

Our nonparametric deprojection technique does not guarantee smooth temperature and density profiles, so we have devised a method whereby the mass profile is computed from within the error envelope of $[\rho(r),\ T(r)]$. By choosing a statistically reasonable realization of the model, we are able to compute mass profiles that not only are smoother than those obtained from the unconstrained $(\rho,\ T)$ set but also avoid the unphysical results that arise owing to the statistical fluctuations inherent in measurements. Essentially, our procedure imposes physically motivated constraints to reduce the uncertainty in the temperature and mass profiles and provides a statistic that characterizes the reliability of the mass reconstruction.

The standard procedure for extracting the gravitating mass profile of a galaxy cluster is to insert its deprojected temperature and density profiles into the hydrostatic equation (Sarazin 1988):

$$M_r = -\frac{kT}{G\mu m_p/r}\left(\frac{d\log T}{d\log r} + \frac{d\log\rho}{d\log r}\right). \qquad (2)$$

Here T and ρ are the local (baryonic) plasma temperature and density, r is the spherical radius, M_r is the total mass enclosed within r (i.e., baryons plus dark matter), and m_p and μ are the proton mass and mean particle weight, respectively. Implicit here is the assumption that the cluster is supported solely by an isotropic thermal pressure gradient, i.e., that random motions greatly exceed the bulk rotational motion, and that the magnetic field energy density is negligible in comparison with the thermal energy content of the plasma. We further assume that no recent merger event has caused a disruption in the pressure and density. Given the run of density and temperature in our binning scheme (r_i, ρ_i, and T_i, $i = 1, 2, \ldots, N$), we can calculate M_i using a simple difference equation version of equation (2):

$$M_i = -Ar_iT_i\left(\frac{t_{i+1} - t_{i-1} + d_{i+1} - d_{i-1}}{x_{i+1} - x_{i-1}}\right), \qquad (3)$$

where $A = k/G\mu m_p$, $x_i = \log_{10}(r_i)$, $d_i = \log_{10}(\rho_i)$, and $t_i = \log_{10}(T_i)$. Since the goal of this technique is to derive enclosed mass profiles, we define r as the outer radius of each shell.

Because of measurement error, a mass profile calculated in this way is not guaranteed to be physically reasonable. Even if the assumptions of spherical symmetry and hydrostatic equilibrium were valid, statistical fluctuations in the temperature measurements could result in unphysical points in our derived mass profile, for example, $dM_r/dr < 0$ or even $M_r < 0$. To deal with the nonphysical fluctuations, we proceed under the assumptions that all unphysical values in the mass profile are due to measurement uncertainty in either ρ or T. Specifically, to estimate the $(\rho,\ T)$ profiles, we impose the condition that the computed total gravitating mass profile is consistent with

$$\rho(\text{total}) = \rho(\text{baryons}) + \rho(\text{dark matter}) \geq 0 \quad \text{for } r \geq 0. \qquad (4)$$

Since $\rho(\text{total}) = (1/4\pi r^2)(dM_r/dr)$ and $1/4\pi r^2$ is positive definite, this is equivalent to the constraint

$$dM_r/dr \geq 0. \qquad (5)$$

Combining equation (5) with the boundary condition $M_r(0) \geq 0$ (e.g., allowing for the presence of an unresolved central object), we obtain a second constraint:

$$M_r(r) \geq 0. \qquad (6)$$

Perhaps the most natural way to impose these conditions would be to invoke them as a Bayesian prior in the spectral fitting procedure. Bayes's theorem (Bayes 1763; Papoulis 1984) states that the probability of model M given data set D [the posterior distribution $P(\mathsf{M}|\mathsf{D})$] is proportional to the product of the probability of that data set given the model [the likelihood function $P(\mathsf{D}|\mathsf{M})$] and the probability of the model itself [the prior knowledge function $P(\mathsf{M})$]:

$$P(\mathsf{M}|\mathsf{D}) \propto P(\mathsf{D}|\mathsf{M}) \cdot P(\mathsf{M}). \qquad (7)$$

The model prior could be chosen to enforce the constraints in equations (5) and (6). Thus, for certain combinations of model parameters T_i, ρ_i, we could choose a prior such that $P(\mathsf{M}(T_i,\ \rho_i)) = 0$.

Unfortunately, this approach is computationally impractical. The difficulty lies in the fact that the relatively simple constraints on $M(r)$ and dM/dr lead to complicated (nonlinear) constraints of the coupled parameters of the spectral models (K_i, T_i). The problem is more tractable if we instead apply the Bayesian constraints after unconstrained density and temperature profiles have been determined using the spectral deprojection method described in § 2. Thus, we compute an unconstrained mass profile from equation (3) and check it for consistency with the constraints in equations (5) and (6). These three equations yield constraints on the solutions to the difference equations that we define using Y_i and Z_i:

$$Y_i = -(t_{i+1} - t_{i-1} + d_{i+1} - d_{i-1}) \geq 0, \qquad (8)$$

$$\begin{aligned} Z_i = &- r_{i+1}T_{i+1}\left(\frac{t_{i+2} - t_i + d_{i+2} - d_i}{x_{i+2} - x_i}\right) \\ &+ r_iT_i\left(\frac{t_{i+1} - t_{i-1} + d_{i+1} - d_{i-1}}{x_{i+1} - x_{i-1}}\right) \geq 0. \end{aligned} \qquad (9)$$

Physically these constraints prohibit the pressure of the X-ray–emitting plasma from rising with radius. Using the model parameter error estimates as a guide, we allow the temperature and normalization of each model component to vary if either of these constraints is violated by the profile. Our goal is thus to find a new set of density and temperature values $\tilde{\rho}$ and $\tilde{T}$ that are as close as possible to the original profiles (“fidelity”) but obey the constraints of equations (8) and (9) (“physicality”).

At first glance this appears to be a standard problem in constrained optimization. Two features of the problem suggest than an alternate route is preferable, however. First, as mentioned above, the constraints are nonlinear functions of the independent variables, so the constrained optimization approach is very complex. Second, if we exercise control over the competing interests of fidelity and physicality, as incorporated in a penalty function, rather than simply eliminating solutions that violate equations (8) and (9), we retain the ability to modulate departures from the unconstrained profiles. For example, it may be the case that the temperature profile would need substantial alteration in order to satisfy the rigid imposition of constraint equations (8) and (9), but that a minor violation of equation (9) (say, at only one point in the mass profile) would allow a temperature profile of much greater fidelity. In this instance it is to our advantage to have the ability to set the relative weighting between the fidelity and physicality terms in the penalty function.

Implicit in our choice of a penalty function to represent the constraints is the assumption that our estimate of the location of the peak of the model probability distribution function $P_{\rm est}(\mathsf{M}(T_i,\ \rho_i))$, obtained by enforcing fidelity, is close to the actual peak of the distribution $P_0(\mathsf{M}(T_i,\ \rho_i))$, which would obtain if the Bayesian priors were utilized during the spectral deprojection. This idea is shown schematically in Figure 1. As the figure illustrates, in the absence of significant small-scale structure in $P(\mathsf{M}(T_i,\ \rho_i))$, $P_{\rm est}(\mathsf{M}(T_i,\ \rho_i))$ will indeed lie near $P_0(\mathsf{M}(T_i,\ \rho_i))$ in parameter space. We therefore cast our physicality constraints as terms in a cost function and henceforth refer to the optimized solution for M as the *constrained profile*, with the understanding that it is an estimate of the fully Bayesian solution and that the constraints do not rigorously forbid excursion into disfavored regions of parameter space. Our formulation of the cost function will reflect our subjective assessment of the relative importance of fidelity versus physicality and use standard optimization algorithms to minimize it.

We use the likelihood to characterize the fidelity term in the cost function. The probability density P at the vector $(\tilde{\rho},\ \tilde{T}) = (\tilde{\rho}_1,\ \tilde{\rho}_2,\ \ldots,\ \tilde{\rho}_N,\ \tilde{T}_1,\ \tilde{T}_2,\ \ldots,\ \tilde{T}_N)$, which is near the unconstrained profiles $(\rho,\ T)$, is given by

$$P = (2\pi)^{N/2}\prod_{i=1}^{N}\left(\sigma_{\rho_i}\sigma_{T_i}\right)^{-1}e^{-(\tilde{\rho}_i-\rho_i)^2/2\sigma^2_{\rho_i}}e^{-(\tilde{T}_i-T_i)^2/2\sigma^2_{T_i}}, \quad (10)$$

where we have assumed that the errors in ρ and T are uncorrelated and normally distributed, characterized by σ_{ρ_i} and σ_{T_i}, respectively.[4] In maximum likelihood methods, cost functions are conveniently defined as the negative log of the likelihood (e.g., von Mises 1964), so we define the fidelity penalty function Q as

$$\begin{aligned} Q &= -2\log P \\ &= N\log(2\pi) \\ &\quad + 2\sum_{i=1}^{N}\left[\log\sigma_{\rho_i} + \log\sigma_{T_i} + \frac{(\tilde{\rho}_i-\rho_i)^2}{2\sigma^2_{\rho_i}} + \frac{(\tilde{T}_i-T_i)^2}{2\sigma^2_{T_i}}\right] \\ &= Q_0 + \chi^2. \end{aligned} \quad (11)$$

[4] In principle, it would be more accurate to make use of the fact that the luminosity of a shell i, $L_i \propto \rho_i^2 T_i^{1/2}$, is tightly constrained by the observations and therefore that deviations in d and t are correlated: $\delta d = -\delta t/4$. In practice, however, this is unnecessary since the relative uncertainty in the temperature of a shell is enormous compared to the uncertainty in its density.

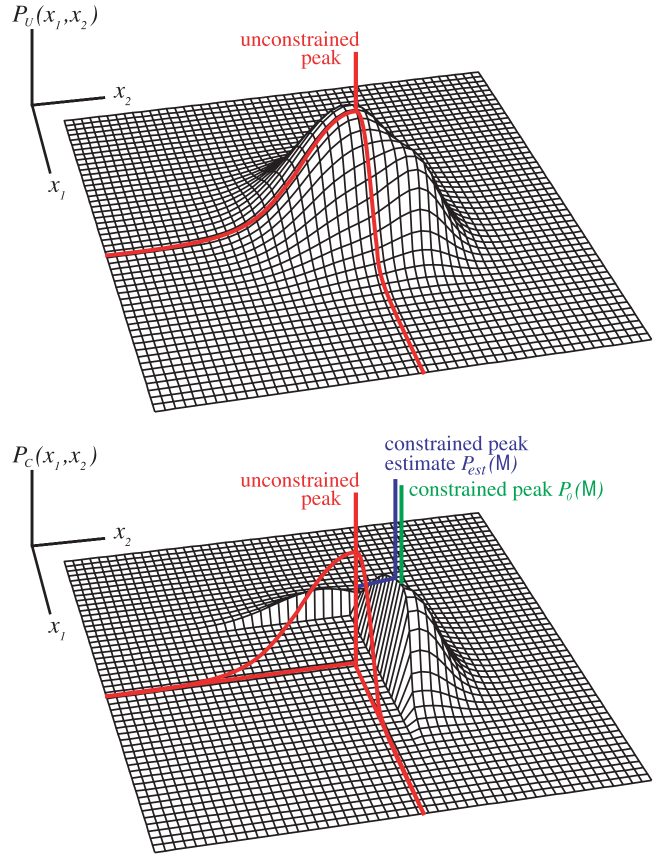

Fig. 1.—Schematic representations of the unconstrained probability distribution function (P_U; *top*) and the constrained distribution (P_C; *bottom*) for a two-dimensional parameter space. The estimated peak of the constrained distribution is shown in blue; the true peak is shown in violet. (For clarity we have not renormalized P_C after removing the physically disallowed region of parameter space.)

Here Q_0 is a constant that depends only on the measurement errors and χ^2 is the $2N$-dimensional variance:

$$\begin{aligned} \chi^2 &= \sum_{i=1}^{N}\left[\frac{(\tilde{\rho}_i-\rho_i)^2}{\sigma^2_{\rho_i}} + \frac{(\tilde{T}_i-T_i)^2}{\sigma^2_{T_i}}\right] \\ &= \chi^2_\rho + \chi^2_T. \end{aligned} \quad (12)$$

We ignore Q_0 since it will not affect the minimization.

We follow a similar procedure to derive the physicality terms in the cost function. We wish to incorporate the physicality constraints by penalizing negative values of Y_i and Z_i in equations (8) and (9) without attempting to maximize them if they are positive. Consider the penalty function

$$\begin{aligned} g(x) &= \left(\left|\frac{x}{x_0}\right| - \frac{x}{x_0}\right)^\eta + \left(\left|\frac{x}{x_0} - 1\right| + \frac{x}{x_0} - 1\right)^\eta \\ &\quad - \ln\frac{1}{2} + \ln C \quad (\eta > 1). \end{aligned} \quad (13)$$

This function has several useful properties: (1) it has a value of zero for $0 \le x \le x_0$; (2) it increases rapidly for $x < 0$ or $x > x_0$; (3) it has a continuous first derivative everywhere for $\eta \ge 2$ (or everywhere but $x = 0$ or $x = x_0$ for $1 < \eta < 2$), making it well suited for optimization routines that rely on gradient information to increase their efficiency; and (4) $h(x) = e^{-g(x)}$ is a

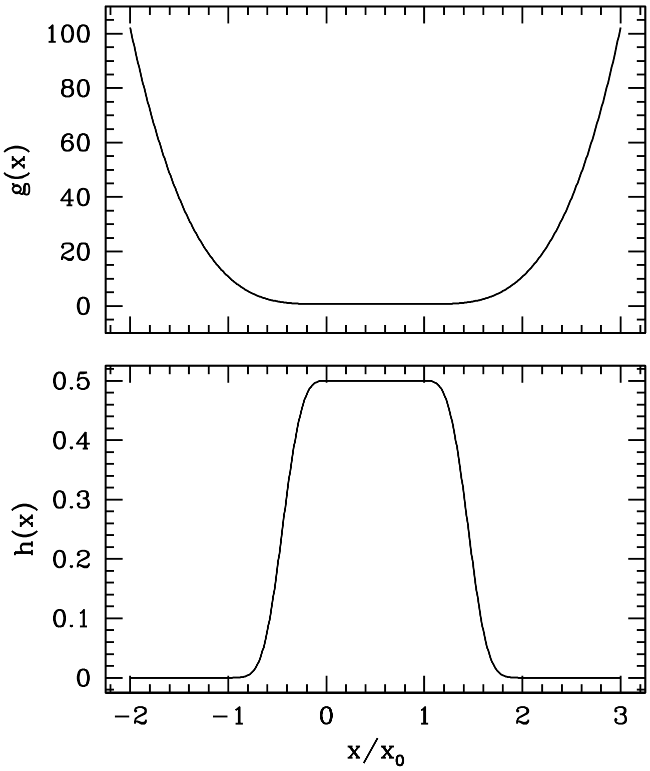

FIG. 2.—Penalty function $g(x)$ and the associated (approximate) probability distribution function $h(x) = e^{-g(x)}$ for $\eta = 7/2$. Here $h(x)$ is nearly, but not precisely, normalized; we have set $C = 1$ (see eq. [13]).

well-behaved probability distribution function. C is of order 1; for arbitrary values of $\eta > 1$ it can be determined numerically. Functions $g(x)$ and $h(x)$ are shown in Figure 2.

We employ $g(x)$ to characterize the two physicality requirements, $M_r \geq 0$ and $dM_r/dr \geq 0$, in their difference equation form (eqs. [8] and [9]). In both cases x_0 is set to a large value that we do not expect to exceed. In the former case we set it to $10^{18}\ M_\odot$; in the latter we can use the equivalent of $d \log M_r/d \log r \leq 3$, corresponding to the upper limit for a monotonically decreasing density profile. We can ignore the normalization and set $K = 1$ since we will be weighting the physicality penalties against the fidelity penalty. The complete cost function f for the minimization is thus

$$f = \sum_{i=1}^{N} [A_1 g(Y_i) + A_2 g(Z_i)] + \chi^2. \tag{14}$$

The weights A_1 and A_2 and the exponent η are chosen to determine whether the departure from the unconstrained profile is reasonable to achieve a physically acceptable solution.

Conveniently, this technique has a built-in gauge of the fidelity, the χ^2 value of the constrained profiles (although it should be noted that it is not distributed as the classical χ^2). If $\chi^2 \lesssim 1$, then we consider the new profile to be a reasonable excursion within the profile's uncertainty envelope. Moreover, we identify the constrained profiles ($\tilde{\rho}$, $\tilde{T}$, $\tilde{M}_r$) as *a closer approximation to the true profiles*, under the assumption that our model is correct. Conversely, $\chi^2 \gg 1$ indicates that the (ρ, T) profiles require excessive alteration to produce a physically sensible mass profile. This could indicate that our assumptions of spherical symmetry and/or hydrostatic equilibrium are invalid, that there is significant spatial or spectral substructure, or that our thermal emission model is inadequate. Regardless of the root cause, the accuracy of such a profile is suspect.

Nulsen & Böhringer (1995, hereafter NB95) developed a nonparametric approach to this problem that also makes use of the fact that the enclosed gravitating mass of a cluster must be monotonically increasing. With their method one obtains a series of interdependent constraints on the mass at each point in r. These constraints are translated into a set of likelihood functions that are assumed to be independent for computational ease. These likelihood functions are jointly maximized to derive the mass profile. Our method is similar in that it makes use of a likelihood function based on excursions within an uncertainty envelope. The methods differ significantly, however, in the deprojection algorithm (NB95 use onion peeling) and in the enforcement of the mass constraints. In NB95, the mass constraints are absolute; in our method they can be invoked to any degree: they can be rigidly enforced, completely ignored, or somewhere in between. The χ^2 metric of the fidelity is a powerful tool. If the physicality weight is set to an arbitrarily large value, we force the profile to monotonically increase, in effect obtaining the result of the NB95 technique. As a by-product, we obtain the plasma temperature and density profiles that are required, and the χ^2 value immediately tells us how likely it is that the temperature and density measurements conspired to achieve this condition. In cases in which rigid imposition of the constraints yields an unacceptably high χ^2 value, one can experiment with different weights and examine the resulting profiles to ascertain, for example, if the problem is due to a single outlier or is more systemic, possibly indicative of a breakdown of hydrostatic equilibrium. Thus, the flexibility of our method with regard to the mass constraints allows us to extract information about the dynamical state of the X-ray plasma and the presence of statistical anomalies in the data.

4. CHOOSING A MODEL: THE *F*-TEST AND MARKOV CHAIN MONTE CARLO SAMPLING

In the previous section we assumed that the underlying model is correct. There is some uncertainty, however, about the form the model should take. As mentioned previously, the best candidates for this type of analysis are relaxed cooling flow clusters. These systems are currently not well understood, yet the way we model the cool plasma can have a significant effect on the resulting mass profile and conclusions we may draw about the properties of dark matter particles or the evolution of large-scale structure (Arabadjis et al. 2002).

A standard procedure for choosing between a simple model M^s and a complex model M^c is to utilize the F-test (Bevington 1969). This is done by computing the F-value for the data set D,

$$F = \frac{\chi^2(\mathsf{M}^\mathsf{s}|\mathsf{D}) - \chi^2(\mathsf{M}^\mathsf{c}|\mathsf{D})}{\chi^2(\mathsf{M}^\mathsf{s}|\mathsf{D})/\nu(\mathsf{M}^\mathsf{s})}, \tag{15}$$

and comparing it to the standard F-distribution. Here $\chi^2(\mathsf{M}^\mathsf{s}|\mathsf{D})$ and $\chi^2(\mathsf{M}^\mathsf{c}|\mathsf{D})$ are the sum of the squares of the error-weighted residuals in the spectroscopic least-squares fit to the simple and complex models, respectively, and $\nu(\mathsf{M}^\mathsf{s})$ is the number of degrees of freedom in the simple model.

Unfortunately, the standard F-test is not applicable in this context. As Protassov et al. (2002) have pointed out, the standard F-test is valid only in cases in which the simple model is nested within the complex model. In the present case, the simple model lies on a boundary of the complex model. That is, the simple model is a special case of the complex model, with

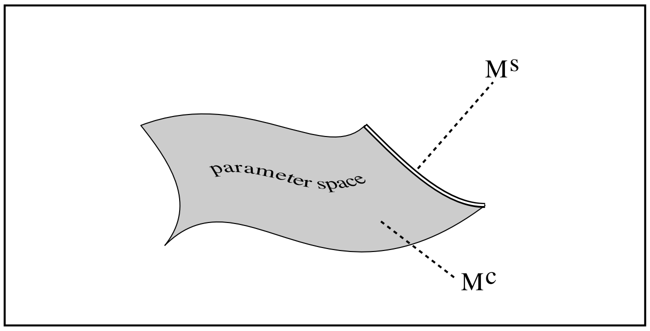

FIG. 3.—Nonnested relationship between the simple and complex models. The simple model corresponds to the complex model with one of the normalizations set to 0.

the normalization of the second emission component in the core set to zero. This means that its F-distribution may deviate significantly from the norm (see Fig. 3). Instead, we must construct an *empirical* F-distribution by sampling the probability distribution function of the simple model, simulating a data set for each sample item, and applying both models to the simulated data. Once our F-distribution has been constructed, we can judge the significance of the extra emission component based on the location within the distribution of the F-value for the data (Protassov et al. 2002).

We employ the Markov Chain Monte Carlo (MCMC) sampling technique (Neal 1993; van Dyk et al. 2001; Lewis & Bridle 2002; Hobson & McLachlan 2003) to build a large sample of data realizations from which to construct an empirical F-distribution. Let $P(\boldsymbol{x})$ represent the posterior probability distribution function of the parameters $\boldsymbol{x} = x_1, x_2, \ldots, x_N$ determined by fitting the simple model M^{S} to a real data set D_0. We can sample $P(\boldsymbol{x})$ by taking a rejection-based random walk through the parameter space. We define a transition probability $T(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})$ as the probability of moving from an initial set of parameters $\boldsymbol{x}_n$ to a new set of parameters $\boldsymbol{x}_{n+1}$. T depends on the value of the posterior distribution at the original and new parameter sets. Let $q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})$ be an arbitrary proposal distribution, that is, the probability that the new proposed parameter set is $\boldsymbol{x}_{n+1}$ given that we are presently at $\boldsymbol{x}_n$. If we accept the proposed parameter set with probability α, then

$$\alpha(\boldsymbol{x}_n, \boldsymbol{x}_{n+1}) = \frac{T(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})}{q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})}, \tag{16}$$

which takes into account the odds of actually stepping to the new location in parameter space, $q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})$, and the odds of such a transition between the two locations being accepted, $T(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})$. The acceptance probability α is calculated from

$$\alpha(\boldsymbol{x}_n, \boldsymbol{x}_{n+1}) = \min\left[1, \frac{P(\boldsymbol{x}_{n+1})q(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n)}{P(\boldsymbol{x}_n)q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})}\right]. \tag{17}$$

In our case we use a particular form of MCMC sampling called the Metroplis algorithm (see, e.g., Neal 1993), which uses a symmetric proposal distribution function:

$$q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1}) = q(\boldsymbol{x}_{n+1}, \boldsymbol{x}_n). \tag{18}$$

This prescription for wandering through parameter space constitutes a Markov Chain since each new parameter set is chosen according to a probability distribution function that depends only on the previous set of values. In the case of the Metropolis algorithm, it is straightforward to show that $P(\boldsymbol{x})$ is an invariant distribution of the Markov Chain. Using equations (16) and (17), we have

$$\begin{aligned} P(\boldsymbol{x}_n)T(\boldsymbol{x}_n, \boldsymbol{x}_{n+1}) &= P(\boldsymbol{x}_n)q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})\alpha(\boldsymbol{x}_n, \boldsymbol{x}_{n+1}) \\ &= q(\boldsymbol{x}_n, \boldsymbol{x}_{n+1})\min[P(\boldsymbol{x}_n), P(\boldsymbol{x}_{n+1})]. \end{aligned} \tag{19}$$

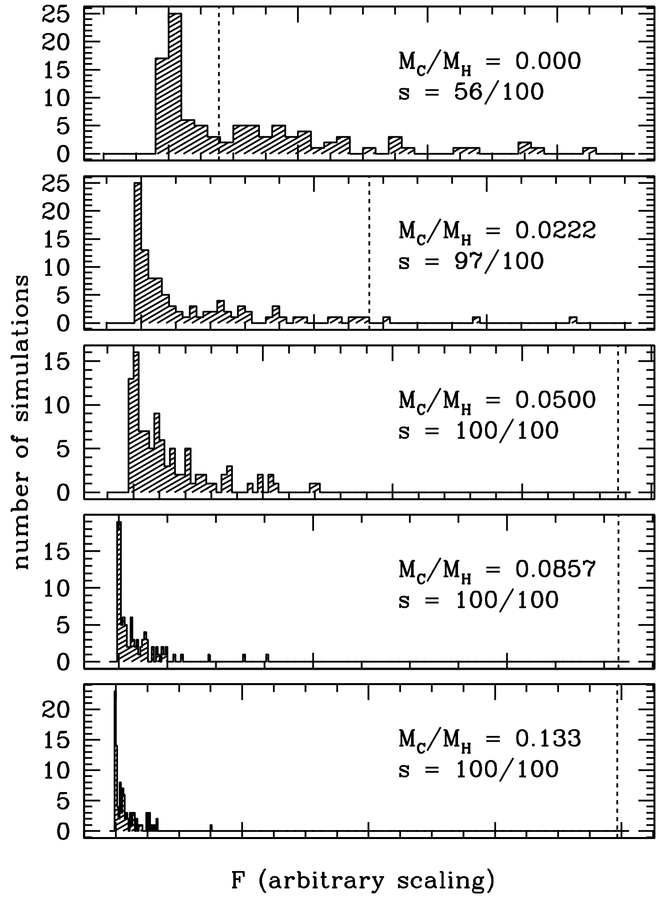

FIG. 4.—Empirical F-distributions for models M^{S} and M^{C} of five simulated data sets. The cold-to-hot plasma mass ratio, assuming that the two are in pressure equilibrium, is shown for each case, along with the significance of the presence of the cool component as determined through 100 MCMC simulations (see eq. [21]). The F-value of the original data set is indicated by a vertical dashed line.

TABLE 1

GALAXY CLUSTER SAMPLE

Cluster	z
A1689	0.181
A1795	0.0631
A1835	0.2523
A2029	0.0765
A2104	0.1554
A2204	0.1523
Hydra A	0.0522
MS 1358	0.328
MS 2137	0.313
ZW 3146	0.2906

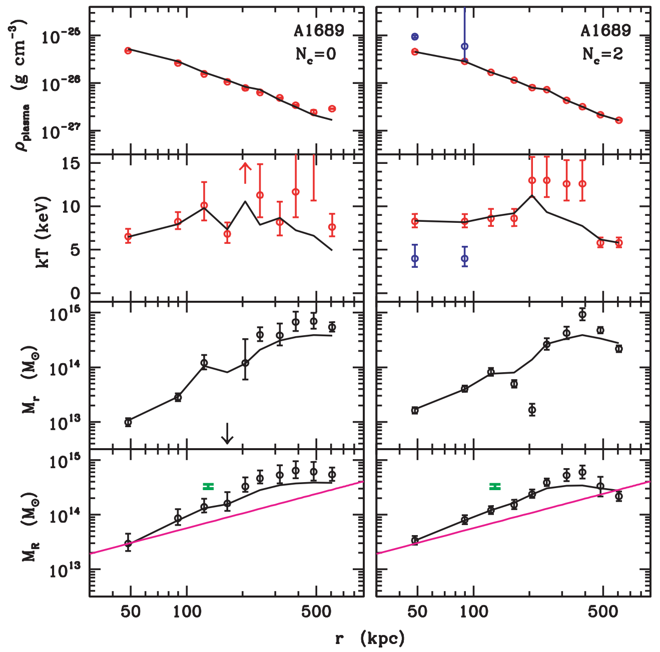

FIG. 5.—Baryon density, baryon temperature, spherically enclosed mass, and cylindrically enclosed (projected) mass profiles of A1689, for $N_c = 0$ (*left*) and 2 (*right*). The hot plasma is shown in red, the cool component in blue. Arrows represent points that lie outside the ordinate range. Reconstructions adhering to constraint eq. (6) are shown by solid black lines. Weak (*violet*) and strong (*green*) gravitational lensing measurements are shown for comparison in the bottom panels. (A solid violet line represents an isothermal sphere fit to the weak-lensing data set.) See Table 3 for lensing references.

Making use of the symmetry of q in the Metropolis algorithm, we have

$$\begin{aligned} P(\boldsymbol{x}_n)T(\boldsymbol{x}_n,\ \boldsymbol{x}_{n+1}) &= q(\boldsymbol{x}_{n+1},\ \boldsymbol{x}_n)\min[P(\boldsymbol{x}_n),\ P(\boldsymbol{x}_{n+1})] \\ &= P(\boldsymbol{x}_{n+1})q(\boldsymbol{x}_{n+1},\ \boldsymbol{x}_n)\alpha(\boldsymbol{x}_{n+1},\ \boldsymbol{x}_n), \end{aligned}$$

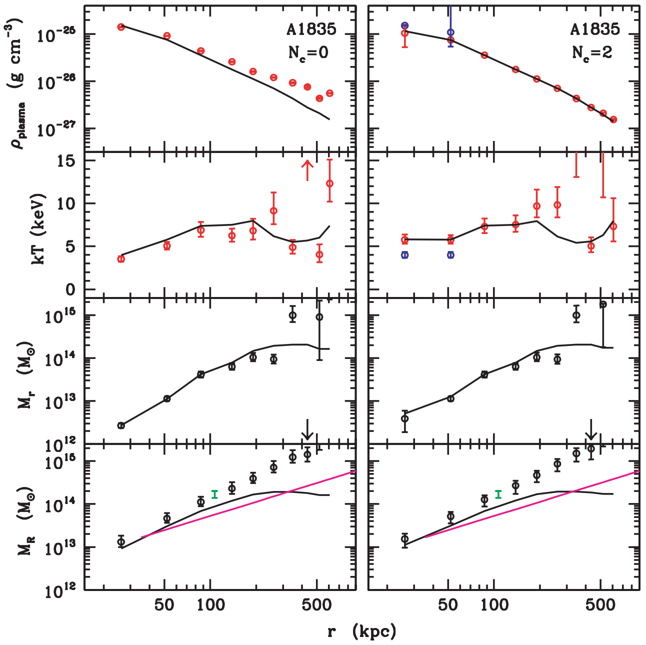

FIG. 7.—Same as Fig. 5, but for A1835.

resulting in

$$P(\boldsymbol{x}_n)T(\boldsymbol{x}_n,\ \boldsymbol{x}_{n+1}) = P(\boldsymbol{x}_{n+1})T(\boldsymbol{x}_{n+1},\ \boldsymbol{x}_n). \tag{20}$$

This statement of detailed balance demonstrates that $P(\boldsymbol{x})$ is a stationary distribution of the Markov Chain. This is necessary, although not sufficient, to ensure that we can sample $P(\boldsymbol{x})$ directly using an appropriately selected chain of Monte Carlo simulations. The other necessary condition, ergodicity, ensures that any substring of the Markov Chain will asymptotically approach $P(\boldsymbol{x})$ regardless of the initial conditions, although a

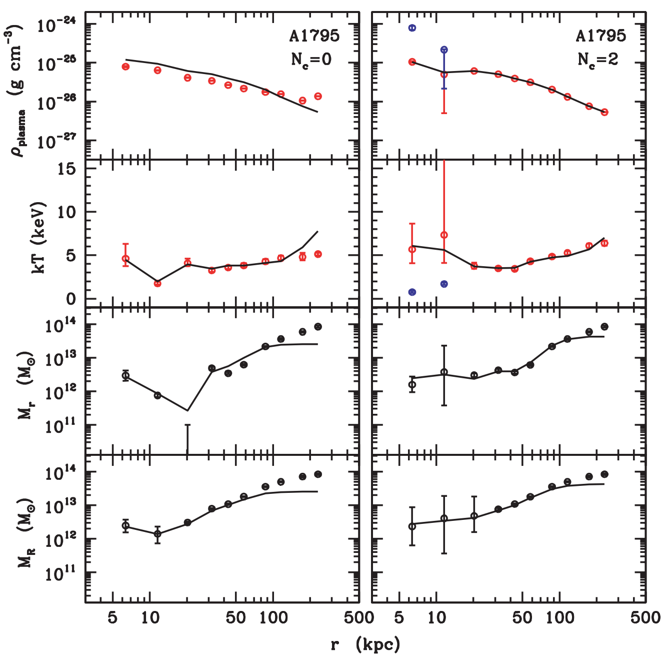

FIG. 6.—Same as Fig. 5, but for A1795.

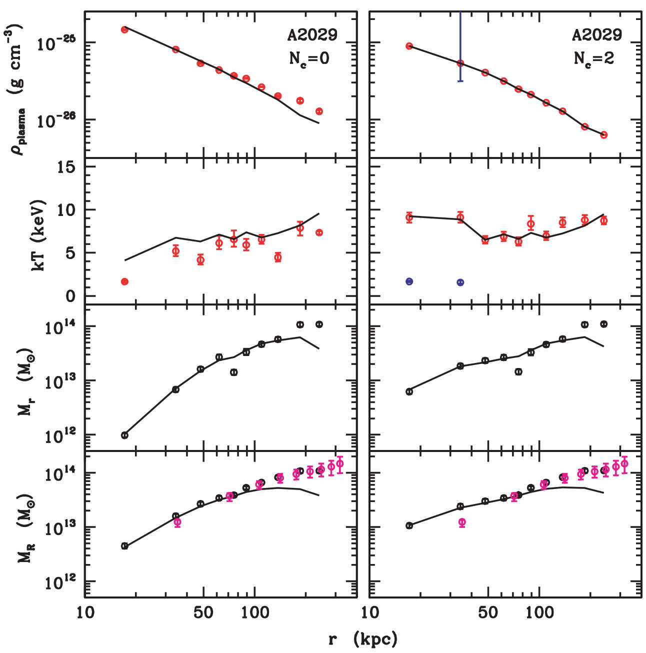

FIG. 8.—Same as Fig. 5, but for A2029.

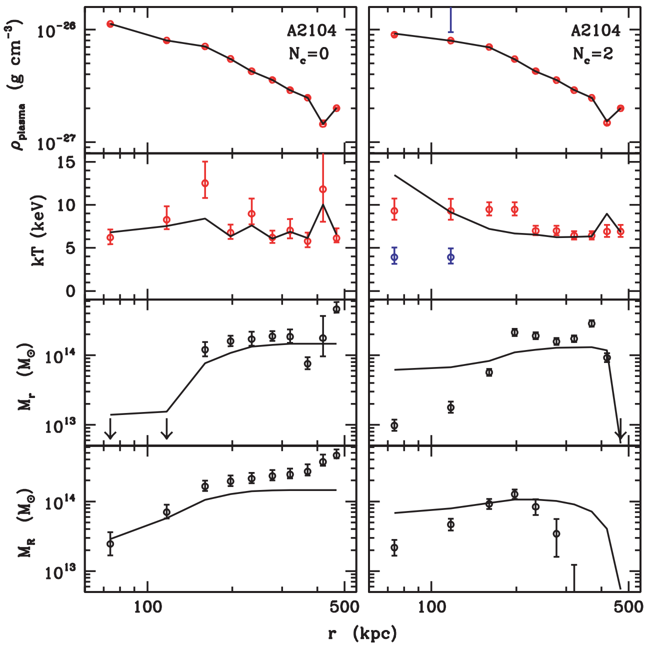

FIG. 9.—Same as Fig. 5, but for A2104. Note that there is no reconstruction solution in the $N_c = 2$ case that is consistent with constraint eqs. (5) and (6).

derivation of this property is beyond the scope of this paper. For a complete discussion see Neal (1993).

In many applications of MCMC sampling one pays special attention to the finite "burn-in" period during which the Markov Chain equilibrates. The length of the burn-in phase depends on the sensibility of the starting point and the appropriateness of the scale chosen for the proposal probability distribution step. This is not a consideration in our case because we start each MCMC sample at the (already known) peak of the probability distribution function $P(\boldsymbol{x})$.

Given $P(\boldsymbol{x})$ computed by the MCMC process, we can construct an empirical F-distribution and perform an F-test on the

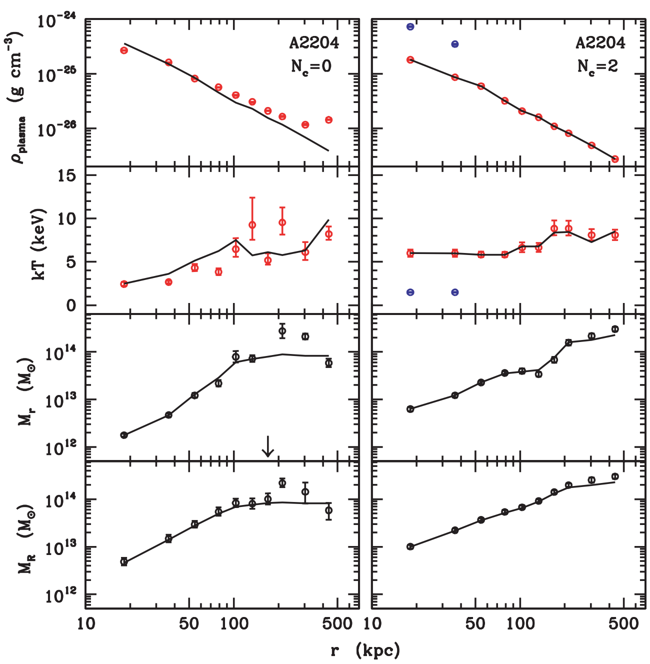

FIG. 10.—Same as Fig. 5, but for A2204.

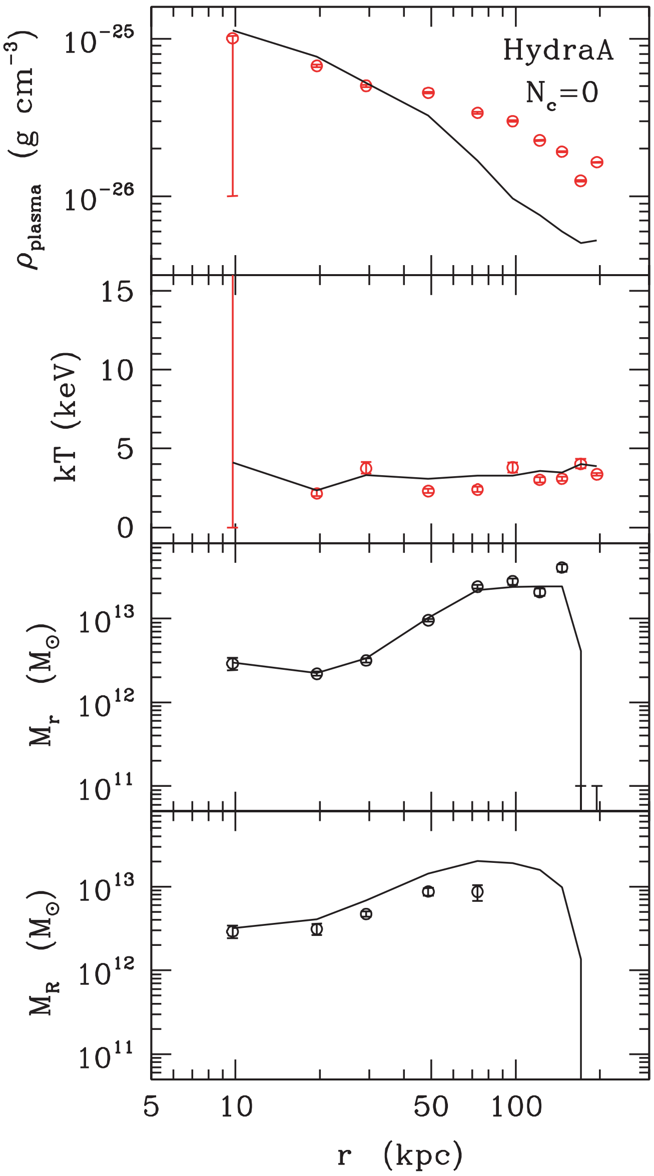

FIG. 11.—Same as Fig. 5, but for Hydra A. Note that there is no reconstruction solution in the $N_c = 1$ case that is consistent with constraint eqs. (5) and (6) and that the Hydra A data set did not admit $N_c = 2$ emission models.

significance of a second emission component in the core. The entire procedure is as follows:

1. Model the real data set D_0 with M^S; call the best-fit parameters $\boldsymbol{x}_0^\mathsf{S}$.
2. Use XSPEC to calculate $P(\mathsf{D}_0|\boldsymbol{x})$ (i.e., the likelihood).
3. Use Bayes's theorem to calculate $P(\boldsymbol{x}|\mathsf{D}_0)$ (see eq. [7]). We discard all unphysical excursions in parameter space, i.e., where $T < 0$ or $\rho < 0$. (In practice, we discard at the level of the model normalization, not ρ, but ρ is simply a function of the normalization and the binning geometry.)
4. Create a large sample of model parameters $\boldsymbol{x}_i^\mathsf{S}$ using $P(\boldsymbol{x}|\mathsf{D}_0)$ and the Metropolis algorithm form of the MCMC technique. For each $\boldsymbol{x}_i^\mathsf{S}$, compute a fake data set D_i, including instrumental effects of the *Chandra* telescope and detectors, as well as counting statistics.
5. Model each D_i using both M^S and M^C.
6. For each pair of models tabulate its F-value given by equation (15).

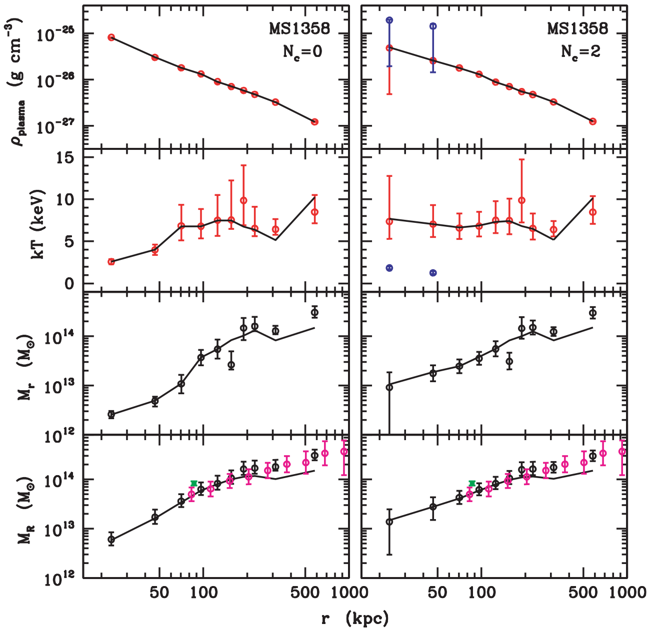

FIG. 12.—Same as Fig. 5, but for MS 1358.

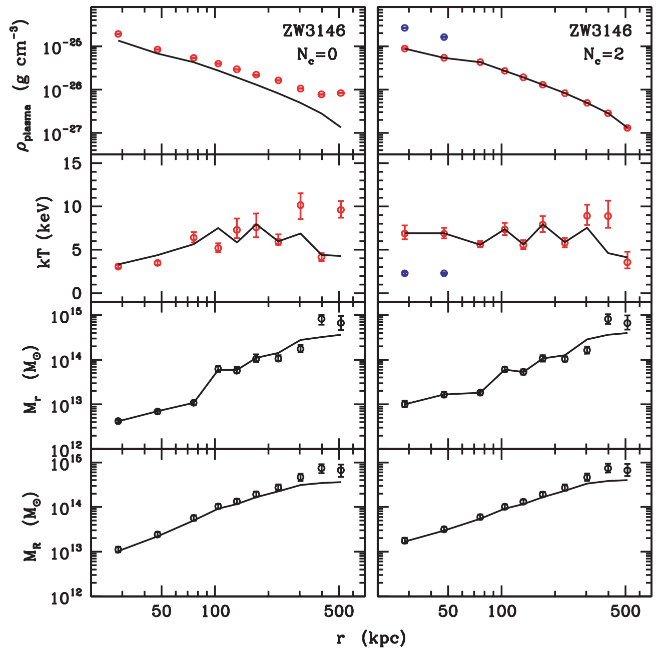

FIG. 14.—Same as Fig. 5, but for ZW 3146.

7. Bin up the set of F-values, creating an unnormalized histogram, and superimpose the F-value of the original data.

In practice, this recipe is computationally intensive, not because of any features of the MCMC sampling per se, but because each of the faked spectra must be modeled twice. For a sample size of 1000 simulations, XSPEC must simulate 1000 spectra and calculate 2002 sets of best-fit values for the model parameters (including the original data). This fact leads us to simplify the method. First, in order to reduce the modeling time, we have adopted a simplified core-halo geometry. In this scheme the "core" is represented by a single shell (in this case a sphere), while the halo is represented by another shell. Thus, M^{S} contains four parameters, the temperature and density of each of the two shells, while M^{C} contains six, the additional two parameters representing the temperature and density of a second cospatial emission component in the core. This simplification also greatly improves the numerical stability of the fitting procedure. The algorithm (steps 1–6 above) is implemented in a Tcl script run within XSPEC.

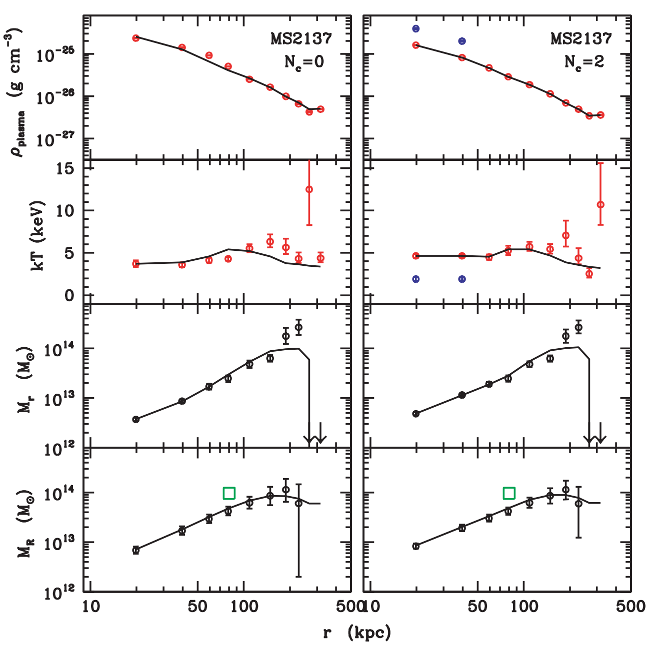

FIG. 13.—Same as Fig. 5, but for MS 2137.

Once we have completed step 7, we can distinguish between the models. The location of the F-value of the data within this empirical F-distribution contains information regarding the relative merit of M^{S} and M^{C}. We define the significance S of the distribution as

$$S = \frac{\int_0^{F_{\rm data}} N(F)\,dF}{\int_0^{\infty} N(F)\,dF}. \tag{21}$$

The significance $S = 1 - P_f$, where P_f is the probability that the simple model constitutes the better description and that the F-value of the data is this large strictly by chance. Thus, for a one-parameter model, $S = 0.68$, 0.90, and 0.99 may be interpreted as 1, 2, and 3 σ detections of the additional component. We checked the sensitivity of this method by applying it to five simulated data sets with known mixtures of hot ($T = 5.0$ keV) and cold ($T = 1.0$ keV) X-ray plasmas in pressure equilibrium. We use mass ratios of $M_{\rm C}/M_{\rm H} = 0.000$, 0.0222, 0.0500, 0.0857, and 0.133 and run 100 MCMC simulations for each. Figure 4 shows an empirical F-distribution for each of these cases. The F-value of the data is shown by a dashed vertical line. For a multiphase plasma of which only 2.2% is in the cold component, the detection of the multiphase plasma is better than 2 σ. At 5% and greater, the detection is statistically highly significant.

5. APPLICATION TO *CHANDRA* CLUSTERS

We illustrate these techniques using X-ray observations of a sample of bright, apparently relaxed galaxy clusters (see Table 1). Each of these clusters (except A1689) contains a significant amount of cooler plasma in its core and is known as

TABLE 2

Fidelity Measures of Baryon Density (χ^2_ρ) and Temperature (χ^2_T) of Each Constrained Mass Profile (see eq. [12]) and Multiphase Core Plasma Significance S

Cluster	χ^2_ρ ($N_c = 0$)	χ^2_T ($N_c = 0$)	χ^2_ρ ($N_c = 2$)	χ^2_T ($N_c = 2$)	S
A1689	0.00150	0.467	0.00825	1.14	0.342
A1795	0.0777	7.45	0.0156	1.29	0.823
A1835	0.00229	1.13	1.188	1.41	0.483
A2029	0.00774	1.17	0.00773	1.20	0.998
A2104	0.00251	0.624	0.190	4.51	0.661
A2204	0.000994	0.759	0.00947	0.284	0.999
Hydra A	0.197	2.43	...	...	...
MS 1358	0.00663	0.381	0.00454	0.356	0.987
MS 2137	0.0101	1.08	0.0110	1.38	0.271
ZW 3146	0.0293	1.25	0.0150	0.913	0.989

Note.—Hydra A data did not admit an $N_c = 2$ emission model.

a classical "cooling flow cluster" since the core radiative cooling time is shorter than the age of the cluster. We prepared each archived data set as described in Arabadjis et al. (2002) and modeled the emission using the two models described in § 2. In the first model, each shell contains isothermal plasma ($N_c = 0$), while in the other, the central two shells are also allowed a second (cooler) emission component ($N_c = 2$), and the best-fit parameter values are obtained iteratively using XSPEC. Figures 5–14 show baryon density, baryon temperature, enclosed spherical gravitating mass, and enclosed cylindrical gravitating mass profiles for each cluster in the sample. The left panels show $N_c = 0$ models; the right panels show $N_c = 2$ models. The unconstrained profiles are shown as data points, with red (blue) symbols representing the hot (cool) plasma components. (Data points that lie outside the ordinate range are indicated by arrows.) Constrained profiles are shown by solid lines; the density and temperature contributions to the χ^2 value of the constrained solutions are listed in Table 2. The last column in that table lists the MCMC significance S of the presence of multiphase gas in each cluster core as determined from the MCMC-derived empirical F-distributions shown in Figure 15. For the subset of five clusters that are arguably the most relaxed and are best described by an NFW profile (A1689, A1835, A2029, MS 1358, and MS 2137; see AB04), we overlay the projected mass profiles with weak- and/or strong-lensing measurements from the literature (see Table 3 for references). Weak-lensing mass profiles or isothermal sphere fits to weak-lensing data are shown in violet; strong-lensing measurements are shown

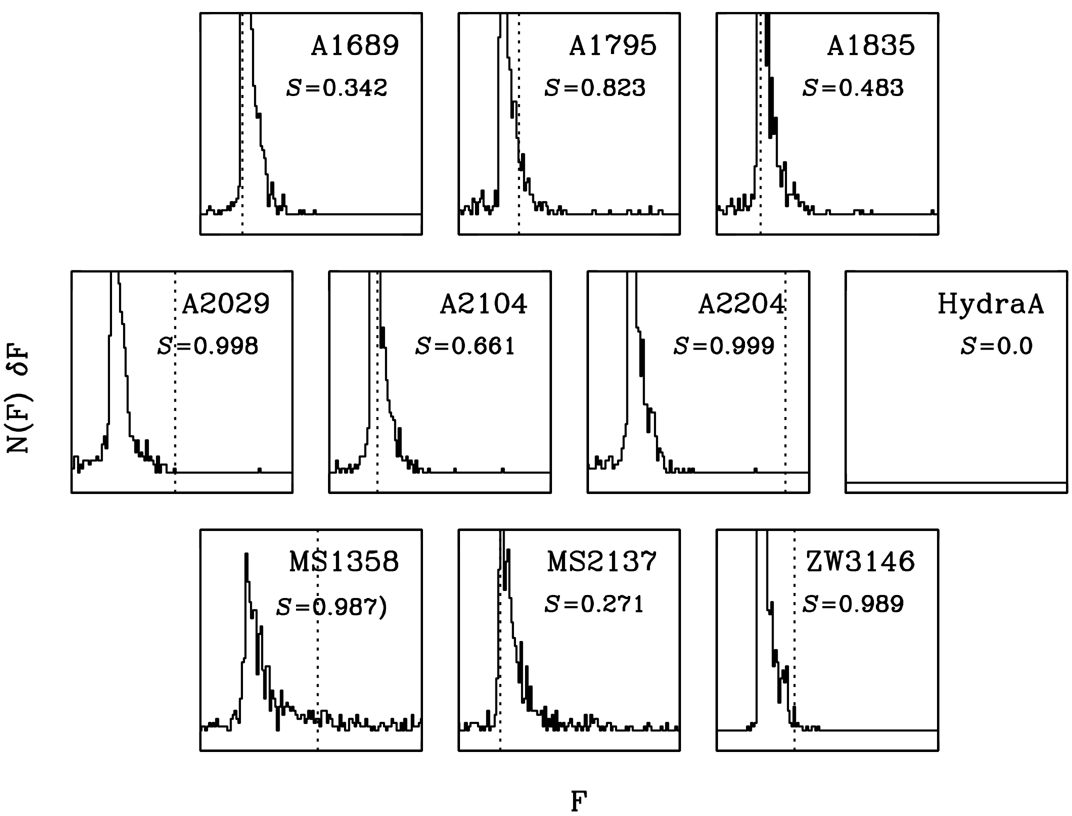

Fig. 15.—Empirical F-distributions and the MCMC significance S of a second cospatial core plasma component for each cluster in the sample. The F-value of the original *Chandra* data set is denoted by a vertical dashed line. (Note that Hydra A data did not admit an $N_c = 2$ emission model.)

TABLE 3
Lensing Comparisons for the Very Relaxed Cluster Subset

Cluster	Weak Lensing	Strong Lensing
A1689	King et al. (2002)	Wu (2000)
A1835	Clowe & Schneider (2002)	Allen (1998)
A2029	Menard et al. (2003)	...
MS 1358	Hoekstra et al. (1998)	Franx et al. (1997); Allen (1998)
MS 2137	...	Sand et al. (2002)

in green. We adopt $H_0 = 67$ km s^{-1} Mpc^{-1}, $\Omega_\Lambda = 0.7$, and $\Omega_m = 0.3$.

A detailed analysis of these profiles and their consequences for cosmology and dark matter candidates is in preparation (AB04), so we make only a few brief comments here. Of the 10 clusters presented here, only Hydra A did not admit a second emission component in the core: attempts to add one resulted in the temperature of the second component being set to the temperature of the first in the fitting procedure. This is consistent with previous *Chandra* (David et al. 2001) and *XMM-Newton* (Kaastra et al. 2004) results. Four of the clusters show evidence for multiphase core plasma at the 99% significance level: A2029, A2204, MS 1358, and ZW 3146 (see Table 2). This is consistent with the study of Kaastra et al. (2004), which finds evidence for multiphase plasma in many clusters. (Note that, with 1000 MCMC simulations per cluster, the precision of the significance estimate is $\sim$3%.) For comparison, a cluster that contained no second plasma component would, on the average, show an MCMC significance of order 0.5, or 50%. These clusters show a greater difference in their core masses between the uni- and multiphase models. This phenomenon is illustrated for two clusters, MS 1358 and MS 2137, in Figures 16 and 17, respectively. The mass of MS 1358 in the multiphase model is about a factor of 2 larger than in the uniphase mass, and its MCMC significance $S = 0.987$. MS 2137, on the other hand, with an MCMC significance of 0.271, shows very little difference between the two mass models. This is perhaps unsurprising, as one would expect that a significant amount of cospatial cool plasma in its core would not only display a clear observational signature but would also affect the equilibrium configuration of the plasma. MS 2137 and ZW 3146 also provide an interesting contrast. In the pre-*Chandra*/*XMM-Newton* era, both clusters were reported to harbor cooling flows with mass deposition rates in excess of 1000 $M_\odot$ yr^{-1} (Allen 2000), yet they show remarkably disparate evidence for multiphase plasma in their cores [S(MS 2137) = 0.271; S(ZW 3146) = 0.989]. This suggests that there may be more than one mechanism at work responsible for the presence of 1 keV plasma at the center of galaxy clusters.

For each constrained reconstruction we use $\eta = 2.5$ and physicality weights $A_1 = A_2 = A_{12}$ originally set to 1×10^{-6}. While this often will not rigidly enforce constraint equation (5), it is usually sufficient to enforce equation (6). In those cases in which not even equation (6) is satisfied, we increased A_{12} by factors of 10 until the resulting profile was nonnegative everywhere if possible. This is the origin of the unacceptably high values of the χ^2 fidelity measures for the $N_c = 2$ model of A2104 and the $N_c = 1$ model of Hydra A. Obviously, the fidelity of a constrained profile tends to be greater when the unconstrained profile shows only one or two significant outliers. In some cases (e.g., A1795) a very small adjustment to the temperature profile produces a large change in the derived mass. This is due to the competition between the derivatives in equation (2): if a statistical fluctuation in the temperature measurement is large enough (and positive), it will swamp the surface brightness decrement at that radius, resulting in a very low or even negative mass. In cases in which $d \log T/d \log r \gtrsim -d \log \rho/d \log r$, a relatively minor adjustment in the temperature can remove an unphysical point from the mass profile.

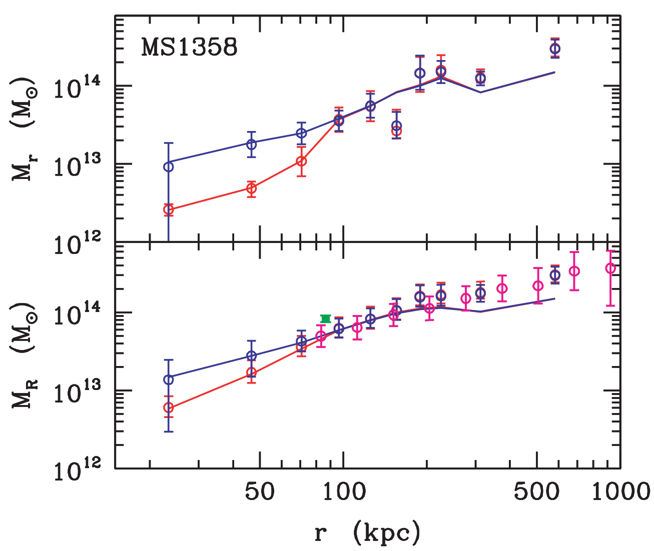

Fig. 16.—Mass profiles for models $N_c = 0$ (*red*) and 2 (*blue*) of MS 1358, overlaid for comparison. Weak (*violet*) and strong (*green*) gravitational lensing measurements are shown for comparison.

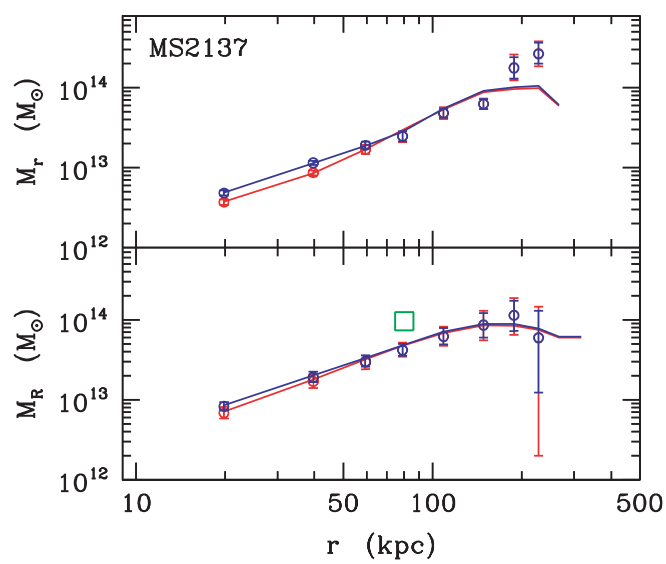

Fig. 17.—Same as Fig. 16, but for MS 2137.

The five very relaxed clusters generally show better agreement with weak-lensing mass measurements than they do with strong lensing. The reprojected profile A1689 shows the least agreement with the weak lensing, differing by up to a factor of 2 at some radii, although the lensing profile is a singular isothermal sphere (IS) fit to weak-lensing measurements (King et al. 2002). The strong-lensing measurement of Wu (2000), however, exceeds our profile again by a factor of 2. The error bars are derived by assuming two values of the (unknown) strong-lensing arc redshift (0.8 and 2.0); the discrepancy is thus unlikely to be due to incorrect source redshift. Our profile is also systematically in excess of the IS fit to weak-lensing observations of A1835 (Clowe & Schneider 2002), although it is consistent with the strong-lensing point of Allen (1998). A2029 and MS 1358 show remarkable agreement with weak-lensing data (Menard et al. 2003; Hoekstra et al. 1998), although the strong-lensing measurement in MS 1358 (Franx et al. 1997; Allen 1998) is moderately discrepant, as is the strong-lensing measurement of MS 2137 (Sand et al. 2002). These issues will be addressed in greater detail in AB04.

6. SUMMARY

We have presented a technique for calculating the dark matter profile of a spherical, relaxed galaxy cluster. We have formulated a technique for coping with statistical uncertainty in the measurement of the cluster plasma temperature. We have also described a method for determining whether there is a statistically significant presence of multiphase plasma in the galaxy cluster core. We have applied these tools to a sample of relaxed galaxy clusters observed with *Chandra* and find that 4/10 require a multiphase treatment of their core plasma. Our masses are in broad agreement with weak-lensing studies, although they are often exceeded by those derived from strong-lensing models.

J. S. A. would like to thank Steve Allen, Aaron Lewis, Jimmy Irwin, and Renato Dupke for many enlightening discussions. The authors would like to thank the anonymous referee for comments that have improved this paper. This work was supported by SAO grant AR3-4016X.

REFERENCES

Alam, S. M. K., Bullock, J. S., & Weinberg, D. H. 2002, ApJ, 572, 34
Allen, S. W. 1998, MNRAS, 296, 392
———. 2000, MNRAS, 315, 269
Allen, S. W., & Fabian, A. C. 1997, MNRAS, 286, 583
Allen, S. W., Fabian, A. C., Johnstone, R. M., Arnaud, K. A., & Nulsen, P. E. J. 2001, MNRAS, 322, 589
Arabadjis, J. S., Bautz, M. W., & Garmire, G. P. 2002, ApJ, 572, 66
Arnaud, K. A. 1996, in ASP Conf. Ser. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes (San Francisco: ASP), 17
Balogh, M. L., Babul, A., & Patton, D. R. 1999, MNRAS, 307, 463
Bayes, T. 1763, Philos. Trans. R. Soc. London, A53, 370
Bevington, P. R. 1969, Data Reduction and Error Analysis for the Physical Sciences (New York: McGraw-Hill)
Borriello, A., Salucci, P., & Danese, L. 2003, MNRAS, 341, 1109
Broadhurst, T., Huang, X., Frye, B., & Ellis, R. 2000, ApJ, 534, L15
Burkert, A. 1995, ApJ, 447, L25
Clowe, D. I., & Schneider, P. 2002, A&A, 395, 385
Dalcanton, J. J., & Bernstein, R. A. 2000, in ASP Conf. Ser. 197, Dynamics of Galaxies: From the Early Universe to the Present, ed. F. Combes, G. A. Mamon, & V. Charmandaris (San Francisco: ASP), 161
David, L. P., Nulsen, P. E. J., McNamara, B. R., Forman, W., Jones, C., Ponman, T., Robertson, B., & Wise, M. 2001, ApJ, 557, 546
Fabian, A. C., Hu, E. M., Cowie, L. L., & Grindlay, J. 1981, ApJ, 248, 47
Firmani, C., D'Onghia, E., Avila-Reese, V., Chincarini, G., & Hernández, X. 2000, MNRAS, 315, L29
Franx, M., Illingworth, G. D., Kelson, D. D., van Dokkum, P. G., & Tran, K.-V. 1997, ApJ, 486, L75
Frenk, C. S. 2002, Philos. Trans. R. Soc. London, A300, 1277
Fukushige, T., & Makino, J. 2001, ApJ, 557, 533
Goodman, J. 2000, NewA, 5, 103
Hennawi, J. F., & Ostriker, J. P. 2002, ApJ, 572, 41
Hicks, A. K., Wise, M. W., Houck, J. C., & Canizares, C. R. 2002, ApJ, 580, 763
Hobson, M. P., & McLachlan, C. 2003, MNRAS, 338, 765
Hoekstra, H., Franx, M., & Kuijken, K. 1998, ApJ, 504, 636
Hogan, C. J., & Dalcanton, J. J. 2000, Phys. Rev. D, 62, 063511
Hu, W., & Peebles, P. J. E. 2000, ApJ, 528, L61
Jing, Y. P., & Suto, Y. 2002, ApJ, 574, 538
Kaastra, J. S. 1992, An X-Ray Spectral Code for Optically Thin Plasmas, Internal SRON-Leiden Rep., Ver. 2.0
Kaastra, J. S., et al. 2004, A&A, 413, 415
Kaplinghat, M., Knox, L., & Turner, M. S. 2000, Phys. Rev. Lett., 85, 3335
Kauffman, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201
King, L. J., Clowe, D. I., & Schneider, P. 2002, A&A, 383, 118
Klypin, A. A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82
Lahav, O., et al. 2002, MNRAS, 333, 961
Lewis, A., & Bridle, S. 2002, Phys. Rev. D, 66, 103511
Lewis, A. D., Buote, D. A., & Stocke, J. T. 2003, ApJ, 586, 135
Lewis, A. D., Stocke, J. T., & Buote, D. A. 2002, ApJ, 573, L13
Liedahl, D. A., Osterheld, A. L., & Goldstein, W. H. 1995, ApJ, 438, L115
Lloyd-Davies, E. J., Ponman, T. J., & Cannon, D. B. 2000, MNRAS, 315, 689
Loewenstein, M. 2000, ApJ, 532, 17
McGaugh, S. S., & de Blok, W. J. G. 1998, ApJ, 499, 41
Menard, B., Erben, T., & Mellier, Y. 2003, in ASP Conf. Ser. 301, Matter and Energy in Clusters of Galaxies, ed. S. Bowyer & C.-Y. Hwang (San Francisco: ASP), 537
Mewe, R., Gronenschild, E. H. B. M., & van den Oord, G. H. J. 1985, A&AS, 62, 197
Mewe, R., Lemen, J. R., & van den Oord, G. H. J. 1986, A&AS, 65, 511
Mohapatra, R. N., Nussinov, S., & Teplitz, V. L. 2002, Phys. Rev. D, 66, 063002
Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., & Tozzi, P. 1999a, ApJ, 524, L19
Moore, B., Quinn, T., Governato, F., Stadel, J., & Lake, G. 1999b, MNRAS, 310, 1147
Mushotzky, R., Figueroa-Feliciano, E., Loewenstein, M., & Snowden, S. L. 2003, preprint (astro-ph/0302267)
Mushotzky, R. F., & Szymkowiak, A. E. 1988, in Cooling Flows in Clusters and Galaxies (Dordrecht: Kluwer), 53
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
———. 1997, ApJ, 490, 493
Navarro, J. F., & Steinmetz, M. 2000, ApJ, 528, 607
Neal, R. M. 1993, Tech. Rep. CRG-TR-93-1, Univ. Toronto, ftp://ftp.cs.utoronto.ca/pub/radford/review.ps
Nulsen, P. E. J., & Böhringer, H. 1995, MNRAS, 274, 1093 (NB95)
Papoulis, A. 1984, Probability, Random Variables, and Stochastic Processes (New York: McGraw-Hill)
Peacock, J. A., et al. 2001, Nature, 410, 169
Peterson, J. R., Kahn, S. M., Paerels, F. B. S., Kaastra, J. S., Tamura, T., Bleeker, J. A. M., Ferrigno, C., & Jernigan, J. G. 2003, ApJ, 590, 207
Pizzolato, F., Molendi, S., Ghizzardi, S., & de Grandi, S. 2003, ApJ, 592, 62
Ponman, T. J., Cannon, T. J., & Navarro, J. F. 1999, Nature, 397, 135
Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2002, ApJ, 571, 545
Sand, D. J., Treu, T., & Ellis, R. S. 2002, ApJ, 574, L129
Sarazin, C. L. 1988, X-Ray Emission from Clusters of Galaxies (Cambridge: Cambridge Univ. Press)
Shapiro, P. R., & Iliev, I. T. 2000, ApJ, 542, L1
Smail, I., Ellis, R., Ritchett, M. J., & Edge, A. C. 1995, MNRAS, 273, 277
Spergel, D. N., & Steinhardt, P. J. 2000, Phys. Rev. Lett., 84, 3760
Sun, M., Jones, C., Murray, S. S., Allen, S. W., Fabian, A. C., & Edge, A. C. 2003, ApJ, 587, 619
Swaters, R. A., Madore, B. F., & Trewhella, M. 2000, ApJ, 531, L107
Swaters, R. A., Madore, B. F., van den Bosch, F. C., & Balcells, M. 2003, ApJ, 583, 732
Tyson, J. A., Kochanski, G. P., & Dell'Antonio, I. P. 1998, ApJ, 498, L107
van Dyk, D. A., Connors, A., Kashyap, V. L., & Siemiginowska, A. 2001, ApJ, 548, 224
von Mises, R. 1964, Mathematical Theory of Probability and Statistics (New York: Academic)
Weisskopf, M. C., Brinkman, B., Canizares, C., Garmire, G., Murray, S., & Van Speybroeck, L. P. 2002, PASP, 114, 1
Wu, K. K. S., Fabian, A. C., & Nulsen, P. E. J. 2000, MNRAS, 318, 889
Wu, X.-P. 2000, MNRAS, 316, 299
Zentner, A. R., & Bullock, J. S. 2002, Phys. Rev. D, 66, 043003