

Parameter Estimation of Selected ACT Clusters from X-Ray and Radio Observations

Jason Pollack
Advisor: Lyman Page

May 3, 2011

Abstract

Hot gas in the halos of galaxy clusters is observed in X-rays from thermal bremsstrahlung emission and in radio waves through inverse Compton scattering (the Sunyaev-Zel'dovich effect). Because the observed fluxes from the two mechanisms have different dependences on cluster parameters, observations of an individual cluster via both effects can be used to constrain properties including its gas fraction, temperature, pressure, and mass. In this paper, we investigate the properties of selected clusters observed by the ACT telescope using the Sunyaev-Zel'dovich effect. Building on recent literature, we perform a maximum likelihood analysis to extract their underlying cluster parameters, including X-ray data where it is available. Although each individual dataset exhibits strong degeneracies, we find that the combination of both radio and X-ray data is able to precisely pinpoint a unique allowed region of parameter space.

1 Introduction

Galaxy clusters are the largest gravitationally bound structures in the universe. Precise measurement of their properties at a given redshift, such as their abundance or mass function, has the potential to constrain cosmological parameters, most importantly the dark energy equation of state w [17, 20]. Calculation of such quantities requires both sensitive observation of the clusters and theoretical models for deriving the desired quantities from astrophysical observables.

Clusters can be observed in multiple different wavelengths. In the radio, they are detected indirectly via the *Sunyaev-Zel'dovich (SZ) effect* [10], the distortion of the cosmic microwave background (CMB) due to the inverse Compton scattering of photons off of the hot ($\sim 10^8$ K) gas of the intracluster medium (ICM). In the visible, clusters can be inferred by observing overdensities of galaxies in a given region of the sky. Finally, X-ray observations detect the thermal bremsstrahlung radiation directly emitted by the hot ICM. Each of these effects comes from a different physical process, meaning that each observable has a different dependence on the underlying parameters of the cluster, such as gas

fraction, mass, or temperature. Ideally, then, the most precise measurements of these parameters can be obtained by observing a given cluster in as many different wavelength regimes as possible.

Sometimes, however, it is not possible to obtain measurements in multiple regimes simultaneously. In particular, as will be discussed in section 3.1 below, the SZ effect is redshift-independent but X-ray and optical methods are not. Alternatively, one might simply want to get information about a cluster before follow-up observations have been performed. In each case, it is desirable to extract as much information as possible about a cluster from a given set of measurements. Cluster properties cannot be exactly determined from realistic data, but they can be constrained using a *maximum likelihood analysis* in which observations are compared to simulations using a particular theoretical model.

This paper applies the techniques of maximum likelihood analysis to estimate the parameters of a set of clusters observed via the SZ effect by the Atacama Cosmology Telescope (ACT), adapting a methodology used to analyze cluster data from the Arcminute Microkelvin Imager [14]. From direct temperature measurements we extract radial profiles of the Compton y -parameter (defined in Section 3.1 below), which quantifies the flux of photons scattering off of the intracluster medium. We then employ the isothermal β model, a simplified theoretical model of the physics of galaxy clusters, to simulate the y -parameter given a particular set of cluster parameters, and perform a maximum likelihood analysis using the simulations and data to extract marginalized distributions of these parameters for each individual cluster. Finally, we extend the model to incorporate measurements of X-ray surface brightness where available, showing how the addition of this data breaks degeneracies present when only radio data is used.

The remainder of this paper is organized as follows. Section 2 provides a brief introduction to the relevant physics of galaxy clusters, and introduces the β model in addition to more realistic empirical models. Section 3 discusses the observables that can be directly derived from observations, and relates them to the underlying cluster parameters. Section 4 introduces the dataset of ACT clusters and the techniques used to derive the y -parameter from them. Section 5 illustrates how to obtain cluster properties from observables using a maximum likelihood analysis. Section 6 gives a worked example of this analysis as applied to one individual ACT cluster (ACT-CL J0516-5430), incorporating available X-ray observations. Finally, we conclude in Section 7.

2 The Physics of Galaxy Clusters

2.1 Cluster Formation

The prevailing cosmological model today is the Λ CDM model [21, 22], which, in addition to baryonic matter and radiation, incorporates nonelectromagnetically interacting and nonrelativistic *cold dark matter* as well as a cosmological constant Λ driving acceleration of the universe at late times. Λ CDM implies

“bottom-up” structure formation, in which structure forms hierarchically starting at the smallest scales. The early universe was radiation-dominated, and at high enough temperatures both ordinary and dark matter were thermalized by repeated interactions with photons and with each other. Once the temperature fell below the weak interaction scale, interactions between dark matter and the rest of the universe became negligible, allowing the dark matter to cool. Because the universe was not perfectly homogeneous, either due to the magnification of early quantum fluctuations via inflation or through some other mechanism, the dark matter fell into gravitational potential wells, amplifying the primordial homogeneities and forming the first structures.

When the temperature cooled sufficiently that the rate of electron-photon interactions fell below the expansion rate of the universe, baryonic matter ceased to become thermalized, an event known as *decoupling*. The onset of decoupling was the last time that photons scattered off of matter in non-negligible quantities, so CMB experiments are observing the so-called *surface of last scattering*. The decoupling redshift and time have been measured precisely by WMAP [1]:

$$z_* = 1090.8 \pm 0.9, t_* = (379 \pm 5) \cdot 10^3 \text{ yr.} \quad (1)$$

After decoupling, baryonic matter falls into the gravitational potential wells formed by dark matter, and structure formation quickly becomes nonlinear as the wells grow rapidly. To a first approximation, the maximum possible size of structure is limited by the horizon size of the universe, as well as competing processes which slow the cooling of matter to the point where gravity becomes dominant. Galaxy clusters were the largest and most recent structures to form; because the universe is now dominated by dark energy instead of matter, no larger structures are forming today.

Dark matter makes up the majority of the material in clusters. Because dark matter has had more time than baryonic matter to cool and clump, we expect the *gas fraction*, the ratio of baryons to total matter in a cluster¹, to be somewhat smaller than the ratio of energy densities in the universe as a whole [2],

$$\frac{\Omega_b}{\Omega_M} = \frac{\Omega_b}{\Omega_b + \Omega_c} = 0.167 \pm 0.008, \quad (2)$$

where Ω_b is the fractional energy density in baryons, Ω_c the fractional energy density in cold dark matter, and $\Omega_M = \Omega_b + \Omega_c$ denotes the total fractional energy density in all matter. We also expect that larger clusters should have a higher proportion of baryons, since they will be able to gravitationally bind particles moving at higher speeds. Both of these expectations are confirmed by X-ray observations, which measure the current gas fraction f_g to be [3]

$$f_g (h/0.72)^{1.5} = 0.125 + 0.037 \log M_{15}, \quad (3)$$

where h is the dimensionless Hubble parameter (the Hubble constant H_0 is $100h$ km/sec/Mpc) and M_{15} is the cluster mass, in units of $10^{15} h^{-1} M_\odot$ (where the

¹The gas fraction includes both the hot ionized gas in the ICM and galaxies. The galaxies themselves act like a gas in hydrostatic equilibrium, as will be discussed in the next subsection.

h^{-1} appears because observations measure redshift rather than distance), out to the radius where the mean density enclosed is 500 times the (current) critical density of the universe.

2.2 The Isothermal β Model

The isothermal β model, first introduced by Cavaliere and Fusco-Femiano [4], considers the simplified case of a spherical, isothermal cluster that is in hydrostatic equilibrium; the model describes *relaxed clusters*, which are not currently undergoing merger events and have had time to settle into equilibrium. Here “equilibrium” means that matter in the cluster has no net force on it, so that the gravitational potential is balanced by the internal pressure of the gas:

$$\frac{d\phi(r)}{dr} = -\frac{1}{\rho(r)} \frac{dP(r)}{dr}. \quad (4)$$

Here $\phi(r)$ denotes the gravitational potential of the cluster, $\rho(r)$ is the gas mass density, and $P(r)$ is the radial pressure profile. The right-hand side of the equation follows from the definition

$$\vec{f} \equiv \vec{\nabla}P, \quad (5)$$

where \vec{f} is the force per unit volume.

First consider the situation for gas in the ICM. The mass density is the electron number density of the gas $n_e(r)$ multiplied by its mean mass,

$$\rho_g(r) = \mu_e n_e(r), \quad (6)$$

where $\mu_e \approx 1.14m_p$ is the gas mass per electron² and m_p is the mass of the proton. The β model assumes that the gas has a polytropic equation of state, so that $P_g(r) \propto \rho_g(r)^\gamma$; in the simplest case, where $\gamma = 1$, the gas follows the ideal gas law,

$$P_g(r) = n_e(r)k_B T, \quad (7)$$

where k_B is the Boltzmann constant and T is the gas temperature, which is constant by the assumption of isothermality. In this case (4) becomes

$$\frac{d\phi(r)}{dr} = -\frac{k_B T}{\mu_e n_e(r)} \frac{dn_e(r)}{dr} = -\frac{k_B T}{\mu_e} \frac{d \ln n_e(r)}{dr}. \quad (8)$$

At this point the right hand side of the equation can be integrated, but the left hand side is not yet known. To find it, we now consider the distribution

²We assume primordial composition of elements, with $\sim 76\%$ H and $\sim 24\%$ He by mass. Then the electron number density n_e is $\rho(0.76 + 2 \cdot \frac{0.24}{4}) = 0.88\rho$, with ρ the total gas mass density, since helium atoms have two electrons but an atomic weight four times greater than hydrogen. In this case $\mu_e = \frac{\rho}{n_e} \approx \frac{\rho}{0.88\rho} \approx 1.14$. A related quantity, the gas mass per particle μ , is given by $\mu = \frac{\rho}{n_e + n_i} \approx \frac{\rho}{\rho(0.88 + 0.82)} \approx 0.6$, since n_i , the ion number density, is $\rho(0.76 + \frac{0.24}{4}) = 0.82\rho$. Finally, we see that $n_i = \frac{0.82}{0.88} n_e \approx 0.93n_e$.

of dark matter, which can be estimated by observations under the assumption that it is tracked by galaxies. We assume that the galaxies have an isotropic velocity dispersion function

$$P(\vec{v}) = f(\vec{v})d^3\vec{v} = 4\pi v^2 f(v)dv, \quad (9)$$

where $P(\vec{v})$ is the probability for a galaxy to have a velocity between \vec{v} and $\vec{v} + d^3\vec{v}$ and $f(v)$, the Maxwellian velocity distribution, has units of inverse velocity cubed. In this case the line-of-sight velocity dispersion σ^2 , which can be directly observed, is

$$\sigma^2 = \frac{1}{3} \int f(\vec{v})v^2 d^3\vec{v} = \frac{4\pi}{3} \int f(v)v^4 dv. \quad (10)$$

Then the dark matter pressure distribution $P_{DM}(r)$ is the flux of momentum through a given area, which is equal to the velocity dispersion multiplied by the dark matter density profile:

$$P_{DM}(r) = \sigma^2 \rho_{DM}(r). \quad (11)$$

Now (4) becomes

$$\frac{d\phi(r)}{dr} = -\sigma^2 \frac{d \ln \rho_{DM}(r)}{dr}. \quad (12)$$

Setting (8) and (12) equal gives

$$\frac{d \ln n_e(r)}{dr} = \beta \frac{d \ln \rho_{DM}(r)}{dr}, \quad (13)$$

where we have defined the dimensionless parameter

$$\beta \equiv \frac{\sigma^2 \mu_e}{k_B T}. \quad (14)$$

Integrating and exponentiating gives the desired expression for the radial profile of the electron number density,

$$n_e(r) = n_e(0) \rho_{DM}^\beta. \quad (15)$$

$\rho_{DM}(r)$, the dark matter density profile, should be proportional to the galaxy number density profile, empirically derived by King from observations of the Coma cluster [5]:

$$\rho_{DM}(r) \propto \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3}{2}}, \quad (16)$$

where r_c , the core radius, is given by [4]

$$r_c = 3 \sqrt{\frac{\sigma^2}{4\pi G \rho_{G0}}}, \quad (17)$$

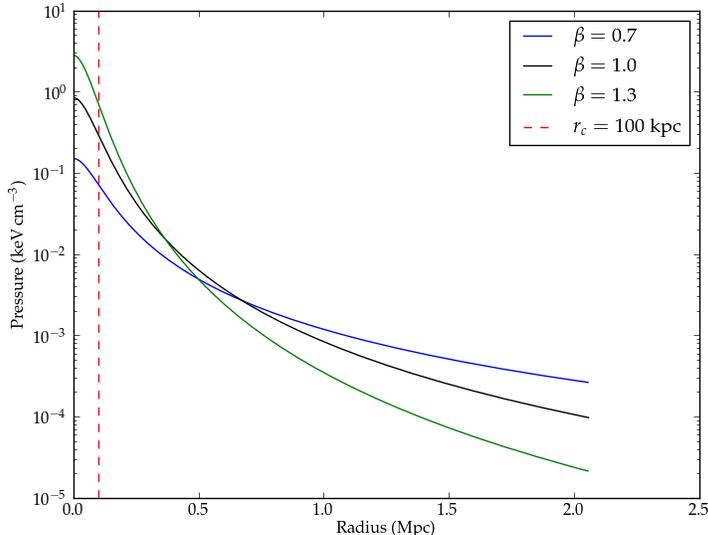


Figure 1: The β -model radial pressure profile (18) for a cluster with typical parameter values $r_c = 100$ kpc, $M_g(r_{200}) = 10^{14}M_\odot$, $z = 0.2$, plotted out to the virial radius $r_{200} \approx 2.05$ Mpc, which is independent of β (see equation (61) below). We plot three different values of β : $\beta = 0.7$ (blue curve), $\beta = 1$ (black), and $\beta = 1.3$ (green). The central pressure $P_g(0)$ is calculated by using equation (59) to find the central density, then using equations (6,7) to calculate the corresponding pressure. Evidently increasing β while leaving all other parameters fixed leads to a higher central pressure but a faster radial falloff. Adjusting β and r_c simultaneously, however, can lead to a degeneracy between large r_c /large β and small r_c /small β if only the inner region of the cluster can be observed, as is the case with ACT data: see Figure 6 on page 23 for an illustration of this degeneracy.

where ρ_{G0} is the central density of galaxies. In this case the gas pressure is given by

$$P_g(r) = P_g(0) \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}}. \quad (18)$$

This is the central result of the isothermal β model.

As discussed in Section 5.2 below, typical β -model parameters are $\beta \sim 1$, $r_c \sim 10^1 - 10^3$ kpc, and total virial gas mass (see Section 2.3.1) $M_g(r_{200}) \sim 10^{13} - 10^{15}M_\odot$. For concreteness, take $\beta = 1$, $r_c = 100$ kpc, $M_g(r_{200}) = 10^{14}M_\odot$. From the equations in Section 5.2, we see that the cluster has virial radius $r_{200} \approx 2.05$ Mpc and temperature 8.151 keV $\approx 9.46 \cdot 10^7$ K. If the cluster has redshift $z = 0.2$, then it is at an angular diameter distance $d_A \approx 665.6$ Mpc,

corresponding to a core angular size $\theta_c \approx 0.52'$ and a virial angular size $\theta_{200} \approx 10.6'$. Figure 1 plots the β -model radial pressure profile given these parameters and varying values of β .

2.3 Improved Models

2.3.1 Inadequacies of the β Model

Recall the assumptions of the isothermal β model: that galaxy clusters are isothermal spheres in hydrostatic equilibrium. It is evident that each of these assumptions can only apply in certain regimes. For example, the β model fails to take ellipticity or other irregularities in the shape of clusters into account; fixing this is a fairly easy extension to make, however. More important are the assumptions of hydrostatic equilibrium and isothermality. It was already mentioned that these assumptions only hold for relaxed clusters; there is no reason to expect the β model to hold for *morphologically disturbed clusters*, such as ones currently undergoing a merger. However, it is clear that they cannot apply universally even for fully relaxed clusters. At large radii, for example, the temperature must fall off to the background temperature of the universe, rather than remaining constant; the β model will therefore underestimate the dropoff of density profiles in this regime. Meanwhile, a large proportion of observed clusters contain so-called *cool cores* [33], central regions of the ICM that have sufficiently high densities that their cooling timescale is shorter than the Hubble time. Because mergers tend to disrupt the central regions of clusters, cool cores are almost always found in relaxed clusters, but not all relaxed clusters have cool cores [34].

Similarly, hydrostatic equilibrium fails to apply at large distances from the cluster center. Equilibrium can apply only in the regime where the matter has become *virialized*, meaning that it is in equilibrium between its own kinetic energy and the gravitational potential energy of the cluster. The *virial radius*, the maximum radius within which matter has become virialized, is usually expressed in terms of the ratio of the mean density within this radius to the critical density of the universe at the time of virialization:

$$\Delta_v = \frac{\rho_T(r_{vir})}{\rho_c(z)} = \frac{8\pi G\rho_T(r_{vir})}{3H^2(z)}. \quad (19)$$

For a universe with zero curvature [16],

$$\Delta_v \approx 18\pi^2 + 82x - 39x^2, \quad (20)$$

where

$$x = \frac{\Omega_M(1+z)^3}{\Omega_M(1+z)^3 + \Omega_\Lambda}, \quad (21)$$

with Ω_M and Ω_Λ the current proportions of total energy density matter (equal to $\Omega_b + \Omega_c$) and the cosmological constant, respectively. Using the WMAP maximum likelihood values [2] of $\Omega_\Lambda = 0.728$, $\Omega_M = 0.272$ results in values

of Δ_v slowly increasing from the current value of 197.07 ($z = 0$) to 220.65 at early times, where $z \rightarrow \infty$, $x \rightarrow 1$. The slow increase in Δ_v is thus essentially negligible, so for observational purposes the virial radius is usually taken to be the radius within which the mean density of the entire cluster is 200 times the critical density of the universe at the time of virialization³ and denoted r_{200} . Beyond r_{200} , the assumption of hydrostatic equilibrium ceases to apply.

2.3.2 The Generalized NFW Profile

Rather than developing new theoretical models to explain the inadequacies of the isothermal β model, astrophysicists have tended simply to extend it with additional parameters in attempts to fit the results of simulations and observations. Navarro, Frenk, and White (NFW) proposed a model [8] based on the results of N -body simulations which allowed pressure profiles that became singular rather than flat at small radii. The NFW model has itself been generalized [9]:

$$\frac{P_g(r)}{P_{500}} = \frac{n_e(r)}{k_B T(r) P_{500}} = \frac{P_0}{x^\gamma (1 + x^\alpha)^{\frac{\beta-\gamma}{\alpha}}}. \quad (22)$$

Here $x \equiv \frac{r}{r_s}$, $r_s = \frac{r_{500}}{c_{500}}$, and r_{500} is the radius at which the overdensity is 500 times the critical density at the time of virialization (used rather than the virial radius r_{200} because observations in this smaller area should have higher signal-to-noise ratios); P_{500} is a parameter with units of pressure. The parameters (γ, α, β) parametrize the slopes at small ($r \ll r_s$), intermediate ($r \sim r_s$) and large ($r \gg r_s$) radii, while the dimensionless c_{500} defines the transition between the regions of different slope and P_0 is a dimensionless normalization parameter.

2.3.3 The Arnaud *et al.* Universal Pressure Profile

Arnaud *et al.* [6] used a sample of 33 local ($z < 0.2$) clusters observed in the X-ray to find the best fitting values of the generalized NFW model (22). They find that

$$[P_0, c_{500}, \gamma, \alpha, \beta] = [8.403 h_{70}^{-\frac{3}{2}}, 1.177, 0.3081, 1.0510, 5.4905], \quad (23)$$

where h_{70} is the ratio of the actual Hubble parameter to the one with $h = 70$,

$$h_{70} = \frac{H_0}{70 \text{ kms}^{-1} \text{Mpc}^{-1}}. \quad (24)$$

Arnaud *et al.* also found the expected scaling relations of P_{500} and $P_g(r)$ with mass and redshift. Collecting everything together, their final empirical result for $P_g(r)$ is

$$1.65 \cdot 10^{-3} h(z)^{\frac{8}{3}} \left[\frac{M_{500}}{3 \cdot 10^{14} h_{70}^{-1} M_\odot} \right]^{\frac{2}{3} + \alpha_p + \alpha'_p (r/r_{500})} \frac{P_0}{(c_{500} x)^\gamma (1 + (c_{500} x)^\alpha)^{\frac{\beta-\gamma}{\alpha}}} h_{70}^2 \text{keVcm}^{-3}, \quad (25)$$

³This is equivalent to the background density in the case of a flat universe (where $\Omega_{tot} = 1$).

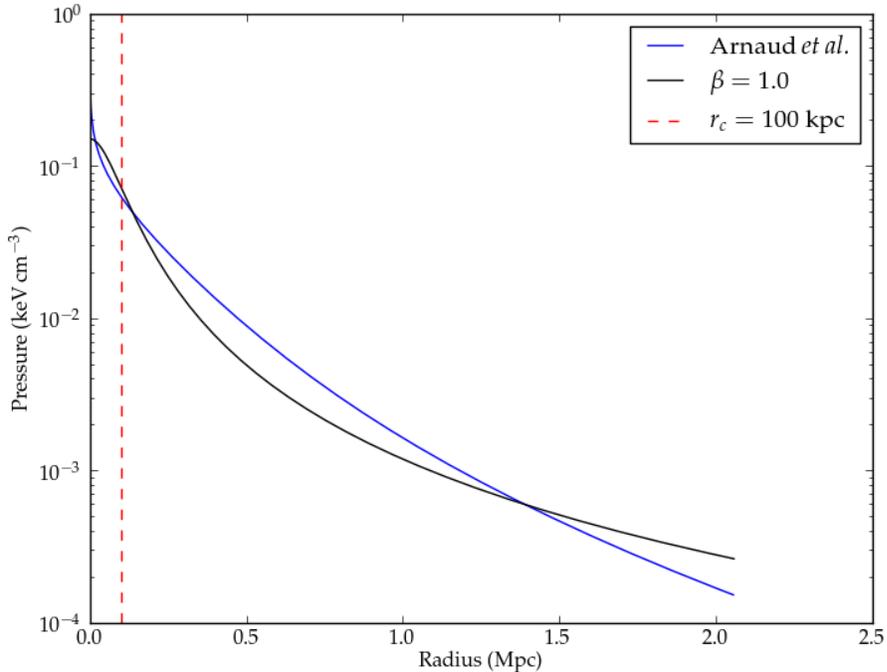


Figure 2: Comparison of the radial pressure profile of a typical β -model cluster discussed in Section 2.2 (black curve), with $\beta = 1$, $r_c = 100$ kpc, $M_g(r_{200}) = 10^{14} M_\odot$, and $z = 0.2$, to the Arnaud *et al.* universal pressure profile with the same values of z , M_{500} , and r_{500} (calculated using the relations in Section 5.2). We see that the Arnaud *et al.* profile has a faster falloff at large distances, as well as becoming singular rather than flat at small radii, in line with the NFW model.

where $x = \frac{r}{r_{500}}$, $h(z) = \frac{H(z)}{H_0}$, M_{500} is the mass contained within r_{500} of the cluster center, and α_p and $\alpha'_p \left(\frac{r}{r_{500}} \right)$ are empirically determined coefficients,

$$\alpha_p = 0.12, \quad \alpha'_p \left(\frac{r}{r_{500}} \right) = 0.10 - (\alpha_p + 0.10) \frac{r^3}{0.125 r_{500}^3 \left(1 + \frac{r^3}{0.125 r_{500}^3} \right)}. \quad (26)$$

Figure 2 compares the β -model and Arnaud *et al.* pressure profiles.

3 Cluster Observables

3.1 The Sunyaev-Zel'dovich Effect and the Compton y -Parameter

The Sunyaev-Zel'dovich (SZ) effect [10] is a distortion of the CMB spectrum due to inelastic (Compton) scattering off of thermal electrons like the ones in the ICM. Consider a region of electrons with an isotropic velocity distribution in the rest frame of the CMB. We wish to find the fractional change in the photon occupation number at a given frequency $n_\gamma(\nu)$ when photons with temperature T_{CMB} inverse Compton scatter off of the electrons at temperature $T \gg T_{CMB}$, thereby gaining energy and increasing in frequency. In the nonrelativistic regime $k_B T, k_B T_{CMB} \ll m_e c^2$, the time derivative of the occupation number is given by the Kompaneets equation [11]:

$$\frac{\partial n_\gamma}{\partial t} = \frac{k_B T}{m_e c^2} \frac{n_e \sigma_T c}{x^2} \frac{\partial}{\partial x} \left[x^4 \left(\frac{\partial n_\gamma}{\partial x} + n_\gamma + n_\gamma^2 \right) \right], \quad (27)$$

with σ_T the Thomson cross-scattering section, m_e the electron mass, and x the dimensionless ratio of energies $\frac{h\nu}{k_B T_{CMB}}$, which is convenient to use because the redshifting and cooling of the CMB occur at the same rate as the universe expands, so that x is redshift-independent. When the electrons are hot ($T \gg T_{CMB}$), as they are in the case of the ICM, only the first term in brackets in (27) is appreciable, and it simplifies to

$$\frac{\partial n_\gamma}{\partial t} = \frac{k_B T}{m_e c^2} \frac{n_e \sigma_T c}{x^2} \frac{\partial}{\partial x} \left[x^4 \frac{\partial n_\gamma}{\partial x} \right]. \quad (28)$$

This expression can be simplified by introducing the dimensionless *Compton y -parameter*,

$$y \equiv \int dy = \int_t^{t_0} \frac{k_B T}{m_e c^2} n_e \sigma_T c dt = \int \frac{k_B T}{m_e c^2} n_e \sigma_T dl, \quad (29)$$

where the last integral represents the integrated pressure over the line of sight. Physically, the y -parameter tells us the average number of times a CMB photon scatters when passing through the cluster. The typical β -model cluster given in Section 2.2 has a central y -parameter of $6.81 \cdot 10^{-4}$, meaning that around 7 of every 10000 CMB photons we observe coming from the region of the sky where the center of the cluster is located have been scattered by it. We see that y is redshift-independent: it depends only on the properties of the cluster itself, not its distance from an observer or the temperature of the CMB.

With this definition of y , (28) simplifies to

$$\frac{\partial n_\gamma}{\partial y} = \frac{1}{x^2} \frac{\partial}{\partial x} \left[x^4 \frac{\partial n_\gamma}{\partial x} \right]. \quad (30)$$

If the incident spectrum is unperturbed, then the photons occupy states according to the Planck function,

$$n_\gamma = \frac{1}{e^x - 1}, \quad (31)$$

so (see Appendix B for the derivation)

$$\frac{\partial n_\gamma}{\partial y} = \frac{xe^x}{(e^x - 1)^2} \left[x \coth\left(\frac{x}{2}\right) - 4 \right], \quad (32)$$

and

$$\frac{\Delta n_\gamma}{n_\gamma} = y \frac{xe^x}{e^x - 1} \left[x \coth\left(\frac{x}{2}\right) - 4 \right]. \quad (33)$$

But according to Planck's law the CMB radiation intensity is proportional to its occupation number,

$$I_{CMB} = \frac{2h\nu^3}{c^2} n_\gamma, \quad (34)$$

so the change in intensity due to inverse Compton scattering is given by

$$\frac{\Delta n_\gamma}{n_\gamma} = \frac{\Delta I_{CMB}}{I_{CMB}} = \frac{\Delta I_{SZ}}{I_{CMB}} = g(x)y, \quad g(x) = \frac{xe^x}{e^x - 1} \left[x \coth\left(\frac{x}{2}\right) - 4 \right], \quad (35)$$

where we identify the intensity shift in the CMB ΔI_{CMB} as ΔI_{SZ} since it is due to the SZ effect. The corresponding temperature change is given by differentiating (34) with respect to T_{CMB} and multiplying (35) by the resulting expression:

$$\frac{\Delta T_{SZ}}{T_{CMB}} = f(x)y, \quad f(x) = x \coth\left(\frac{x}{2}\right) - 4. \quad (36)$$

This is the desired result, relating the temperature change to the properties of the gas which causes the inverse Compton scattering. Since we used the nonrelativistic Kompaneets equation (27), this expression neglects relativistic corrections, which will be of order $\frac{k_B T}{m_e c^2}$; since a typical cluster temperature is $\sim 5 - 10$ keV, the corrections will be on the 1 - 2% level. Consider the effects of this distortion on a given frequency range (and therefore a given energy range) in the initially undistorted CMB: some photons of lower frequency will interact with the ICM, gain energy, and therefore enter the frequency range; on the other hand, some of the photons initially in the range will themselves gain energy from the ICM and exit it. Because the blackbody distribution is peaked, rather than monotonically increasing or decreasing, there will be some point where the numbers of incoming and outgoing photons are identical, so that there is zero net intensity change from the SZ effect. From (35), this happens when $f(x) = g(x) = 0$, which occurs at $x \approx 3.83$, or 218 GHz at the current temperature of the CMB; at frequencies below this point, like the 148 GHz ACT observation band, the SZ effect will be seen as a temperature and intensity *decrement*.

The y parameter (29) is a direct measurement of the electron number density and temperature of a cluster. Combining equations (18) and (36), it relates the observed temperature change directly to the cluster properties in the isothermal β model:

$$\Delta T_{SZ} = \frac{\sigma_T n_e(0) k_B T T_{CMB} f(x)}{m_e c^2} \int \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}} dl, \quad (37)$$

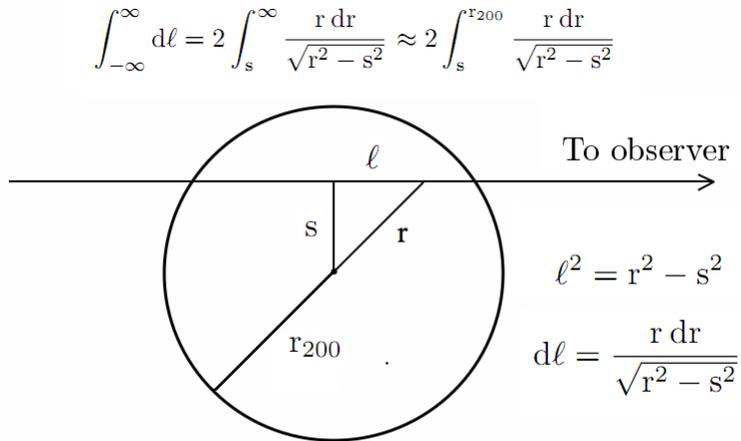


Figure 3: Geometry for transforming between a two-dimensional projection along the line of sight ℓ and a three-dimensional sphere, where the line of sight lies a distance s from the cluster center. In the β model we treat each cluster as a sphere of radius r_{200} , so that the upper limit of integration can be set there rather than at infinity.

where ℓ ranges along the line of sight. Then the temperature decrement observed a projected distance s from the cluster center is given by

$$\Delta T_{SZ} = \frac{2\sigma T n_e(0) k_B T T_{CMB} f(x)}{m_e c^2} \int_s^{r_{200}} \frac{r}{\sqrt{r^2 - s^2}} \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}} dr, \quad (38)$$

where r is the actual distance from the cluster center and r_{200} is the virial radius (see Section 2.3.1 above) which marks the edge of the cluster in the β model. See Figure 3 for the geometry involved in changing the variable of integration from ℓ to r . In the limit that $r_{200} \gg s$, so that the upper limit can be set at ∞ , we get the approximate scaling relation

$$\Delta T_{SZ} \propto \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{\frac{1-3\beta}{2}}. \quad (39)$$

3.2 X-Ray Luminosity and Surface Brightness

In addition to scattering CMB photons, the hot gas of the ICM also emits X-ray photons directly via free-free emission (thermal brehmsstrahlung). The emission is quantified by the *emissivity*, which is the emitted luminosity L_X per unit volume V . For free-free emission between ions of charge Z and electrons

at temperature T , the emissivity at a frequency ν is given by [12]

$$\epsilon_{\nu}^{ff} = \frac{dL_X}{dV d\nu} \frac{2^5 \pi e^6}{3 m_e c^3} \sqrt{\frac{2\pi}{3 m_e k_B}} Z^2 n_e n_i g_{ff}(Z, T, \nu) T^{-\frac{1}{2}} e^{-\frac{h\nu}{k_B T}}, \quad (40)$$

where n_e and n_i are the number density of electrons and ions, respectively, and g_{ff} , the *Gaunt factor*, of order unity, parameterizes the quantum corrections to the semiclassical expression. At the high temperatures of the ICM, free-free emission is the dominant contribution to the X-ray emissivity; other factors, such as radiative recombination (bound-free emission) and decay of metastable states, can be neglected. Integrating over the frequency gives

$$\epsilon = \frac{dL_X}{dV} = 0.93 \Lambda_0 n_e^2 T^{\frac{1}{2}}, \quad (41)$$

where the constants have been collected into the single normalization constant Λ_0 , of order $2 \cdot 10^{-27}$ ergs \cdot cm³s⁻¹K^{- $\frac{1}{2}$} for a typical cluster [13] and we have used the fact that $n_i \approx 0.93 n_e$ for gas with a primordial composition of hydrogen and helium (see Footnote 2 on page 4). Now it is straightforward to find the X-ray observable, the *surface brightness* S_X , defined as the X-ray F_X per unit solid angle. Integrating over along the line of sight gives the luminosity per unit surface area,

$$\frac{L_X}{dA} = \frac{L_X}{d_A^2 d\Omega} = \int 0.93 \Lambda_0 n_e^2 T^{\frac{1}{2}} dl, \quad (42)$$

where d_A is the angular diameter distance and $d\Omega$ denotes the subtended solid angle. Now flux is related to luminosity by

$$F_X = \frac{L_X}{4\pi d_A^2 (1+z)^4}, \quad (43)$$

with z the redshift, so

$$S_X = \frac{F_X}{d\Omega} = \frac{0.93 \Lambda_0}{4\pi (1+z)^4} \int n_e^2 T^{\frac{1}{2}} dl. \quad (44)$$

Note the different dependences of y (29) and S_X (44) on temperature and electron number density.

Like the SZ temperature decrement (38), the surface brightness can be directly expressed in terms of the parameters of the β model:

$$S_X = \frac{0.93 \Lambda_0 T^{\frac{1}{2}}}{4\pi (1+z)^4} \int \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-3\beta} dl. \quad (45)$$

The surface brightness a projected density s from the cluster center is given by

$$S_X = \frac{0.93 \Lambda_0 T^{\frac{1}{2}}}{4\pi (1+z)^4} \int_s^{r_{200}} \frac{r}{\sqrt{r^2 - s^2}} \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-3\beta} dl. \quad (46)$$

In the limit of large virial radius r_{200} , this gives the approximate scaling relation

$$S_X \propto \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{\frac{1-6\beta}{2}}. \quad (47)$$

4 Dataset

4.1 Radio Data

4.1.1 Data Collection and Mapmaking

The Atacama Cosmology Telescope [23] is a six-meter off-axis Gregorian telescope located at an altitude of 5190 m near the peak of Cerro Toco in the Atacama Desert in northern Chile (latitude $22^{\circ}57'31''\text{S}$, longitude $67^{\circ}47'15''\text{W}$). The focal-plane instrument is the Millimeter Bolometer Array Camera (MBAC), which contains three 1024-element arrays of transition edge sensor (TES) bolometers operating at 148 GHz, 218 GHz, and 277 GHz. This analysis, following [24], uses data from a subset of the 2008 southern survey, covering a 455 deg^2 region lying between right ascensions $00^{\text{h}}12^{\text{m}}$ and $07^{\text{h}}08^{\text{m}}$ and declinations $-56^{\circ}11'$ and $-49^{\circ}00'$. As described in [25], ACT makes 4.5° scans in azimuth at a constant 50° elevation, sampling each detector at 399 Hz. Data are written in 15-minute segments called time-ordered data sets (TODs); each detector TOD uses 1.6 GB of storage per array.

Once the data has been collected, bad TODs are cut on the basis of criteria such as bad weather, high receiver temperature, and detector malfunctions [26]. The resulting data is then calibrated based on the measured temperature of Uranus, as well as cross-correlation with WMAP 94 GHz measurements; pointing is calibrated based on observations of Saturn and known radio sources [25]. Finally, maps are constructed iteratively by solving for a maximum likelihood solution based on a preconditioned conjugate gradient method [27].

4.1.2 Calculation of y

This analysis uses portions of the 148 GHz map containing 21 of the clusters presented in the Marriage *et al.* ACT cluster catalog [24]. The submaps vary slightly in size and resolution but tend to measure around half a square degree, with pixels approximately $0.5'$. At this stage the maps show specific intensity (also referred to as “radio brightness”), measured in units of $\frac{\text{Jy}}{\text{sr}}$ ($1 \frac{\text{Jy}}{\text{sr}} = 10^{-26} \frac{\text{W}}{\text{m}^2 \cdot \text{Hz} \cdot \text{sr}}$). To convert from intensity to temperature, we use equations (34,35,36):

$$\frac{\Delta T_{SZ}}{T_{CMB}} = \frac{\Delta I_{SZ}}{I_{CMB}} \frac{f(x)}{g(x)} = \Delta I_{SZ} \frac{2k_B T_{CMB} c^2}{h^2 \nu^4} \sinh^2 \left(\frac{h\nu}{k_B T_{CMB}} \right), \quad (48)$$

where $\nu = 148 \text{ GHz}$. The maps are then converted from temperature to y -parameter, using (36). Each submap is accompanied by an equivalent weight

map which gives the uncertainty on each intensity measurement; the same steps are followed to convert the weight map to y -parameter uncertainties.

The map is then filtered to remove noise from the CMB, using the Flipper Python package [28], specifically designed for working with CMB data. This is done by removing all power where $\ell < 180$ from the two-dimensional Fourier transform of the map, then returning to real space⁴. From the filtered map, we find the cluster center by searching for the maximum of the Gaussian-smoothed map.

Figure 4 shows the y -parameter maps for the Bullet Cluster (ACT-CL J0658-5557) before and after filtering. As desired, the large-scale features of the map, including an overall nonzero mean, are removed, leaving the cluster to clearly stand out. Note that we do not mask the immediate area of the cluster during the filtering process: as a result, some power is removed from the cluster along with the rest of the map, although presumably a significant portion of this power was due to CMB fluctuations rather than the cluster itself.

Using the measured optical redshifts from [29], we calculate the angular diameter distance:

$$d_A = \frac{c}{1+z} \int_0^z \frac{dz'}{H(z')} = \frac{c}{H_0(1+z)} \int_0^z \frac{dz'}{\sqrt{\Omega_\Lambda + \Omega_M(1+z)^3}}. \quad (49)$$

Using d_A , the data is then binned in 0.1 Mpc-wide annuli centered at the cluster center, moving outward at 0.1 Mpc intervals⁵, in each of which a weighted average is taken to produce the final value and uncertainty $[y(s), \sigma_y(s)]$. We continue moving outward from the cluster by incrementing s until there is no detectable excess in y above its background level: that is, when increasing the radius results in an increased y , which is not compatible with the β model for physical values of β (see Section (2.2)). Typically, this results in $\sim 4 - 8$ points, meaning that the radio data is taken from a region inside the cluster that is well less than 1 Mpc in radius. The collected y -parameter and uncertainties are collected into the data vector \mathbf{D} and uncertainty vector $\boldsymbol{\sigma}$, respectively.

4.2 X-Ray Data

4.2.1 Data

Two of the clusters we consider in this analysis also have X-ray data available, provided by Professor Jack Hughes at Rutgers: ACT-CL J0658-5557 (1E0657-56, the Bullet cluster) was observed by the Advanced CCD Imaging Spectrometer (ACIS) on the Chandra X-ray Observatory [30] and ACT-CL J0516-5430

⁴When removing power, we would like to remove the CMB on large scales, which can create gradient patterns that distort y -parameter values, while keeping as much power from the cluster itself as possible. Experimenting with different cutoff ℓ values showed that $\ell = 180$ was the best compromise between these two goals; just removing the mean ($\ell = 0$) left gradient patterns, while removing as much CMB power as possible ($\ell = 1800$) also destroyed the cluster signature itself.

⁵Some of the clusters with higher redshifts have sufficiently large angular diameter distances that the innermost annuli do not intersect any of the pixel centers. In this case the radial interval is increased to 0.2 or 0.3 Mpc, as necessary.

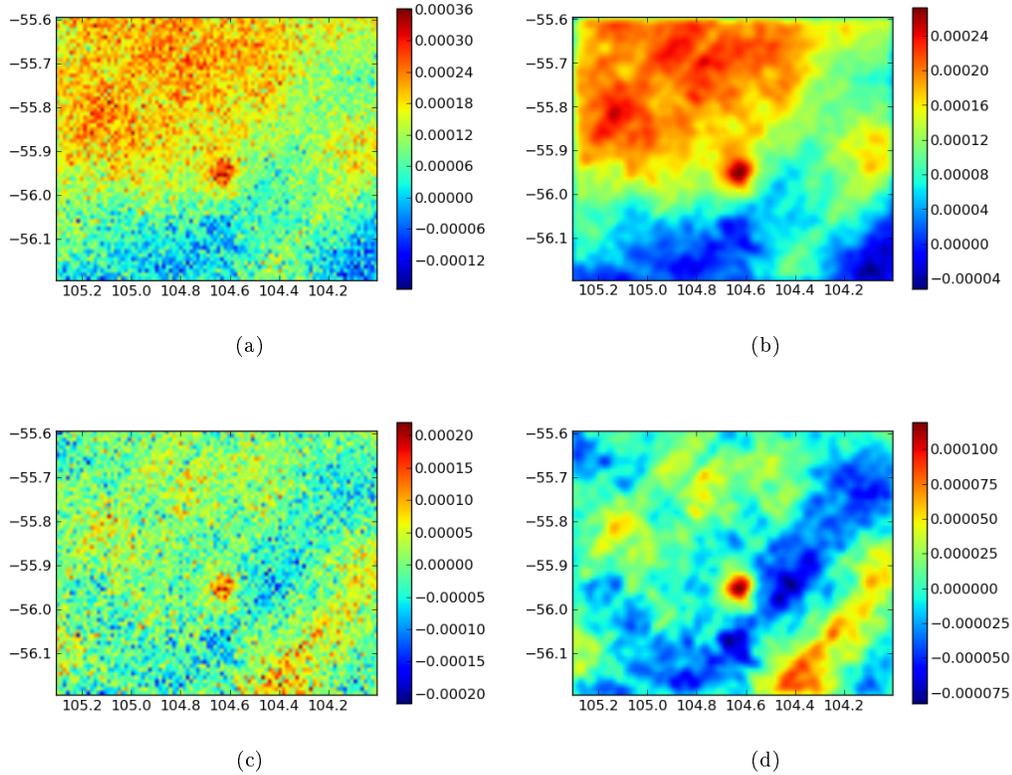


Figure 4: y -parameter maps of the Bullet cluster (ACT-CL J0658-5557). The x and y axes give degrees of declination and right ascension, respectively. The color denotes the y -parameter (dimensionless). (a,b) Unfiltered map, both (a) unsmoothed and (b) Gaussian smoothed to allow easier identification of large-scale features. Note the large northwest-southeast gradient. (c,d) Filtered map with power below $\ell = 180$ removed, both (c) unsmoothed and (d) Gaussian smoothed. The large-scale features of the map are removed, and the mean is set to zero since the zero-frequency mode is removed. (c) is the map used for the maximum likelihood analysis.

(Abell S0520, SPT-CL 0517-5430) was observed by the European Photon Imaging Camera (EPIC) Metal Oxide Superconductor camera 1 (MOS1) on the X-ray Multi-Mirror Mission – Newton (XMM-Newton) [31]. As is standard for X-ray observations, the data is given in the form of a count map, with units of $\frac{\text{counts}}{\text{arcmin}^2 \cdot \text{cm}^2 \cdot \text{sec}}$. The maps used in this analysis have had point sources removed, adaptive smoothing applied, and the background level subtracted.

4.2.2 Calculation of S_X

Conversion of the count maps to surface brightness maps, with units of flux per solid angle, or $\frac{\text{ergs}}{\text{arcmin}^2 \cdot \text{cm}^2 \cdot \text{sec}}$, requires estimation of the energy of each count. In this analysis we adopt the simplest possible conversion scheme: we assume that each count corresponds to a photon with energy $k_B T$, where T is the temperature of the cluster as measured in [29]. Hence we multiply each map by $k_B T$ and then bin in annuli as above to find the surface brightness as a function of projected radius $S_X(s)$. Lacking a weight map to calculate uncertainties, we instead take the error on each point to be equal to the root-mean-square value of the entire map⁶.

5 Cluster Properties via Maximum Likelihood

5.1 Description of the Problem

Following the notation of [14], consider an N -element set of data \mathbf{D} to be described by a set of p parameters Θ from a model H . According to Bayes' theorem [15],

$$\Pr(\Theta|\mathbf{D}, H) = \frac{\Pr(\mathbf{D}|\Theta, H) \Pr(\Theta|H)}{\Pr(\mathbf{D}|H)} \quad (50)$$

where $\Pr(\mathbf{A}|\mathbf{B})$ is the conditional probability of \mathbf{A} given \mathbf{B} . Here $\Pr(\Theta|\mathbf{D}, H) \equiv P(\Theta)$ is the posterior probability of Θ , $\Pr(\mathbf{D}|\Theta, H) \equiv \mathcal{L}(\Theta)$ the likelihood, $\Pr(\Theta|H) \equiv \pi(\Theta)$ the prior probability of Θ , and $\Pr(\mathbf{D}|H) \equiv \mathcal{Z}$ the Bayesian evidence factor. For the purposes of parameter estimation within a given model, \mathcal{Z} can be neglected, so that

$$P(\Theta) \propto \mathcal{L}(\Theta) \pi(\Theta). \quad (51)$$

This analysis in this paper employs only uniform priors, so

$$P(\Theta) \propto \mathcal{L}(\Theta). \quad (52)$$

In this case $\Pr(\Theta|\mathbf{D}, H)$, the posterior probability distribution of the β model parameters given the data, which is the objective of the analysis, can be directly calculated by finding $\Pr(\Theta|H)$, the likelihood of the data given the parameters. Hence we can estimate the probability that a cluster has a

⁶Since the map has been masked to remove point sources, we in fact calculate the root-mean-square value of only the *unmasked* points.

given set of β -model parameters by evaluating the likelihood function for these parameters. Making the usual assumption that the observed values of the data are normally distributed around their true values implies that the likelihood itself is given by a Gaussian,

$$\mathcal{L}(\Theta) = Ae^{-\frac{\chi^2}{2}}, \quad (53)$$

with A a normalization constant, and χ^2 given by

$$\chi^2 = \sum_{i=1}^N \left(\frac{D_i - D_i^p(\Theta)}{\sigma_i} \right)^2 \quad (54)$$

where $D^p(\Theta)$ is the N -element set of predicted data calculated from the parameters Θ and σ_i is the uncertainty on the i th data point D_i . Then the p -dimensional likelihood distribution for the parameters is given by normalizing $\exp(-\chi^2/2)$, and the likelihood distribution for a subset of parameters $\tilde{\Theta} \subset \Theta$ is given by marginalizing the distribution over $\Theta \setminus \tilde{\Theta}$ (the set Θ with the elements in $\tilde{\Theta}$ removed). In particular, the one-dimensional marginalized likelihood distribution for the parameter Θ_j is given by

$$\mathcal{L}(\Theta_j) = \int \dots \int Be^{-\frac{\chi^2}{2}} d\Theta_1 \dots d\Theta_{j-1} d\Theta_{j+1} \dots d\Theta_p, \quad (55)$$

with the integrals over the entire parameter space and B a normalization constant. Once the data and uncertainty values (described in Section 4 above) have been provided, all that is needed for the calculation is a method of deriving the predicted data D^p from the parameters Θ .

5.2 Calculation of the Predicted Data

As described in Section 4 above, the data D consists of a set of y -parameter values (and S_X values for clusters with X-ray data available) calculated for annuli extending outward from the cluster centers in intervals of 100 kpc or greater. We wish to calculate the corresponding predicted data D^p for a cluster of known redshift z at the same intervals given a set of β -model parameters Θ . Recall that the gas pressure in the β model is given by (18)

$$P_g(r) = P_g(0) \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}}. \quad (56)$$

Following [14], we work in terms of the cluster gas density, which (from equations (6) and (7)) is proportional to the pressure and thus scales in the same way :

$$\rho_g(r) = \rho_g(0) \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}}. \quad (57)$$

Density is more convenient to work with because the central density $\rho_g(0)$ can easily be related to the gas mass contained within the virial radius, $M_g(r_{200})$:

$$M_g(r_{200}) = \rho_g(0) \int_0^{r_{200}} 4\pi r \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}} dr, \quad (58)$$

so

$$\rho_g(0) = M_g(r_{200}) \left(\int_0^{r_{200}} 4\pi r \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}} dr \right)^{-1}. \quad (59)$$

Then the predicted y -parameter at a radius s is (using equations 18,36,38):

$$y(s) = \frac{2\sigma_T \rho_g(0) k_B T}{\mu_e m_e c^2} \int_s^{r_{200}} \frac{r}{\sqrt{r^2 - s^2}} \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{3\beta}{2}} dr. \quad (60)$$

Using this equation, $\mathbf{D}^{\mathbf{P}}$ can be calculated from only three parameters: β , the core radius r_c , and the virial mass $M_g(r_{200})$. The temperature T and virial radius r_{200} can both be calculated once the total virial mass $M_T(r_{200})$ is known: r_{200} is *defined* by the requirement that the mean density within it is equal to 200 times the critical density,

$$\frac{M_T(r_{200})}{\frac{4}{3}\pi r_{200}^3} = 200\rho_c(z) \implies r_{200} = \left[\frac{M_T(r_{200})}{\frac{800}{3}\pi\rho_c(z)} \right]^{\frac{1}{3}}, \quad (61)$$

while the temperature is given by the virial theorem, which states that the mean kinetic energy of the system is twice its gravitational potential energy,

$$k_B T = \frac{G\mu M_T(r_{200})}{2r_{200}}, \quad (62)$$

where $\mu \approx 0.6m_p$ is the gas mass per particle (see Footnote 2 on page 4).

Finally, it remains to calculate the total virial mass $M_T(r_{200})$. Since $M_g(r_{200})$ is given, this amounts to finding the gas fraction at r_{200} . The gas fraction at r_{500} is measured by X-ray observations as (3)

$$f_g(r_{500}) = \frac{M_g(r_{500})}{M_T(r_{500})} = 0.125 + 0.037 \log \left(\frac{M_T(r_{500})}{10^{15} h^{-1} M_\odot} \right), \quad (63)$$

where we have set $h = 0.72$. To extrapolate this to r_{200} , we use the approximation⁷ $r_{500} \approx 0.65r_{200}$ [2] and apply the virial theorem again at r_{500} to find

$$f_g(r_{200}) \approx \left[0.125 + 0.037 \log \left(\frac{M_T(r_{500})}{10^{15} h^{-1} M_\odot} \right) \right] \frac{5}{2} \cdot 0.65^3, \quad (64)$$

⁷To derive this approximation, integrate (8) using the density profile (6), then take $r_X \gg r_c$ to find $M_T(r_X) \propto r_X \beta T$, where $X = 200, 500$. Combined with (62), this says $r_{200} \propto \sqrt{\frac{\beta T}{200}}$, and similarly $r_{500} \propto \sqrt{\frac{\beta T}{500}}$, since the virial theorem also applies at smaller radii. Then $r_{500} \approx \sqrt{0.4} \approx 0.65$.

or

$$\frac{M_g(r_{200})}{M_T(r_{200})} \approx 0.0858 + 0.0254 \log \left(\frac{\frac{5}{2} M_T(r_{200}) \cdot 0.65^3}{10^{15} h^{-1} M_\odot} \right). \quad (65)$$

This can be solved in terms of special functions⁸ to find $M_T(r_{200})$, as desired.

The procedure for finding \mathbf{D}^p for a given $(\beta, r_c, M_g(r_{200}))$ is then as follows. First, calculate the total virial mass $M_T(r_{200})$ by solving (65). Then find the virial radius r_{200} from (61) and the temperature T from (62), as well as the central density $\rho_g(0)$ from (59). Finally, find $y(s)$ at the desired values of s from the integral⁹ in (60).

Following [14], we consider β -parameters in the ranges $0.3 \leq \beta \leq 3.1$, $50 \text{ kpc} \leq r_c \leq 3000 \text{ kpc}$, $10^{13} M_\odot \leq M_g(r_{200}) \leq 10^{15} M_\odot$. For each cluster, we calculate likelihoods for around 40000 points in the parameter space.

5.2.1 X-Ray Predicted Data

In the β model, the surface brightness is given by (46)

$$S_X = \frac{0.93 \Lambda_0 T^{\frac{1}{2}}}{4\pi(1+z)^4} \int_s^{r_{200}} \frac{r}{\sqrt{r^2 - s^2}} \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-3\beta} dl. \quad (66)$$

The normalization constant Λ_0 , which depends on the detailed chemical composition of the ICM, can vary from cluster to cluster and, unlike the gas fraction, has not been precisely measured by experiments. In this analysis, therefore, we calculate only the shape of the surface brightness profile, not its normalization, using the scaling relation (47)

$$S_X = S_X^0 \left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{\frac{1-6\beta}{2}}. \quad (67)$$

The normalization S_X^0 is given by the central surface brightness found from X-ray observations.

6 Analysis of ACT Clusters

Space constraints prevent a full discussion of each individual cluster. We will therefore discuss one cluster, ACT-CL J0516-5430 (also known as Abell S0520 and SPT-CL 0517-5430) in detail¹⁰, then discuss possible sources of systematic error in the analysis.

⁸The solution is $M_T(r_{200}) \approx 90.653 M_g(r_{200}) W \left(\frac{148599 M_g(r_{200})}{10^{15} h^{-1} M_\odot} \right)^{-1}$, where $W(x)$, the Lambert W function, is defined by $x = W(x)e^{W(x)}$.

⁹The integral can actually be solved exactly in terms of Gauss hypergeometric functions ${}_2F_1(a, b; c; z)$; the solution is not provided here because it takes several lines to write, but in practice it is used in the likelihood analysis to speed up the execution time of the code.

¹⁰The second cluster with X-ray data, ACT-CL J0658-5557 (the Bullet cluster), is discussed in Appendix A below.

6.1 ACT-CL J0516-5430

6.1.1 SZ Analysis

The first step in the analysis is to create a filtered y -parameter map, as described in Section 4.1.2 above. We can then proceed to a maximum likelihood analysis as described in Section 5 above. Figure 5 gives the resulting one- and two-dimensional marginalized distributions for the β model parameters β , r_c , and $M_g(r_{200})$.

The marginalized parameter distributions for the cluster exhibit a trend common to the sample as a whole: the core radius and gas mass are well-constrained by the SZ data, but β is not. This is not entirely surprising: heuristically, the predicted central gas density $\rho_g(0)$ (59) has a β dependence that is the inverse of the β dependence of the rest of the expression for the predicted y -parameter (60), so much of the dependence cancels, especially since the usable data from J0516-5430 only extends to 0.5 Mpc. The result is that the SZ data lacks the power to constrain β , which has the greatest effect at values near the edge of the cluster: essentially, the problem is that there is a degeneracy between high β /high core radius r_c and low β /low r_c if only the inner part of the cluster is observed. Figure 6 illustrates this point by showing three different sets of β -model parameters, all of which fit the data excellently despite having wildly divergent values of β . The same degeneracy is seen in a more sophisticated Bayesian analysis of eight clusters observed via the SZ effect in multiple radio wavelengths by the Arcminute Microkelvin Imager (AMI) [35]; it is an inherent feature of any analysis which uses only small-radius y -parameter measurements, as the curves in Figure 6 make clear.

6.1.2 Adding X-Rays

The predicted surface brightness (67), on the other hand, has power only to constrain the radial profile of the cluster, rather than its absolute normalization. We therefore expect that it should provide significantly tighter limits on the value of β . Where X-ray data is available, as it is in the case of J0516-5430, it provides a larger number of high-significance points with lower relative uncertainties than the corresponding radio data out to larger radii. In the case of J0516-5430, we saw above that using ACT data we could extract y -parameter information out to only 0.5 Mpc; the corresponding XMM-Newton data allows high-significance surface brightness extraction to 0.9 Mpc, almost twice the distance. Figure 7 shows the marginalized parameter distributions that result from taking into account X-ray observations in addition to the SZ data.

The preferred values of β and r_c , which can be extracted from the shape of the distribution, have tightened extremely, picking out much smaller allowed regions in the parameter space: Figure 8 illustrates how the reduced noise and greater radial extent of the XMM-Newton data picks out one particular β -model profile. Meanwhile, as expected, $M_g(r_{200})$ is not significantly further constrained, since the surface brightness data is fit only to the shape of the profile, not its overall normalization (see Section 5.2.1). Note that the 2σ allowed

ACT-CL J0516-5430 Marginalized Parameters (ACT data only)

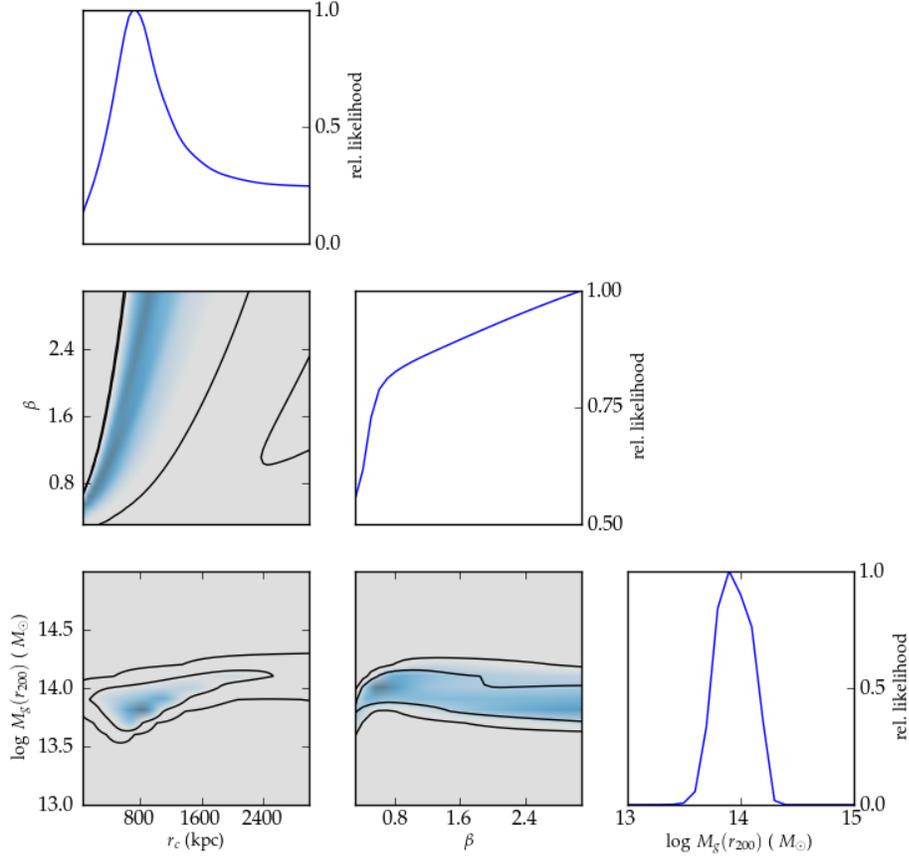


Figure 5: One- and two-dimensional marginalized parameter distributions for ACT-CL J0516-5430 ($z = 0.294$). The y -axis of the one-dimensional plots gives normalized relative likelihoods, with 1 representing the likelihood of the best-fit parameter values. As seen in Figure 6 below, the single set of parameters that best fits the data is $\beta = 2.2$, $r_c = 700$ kpc, $M_g(r_{200}) = 10^{13.8} M_\odot$, which has $\chi^2/\text{D.O.F.} = 0.624/3$. The contours in the two-dimensional plots are 1σ (68%) and 2σ (95%) confidence intervals (taking into account only statistical uncertainties, not systematic ones). The cluster gas mass $M_g(r_{200})$ is well-constrained, and the possible values for the core radius r_c are significantly limited, but β is essentially unconstrained.

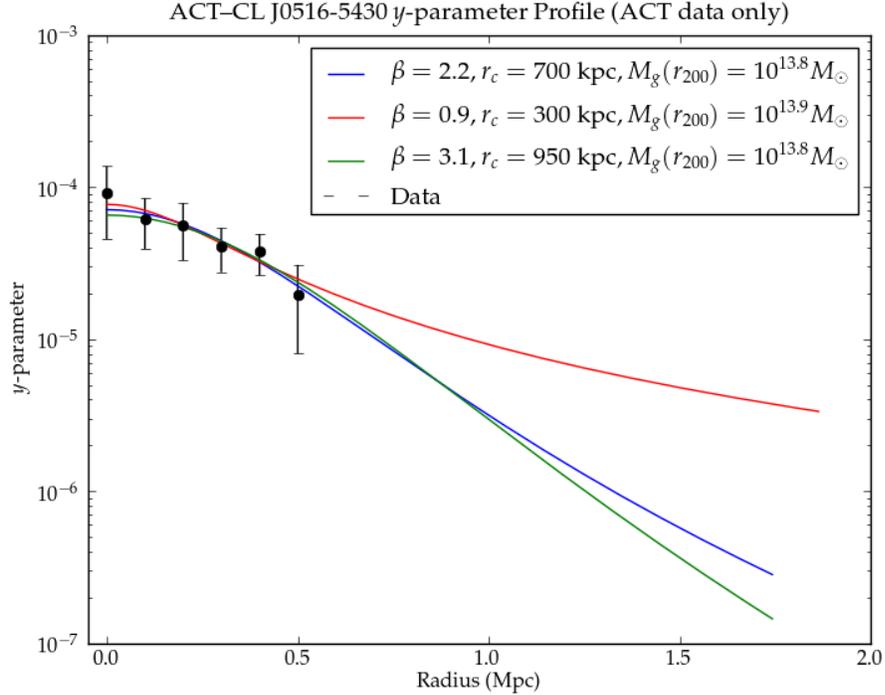


Figure 6: Calculated y -parameter radial profile for ACT-CL J0516-5430, along with several theoretical β -model curves that fit the data well. The blue curve, with $\beta = 2.2$, $r_c = 700$ kpc, $M_g(r_{200}) = 10^{13.8} M_\odot$, is the single set of parameters that best fits the data; it has $\chi^2/\text{D.O.F.} = 0.624/3$. The red curve, with $\beta = 0.9$, $r_c = 300$ kpc, $M_g(r_{200}) = 10^{13.9} M_\odot$, has $\chi^2/\text{D.O.F.} = 0.712/3$. The green curve, with $\beta = 3.1$, $r_c = 950$ kpc, $M_g(r_{200}) = 10^{13.8} M_\odot$, has $\chi^2/\text{D.O.F.} = 0.679/3$. These reduced χ^2 values for J0516-5430 are much smaller than the equivalent values for the Bullet cluster (see Figure 9 on page 30) both because J0516-5430 was observed less often, so the absolute errors are several times larger, and because J0516-5430 is significantly dimmer in the radio, with lower absolute y -parameter values and correspondingly larger error bars in logarithmic space. We see that the data has no strong preference between even wildly-varying β values, since β only strongly affects y beyond the core radius. By contrast, the gas mass $M_g(r_{200})$ is well-constrained, since the overall normalization depends on it strongly, and the core radius r_c is mildly constrained by the falloff in y between the center and the edge of the data.

ACT-CL J0516-5430 Marginalized Parameters (ACT + XMM-Newton data)

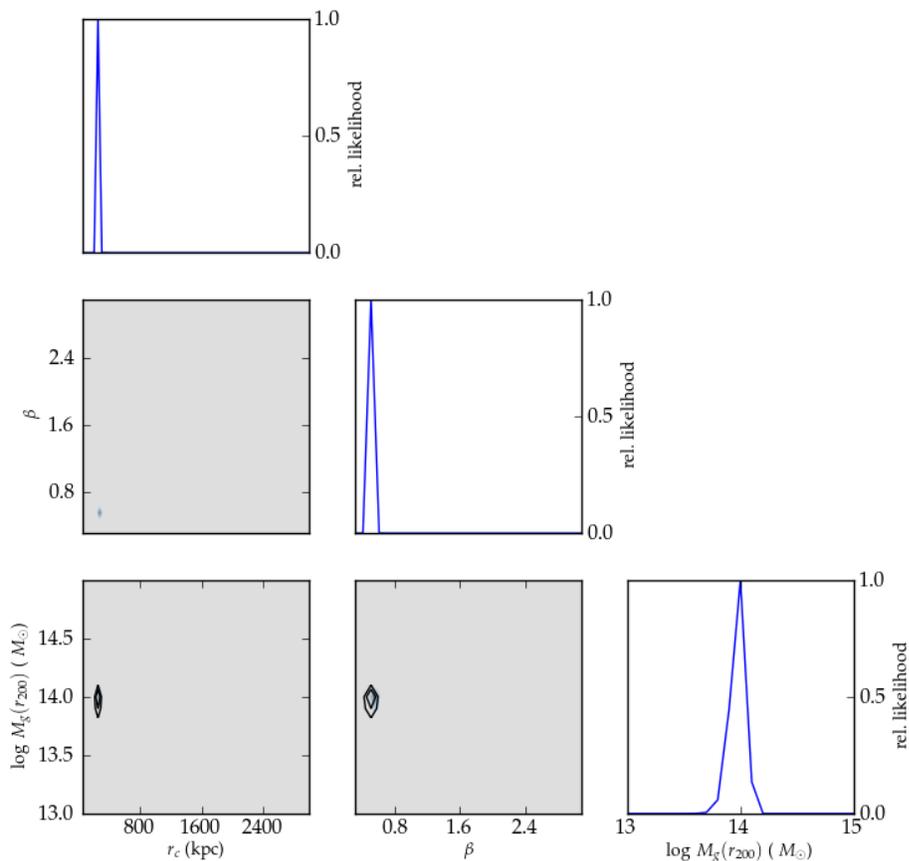


Figure 7: One- and two-dimensional marginalized parameter distributions for ACT-CL J0516-5430, incorporating X-ray data from XMM-Newton. The y -axis of the one-dimensional plots gives normalized relative likelihoods, with 1 representing the likelihood of the best-fit parameter values. The contours in the two-dimensional plots are 1σ (68%) and 2σ (95%) confidence intervals (taking into account only statistical uncertainties, not systematic ones). The values of β and r_c have been further constrained from Figure 5, as desired, while the maximum likelihood values remain close to the previously derived ones, indicating general consistency between the two datasets.

regions in Figure 7 all fall within the corresponding 1σ regions in Figure 5: the X-ray and radio data are complementary, not contradictory, indicating that ACT-CL J0516-5430 is well-described by the isothermal β model¹¹.

6.2 Systematic Errors

The marginalized parameter distributions presented above and in Appendix A take into account the statistical uncertainties from the radio and X-ray observations, but they do not incorporate *systematic* errors which arise during the transition from raw data into distributions. The main goal of this analysis was to exhibit the potential of pure SZ observations to constrain cluster parameters, especially when supplemented by X-ray data, rather than actually giving reliable estimates of the parameters themselves. Hence systematic errors, which can be corrected for by more careful or complicated procedures, do not change the overall message given here; nevertheless, it is important to discuss them.

The first major category of systematic errors arises from modeling the clusters using the isothermal β model. The inadequacies of the β model are discussed in detail in Section 2.3.1 above. To summarize, the β model neglects important physical features such as asphericity, lack of hydrostatic equilibrium, and cool cores: the presences of these features causes the model to fail in certain regimes. Hence the β model will be a poor fit in the case of, for example, a violent merger between clusters.

A second category of errors is due to the filtering applied to the ACT maps to derive values of $y(s)$ (see Section 4.1.2 above). As seen in Figure 4 and discussed in Footnote 4 on page 15, removing low- ℓ power from the entire map, with the intent of removing gradients from the CMB, unavoidably results in removing some of the detected power from the cluster itself. A more sophisticated analysis could partially remedy this analysis by explicitly removing power according to the detected CMB power spectrum itself [32].

Finally, the use of the rough scaling relation (47), which assumes $r_{200} \gg r_c$, to fit the radial surface brightness profile, may have introduced systematic errors into the calculation of S_X . Note, however, that since we use a scaling relation rather than attempting to find the normalization theoretically, incorrect overall normalization between the y and S_X maps, which could occur, for example, from excess removal of power on low angular scales, should *not* effect the results.

7 Conclusion

In this paper, we have used data from the ACT radio telescope to extract the radial profiles of clusters detected via the SZ effect. Using these profiles, we have conducted a maximum likelihood analysis to derive maximum likelihood distributions of underlying cluster parameters in the isothermal β model, taking

¹¹See Appendix A for a case where the radio and X-ray data are contradictory, indicating that the β model does a poor job of describing the cluster.

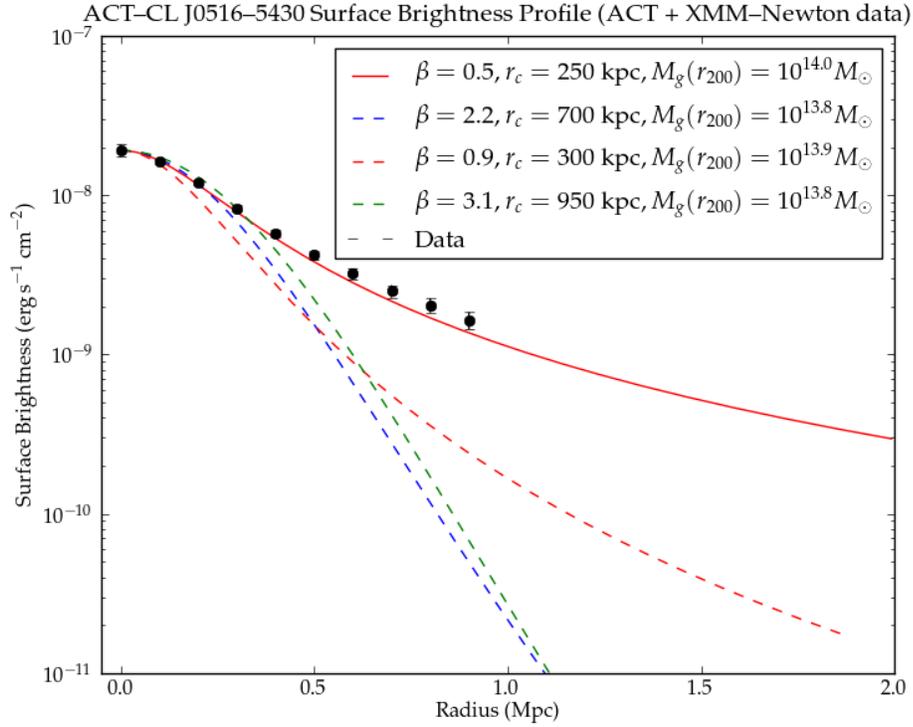


Figure 8: Calculated surface brightness curve for ACT-CL J0516-5430, along with β -model fits. The red solid line, with $\beta = 0.5$, $r_c = 250$ kpc, $M_g(r_{200}) = 10^{14.0} M_\odot$, is the single set of parameters that best fits the combined X-ray and radio data; the total fit has $\chi^2/\text{D.O.F.} = 18.16/13$. The blue, red, and green dashed lines are the three curves previously plotted in Figure 6; with the addition of the X-ray data they fall off too fast and are highly disfavored, with total $\chi^2/\text{D.O.F.}$ of 527.6/13, 851.1/13, and 385.7/13, respectively. Note that the blue and green dashed curves, with high β and r_c , are disfavored to a lesser extent than the red dashed curve, with low β and r_c : this is because their better values at small distances, where the error bars are smaller in log space, partially compensate for their faster falloff at high distances, which have larger log-space errors.

the first steps towards conducting a similar analysis to that of the AMI collaboration [35] on the entire set of ACT clusters. Where it was available, data from X-ray observations was used to supplement the SZ data and further narrow the parameter space. As more data is collected in the X-ray and optical regimes, it can be used to break degeneracies present in the radio observations and provide precise parameter estimates for each cluster.

This analysis also illustrates that, in cases when observation across the electromagnetic spectrum is unavailable, radio observations of the CMB alone have the power to significantly constrain cluster properties, though of course the allowed regions are further narrowed by adding additional datasets. In the case of high-redshift clusters, which are cosmologically most interesting, observation via the Sunyaev-Zel'dovich effect, which is redshift-independent, is the only possible means of detection. The development and refinement of methods, such as the one exhibited here, to extract detailed information about such clusters offers a unique opportunity to learn about the universe on the largest scales¹².

References

- [1] N. Jarosik *et al.*, *Astrophys. J. Suppl. S.* **192**, 14 (2011).
- [2] E. Komatsu *et al.*, *Astrophys. J. Suppl.* **192**, 18 (2011).
- [3] A. Vikhlinin *et al.*, *Astrophys. J.* **692**, 1033 (2009).
- [4] A. Cavaliere and R. Fusco-Femiano, *Astron. Astrophys.* **70**, 677 (1978).
- [5] I. R. King, *Astrophys. J.* **174**, L123 (1972).
- [6] M. Arnaud *et al.*, *Astron. Astrophys.* **517**, A92 (2010).
- [7] E. Komatsu *et al.*, *Astrophys. J. Suppl.* **198**, 18 (2011).
- [8] J. F. Navarro, C. S. Frenk, and S. D. M. White, *Astrophys. J.* **462**, 563 (1996).
- [9] D. Nagai *et al.*, *Astrophys. J.* **668**, 1 (2007).
- [10] Y. A. Zel'dovich and R. A. Sunyaev, *Astrophys. Space Sci.* **4**, 301 (1969).
- [11] A. S. Kompaneets, *Zh. Eksp. Teor. Fiz.* **31**, 876 (1956). Translation in *Sov. Phys. JETP* **4**, 730 (1957).
- [12] C. L. Sarazin, *X-Ray Emission from Clusters of Galaxies* (Cambridge University Press, Cambridge, England, 2009).
- [13] A. Cavaliere, A. Lapi, and Y. Rephaeli, *Astrophys. J.* **634**, 784 (2005).

¹²I am grateful to Jack Hughes for providing me with the X-ray data, and to the entire ACT collaboration for providing feedback on the analysis presented here. I am especially grateful to Lyman Page for concieving the original idea and providing helpful feedback at every stage.

- [14] AMI Consortium: C. Rodríguez-González *et al.*, arXiv:1011.0325.
- [15] T. Bayes and R. Price, *Phil. Trans. R. Soc. Lond.* **53**, 370 (1763).
- [16] G. L. Bryan and Michael L. Norman, *Astrophys. J.* **495**, 80 (1998).
- [17] L. Wang and P. J. Steinhardt, *Astrophys. J.* **508**, 483 (1998).
- [18] M. Birkinshaw, *Phys. Rep.* **310**, 97 (1999).
- [19] S. Frazier, <http://www.physics.princeton.edu/act/papers/sfrazierJP.pdf> (2010).
- [20] S. Majumdar and J. J. Mohr, *Astrophys. J.* **613**, 41 (2004).
- [21] K. Nakamura *et al.* (Particle Data Group), *J. Phys.* **G37**, 075021 (2010), Chapter 19.
- [22] S. Weinberg, *Cosmology* (Oxford University Press, Oxford, England, 2008).
- [23] D. S. Swetz *et al.*, arXiv:1007.0290.
- [24] T. A. Marriage *et al.*, arXiv:1010.1065.
- [25] J. W. Fowler *et al.*, *Astrophys. J.* **722**, 1148 (2010).
- [26] R. Dünner, Ph.D. thesis, Pontificia Universidad Católica de Chile (2009).
- [27] G. Hinshaw *et al.*, *Astrophys. J. Suppl. S.* **180**, 225 (2009).
- [28] S. Das and T. A. Marriage, <http://www.astro.princeton.edu/~act/flipper/doxygen-0.1.5/index.html> (2009).
- [29] F. Menanteau *et al.*, *Astrophys. J.* **723**, 1523 (2010).
- [30] G. P. Garmire *et al.*, *Soc. Photo.-Opt. Instru.* **4851**, 28 (2003).
- [31] M. J. L. Turner *et al.*, *Astron. Astrophys.* **365**, L27 (2001).
- [32] S. Das *et al.*, *Astrophys. J.* **729**, 62 (2011).
- [33] S. Molendi and F. Pizzolato, *Astrophys. J.* **560**, 194 (2001).
- [34] D. S. Hudson *et al.*, *Astron. Astrophys.* **513**, A37 (2010).
- [35] AMI Consortium: T. W. Shimwell *et al.*, arXiv:1101.5590.

A ACT-CL J0658-5557: A Second Cluster with X-Ray Data

As mentioned in 4.2, ACT-CL J0658-5557 (also known as 1E0657-56 or the Bullet cluster) was observed by the ACIS camera on Chandra. Figure 9 shows the marginalized parameters and y -parameter profile using only ACT data; Figure 10 shows the marginalized parameters and surface brightness profile from the combined ACT and Chandra data. The Bullet cluster is very bright in the radio, so ACT data alone is able to constrain the total mass and core radius well, though the β - r_c degeneracy still persists. However, the Chandra data reveals that the Bullet cluster is extremely morphologically disturbed (as its name, given because it is comprised of two rapidly merging clusters, suggests); the surface brightness profile cannot be fit to a β -model profile, and the resulting β -parameter distributions are very narrow, consisting of parameter sets which give $\chi^2/D.O.F.$ of around 300/12 (!!!) instead of even worse fits.

ACT-CL J0658-5557 Marginalized Parameters (ACT data only)

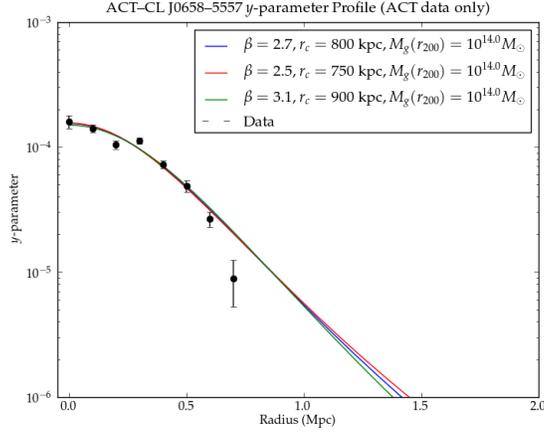
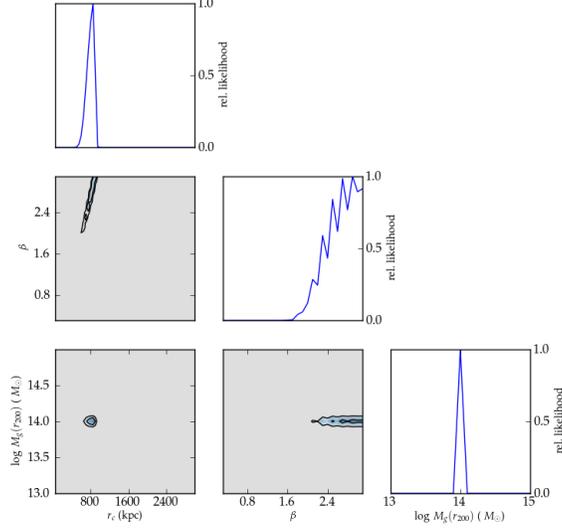


Figure 9: (top) Marginalized parameter distributions for ACT-CL J0658-5557 ($z = 0.296$). The cluster gas mass and core radius are well-constrained, but β is strongly degenerate with r_c . (bottom) Calculated y -parameter radial profile for ACT-CL J0658-5557, along with several of the best-fitting β -model curves. The blue curve is the single set of parameters that best fits the data; it has $\chi^2/\text{D.O.F.} = 26.67/5$. The red and green curves have $\chi^2/\text{D.O.F.} = 26.97/5$ and $27.04/5$, respectively. The reduced χ^2 for J0658-5557 are much higher than those for J0516-5430 (see figure 6) both because the former was observed more often, reducing the absolute error by a factor of several, and because it has higher absolute y -parameter values, further reducing the logarithmic size of the error. The mass $M_g(r_{200})$ is well-constrained by the overall normalization of the data. However, the slow fall-off in the data, and corresponding preference for high r_c , mean that it is very hard to distinguish between differing values of β .

ACT-CL J0658-5557 Marginalized Parameters (ACT + Chandra data)

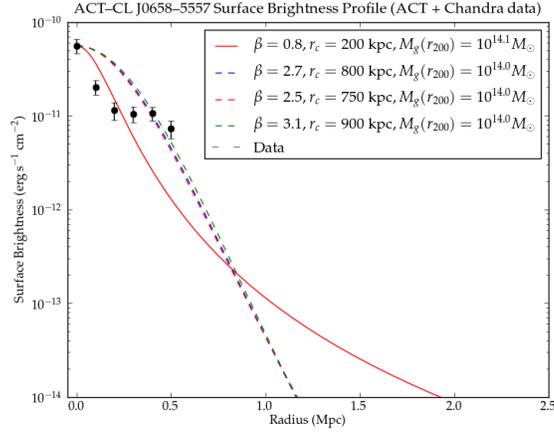
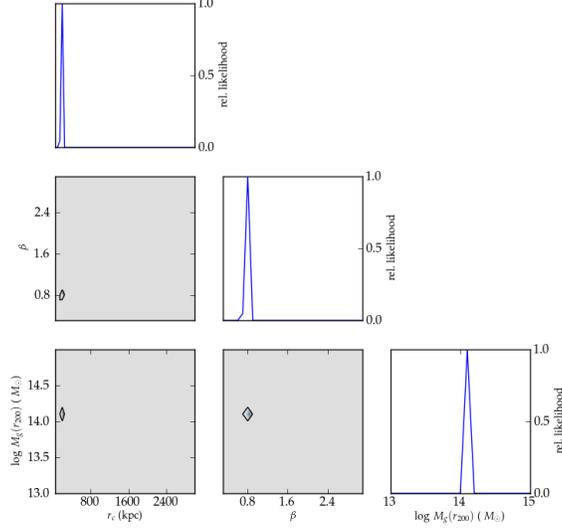


Figure 10: (top) Marginalized parameter distributions for ACT-CL J0516-5430, incorporating X-ray data from XMM-Newton. All parameter values are extremely constrained. (bottom) Calculated surface brightness curve for ACT-CL J0658-5557, along with β -model fits. It is clear that the data cannot be fit well by a β -model curve, which is unsurprising because the Bullet cluster is known to be extremely morphologically disturbed. The red solid line, with $\beta = 0.8$, $r_c = 200$ kpc, $M_g(r_{200}) = 10^{14.1} M_\odot$, is the single set of parameters that best “fits” the combined X-ray and radio data; the total fit has a staggering $\chi^2/\text{D.O.F.} = 340.4/12$. The blue, red, and green dashed lines are the three curves previously plotted in Figure 6; the rapid falloff of the first three points implies a small core radius and thus extremely disfavors them, with even worse total $\chi^2/\text{D.O.F.}$ of 789.7/12, 752.1/12, and 870.0/12, respectively.

B Derivation of $\frac{\partial n_\gamma}{\partial y}$

In Section 3.1 we found the Kompaneets equation in terms of y , the Compton y -parameter (30):

$$\frac{\partial n_\gamma}{\partial y} = \frac{1}{x^2} \frac{\partial}{\partial x} \left[x^4 \frac{\partial n_\gamma}{\partial x} \right], \quad (68)$$

where n_γ , the Planck function, is given by (31)

$$n_\gamma = \frac{1}{e^x - 1}. \quad (69)$$

Substitution into the first equation gives

$$\frac{\partial n_\gamma}{\partial y} = \frac{1}{x^2} \frac{\partial}{\partial x} \left[x^4 \frac{\partial}{\partial x} \left(\frac{1}{e^x - 1} \right) \right] = \frac{1}{x^2} \frac{\partial}{\partial x} \left[\frac{-x^4 e^x}{(e^x - 1)^2} \right]. \quad (70)$$

Performing the differentiation gives

$$\frac{\partial n_\gamma}{\partial y} = \frac{1}{x^2} \left(\frac{-4x^3 e^x}{(e^x - 1)^2} + \frac{-x^4 e^x}{(e^x - 1)^2} + \frac{2x^4 (e^x)^2}{(e^x - 1)^3} \right) = \frac{x e^x}{(e^x - 1)^2} \left(-4 - x + \frac{2x e^x}{e^x - 1} \right). \quad (71)$$

But the term in parentheses can be rewritten as

$$-4 - x + \frac{2x e^x}{e^x - 1} = \frac{-x e^x + x + 2x e^x}{e^x - 1} - 4 = \frac{e^{-\frac{x}{2}} (x + x e^x)}{e^{\frac{x}{2}} - e^{-\frac{x}{2}}} - 4, \quad (72)$$

which simplifies to

$$\frac{x (e^{-\frac{x}{2}} + e^{\frac{x}{2}})}{e^{\frac{x}{2}} - e^{-\frac{x}{2}}} - 4 = \frac{x \cosh\left(\frac{x}{2}\right)}{\sinh\left(\frac{x}{2}\right)} - 4 = x \coth\left(\frac{x}{2}\right) - 4. \quad (73)$$

So the Kompaneets equation becomes

$$\frac{\partial n_\gamma}{\partial y} = \frac{x e^x}{(e^x - 1)^2} \left(x \coth\left(\frac{x}{2}\right) - 4 \right). \quad (74)$$

This completes the derivation of equation (32).

To derive equation (33), note that (74) can be rewritten as

$$\frac{\partial n_\gamma}{\partial y} = \frac{x e^x n_\gamma}{e^x - 1} \left(x \coth\left(\frac{x}{2}\right) - 4 \right), \quad (75)$$

so

$$\frac{\Delta n_\gamma}{n_\gamma} = y \frac{x e^x}{e^x - 1} \left[x \coth\left(\frac{x}{2}\right) - 4 \right], \quad (76)$$

where we have used the approximation $\frac{\partial n_\gamma}{\partial y} = \frac{\Delta n_\gamma}{y}$, valid in the limit of small y . This is precisely equation (33).