# Measurement of Galaxy Cluster Integrated Comptonization and Mass Scaling Relations with the South Pole Telescope

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

| As Published | http://dx.doi.org/10.1088/0004-637x/799/2/137 |
| Publisher | IOP Publishing |
| Version | Final published version |
| Accessed | Sun Dec 30 15:24:59 EST 2018 |
| Citable Link | http://hdl.handle.net/1721.1/96786 |
| Terms of Use | Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use. |
| Detailed Terms |  |
MEASUREMENT OF GALAXY CLUSTER INTEGRATED COMPTONIZATION AND MASS SCALING RELATIONS WITH THE SOUTH POLE TELESCOPE


1 Physics Department, Center for Education and Research in Cosmology and Astrophysics, Case Western Reserve University, Cleveland, OH 44106, USA; benjamin.saliwanchik@case.edu
2 University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
3 Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA
4 Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
5 Kavli Institute for Cosmological Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
6 Enrico Fermi Institute, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
7 Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
8 Department of Physics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
9 Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA
10 Department of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81679 Munich, Germany
11 Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany
12 Department of Physics and Astronomy, University of Missouri, 5110 Rockhill Road, Kansas City, MO 64110, USA
13 Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA
14 NIST Quantum Devices Group, 325 Broadway, Mailcode 817.03, Boulder, CO 80305, USA
15 Departamento de Astronomía y Astrofísica, Pontificia Universidad Catolica, Chile
16 Department of Physics, McGill University, 3600 Rue University, Montreal, QC H3A 2T8, Canada
17 Astronomy Department, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA
18 Department of Physics, University of Illinois at Urbana-Champaign, 1110 W. Green Street, Urbana, IL 61801, USA
19 Department of Physics, University of California, Berkeley, CA 94720, USA
20 Department of Astronomy, University of Florida, Gainesville, FL 32611, USA
21 Department of Astrophysical and Planetary Sciences and Department of Physics, University of Colorado, Boulder, CO 80309, USA
22 Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA
23 Department of Physics, Stanford University, 452 Lomita Mall, Stanford, CA 94305, USA
24 Department of Physics, University of California, One Shields Avenue, Davis, CA 95616, USA
25 Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
26 California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA
27 Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA
28 Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
29 Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, USA
30 Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, D-85748 Garching, Germany
31 Physics Department, University of Minnesota, 116 Church Street S.E., Minneapolis, MN 55455, USA
32 Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
33 Liberal Arts Department, School of the Art Institute of Chicago, 112 S Michigan Avenue, Chicago, IL 60603, USA
34 Institute of Geophysics and Planetary Physics, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA
35 Dunlap Institute for Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
36 Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada
37 Berkeley Center for Cosmological Physics, University of California, Berkeley, CA 94720, USA

Received 2013 December 10; accepted 2014 November 20; published 2015 January 22

ABSTRACT

We describe a method for measuring the integrated Comptonization ($YSZ$) of clusters of galaxies from measurements of the Sunyaev–Zel’dovich (SZ) effect in multiple frequency bands and use this method to characterize a sample of galaxy clusters detected in the South Pole Telescope (SPT) data. We use a Markov Chain Monte Carlo method to fit a $\beta$-model source profile and integrate $YSZ$ within an angular aperture on the sky. In simulated observations of an SPT-like survey that include cosmic microwave background anisotropy, point sources, and atmospheric and instrumental noise at typical SPT-SZ survey levels, we show that we can accurately recover $\beta$-model parameters for inputted clusters. We measure $YSZ$ for simulated semi-analytic clusters and find that $YSZ$ is most accurately determined in an angular aperture comparable to the SPT beam size. We demonstrate the utility of this method to measure $YSZ$ and to constrain mass scaling relations using X-ray mass estimates for a sample of 18 galaxy clusters from the SPT-SZ survey. Measuring $YSZ$ within a 0.75 radius aperture, we find an intrinsic log-normal scatter of 21% ± 11% in $YSZ$ at a fixed mass. Measuring $YSZ$ within a 0.3 Mpc projected radius (equivalent to 0.75 at the survey median redshift $z = 0.6$), we find a scatter of 26% ± 9%. Prior to this study, the SPT observable found to...
have the lowest scatter with mass was cluster detection significance. We demonstrate, from both simulations and SPT observed clusters that $Y_{SZ}$ measured within an aperture comparable to the SPT beam size is equivalent, in terms of scatter with cluster mass, to SPT cluster detection significance.

Key words: galaxies: clusters: general – methods: data analysis – X-rays: galaxies: clusters

1. INTRODUCTION

Galaxy clusters are the largest gravitationally collapsed systems in the observed universe, and their abundance as a function of mass and redshift is a sensitive probe of the growth of structure in the universe. The ability to accurately and precisely estimate cluster masses is essential for using them to constrain cosmological parameters. Typically this is done through cluster observables, which do not directly measure cluster mass, but can be related to it through scaling relations (Vikhlinin et al. 2009b; Rozo et al. 2010; Mantz et al. 2010; Benson et al. 2013). The Sunyaev–Zel’dovich (SZ) effect (Sunyaev & Zel’dovich 1972) is caused by the inverse Compton scattering of cosmic microwave background (CMB) photons off of hot intra-cluster gas. It is a measure of the line-of-sight integral of the cluster pressure and is expected to be a low scatter proxy for cluster mass (Carlstrom et al. 2002; Kravtsov et al. 2006). In particular, the integrated Comptonization of a cluster, $Y_{SZ}$, is expected to have a low intrinsic scatter with cluster mass and to be relatively insensitive to cluster astrophysics (Barbosa et al. 1996; Holder & Carlstrom 2001; Motl et al. 2005; Nagai et al. 2007; Fabjan et al. 2011).

However, for SZ observations where the cluster size is on the order of the instrument beam size or smaller, there is typically a degeneracy in the constraints on the amplitude and shape of the assumed cluster profile (e.g., Benson et al. 2004; Planck Collaboration et al. 2011a, 2013a). One potential way of handling this degeneracy is to employ a Bayesian or Monte Carlo analysis method. A number of experiments have used Monte Carlo methods to characterize the profiles of galaxy clusters in recent years, including Bolocam (Sayers et al. 2013), the Arcminute Microkelvin Imager (AMI; AMI Consortium et al. 2012a, 2012b), the Planck Collaboration (Planck Collaboration et al. 2012; Planck and AMI Collaborations et al. 2013), and the Atacama Cosmology Telescope (ACT) and Sunyaev–Zel’dovich Array (SZA; Reese et al. 2012). In this work, we present a Markov Chain Monte Carlo (MCMC) analysis method for analyzing observations of the SZ effect, which measures $Y_{SZ}$ while marginalizing other SZ model parameters. A feature of the MCMC method is that the $Y_{SZ}$ estimates it produces are well constrained even for clusters with relatively small radii on the sky. This method is related to the method presented in T. E. Montroy et al. (2015, in preparation), which uses the same likelihood, but employs a rapid grid method to directly evaluate the likelihood throughout the parameter space.

We apply this MCMC method to simulated and real observations from the South Pole Telescope (SPT). Various experiments have examined scaling relations between SZ signal and optical or X-ray data. The Planck Collaboration has examined $Y_{SZ}-L_X$ scaling relations (Planck Collaboration et al. 2011b) and $Y_{SZ}-M_X$ scaling relations (Planck Collaboration et al. 2013b). Sehgal et al. (2011), Sifon et al. (2013), and Hasselfield et al. (2013) investigate the scaling between central Comptonization, $n_0$, or $Y_{SZ}$ and mass for ACT clusters. Previous analyses of clusters observed in the SPT-SZ survey used the cluster detection significance, $\xi$, as a proxy for cluster mass (Vanderlinde et al. 2010; Andersson et al. 2011; Benson et al. 2013; Reichardt et al. 2013). Here we show that $Y_{SZ}$ integrated over a fixed angular aperture near the SPT beam size and $\xi$ have comparable fractional scatter in their respective mass scaling relations. $Y_{SZ}$, however, is more easily compared to cluster parameters derived from other measurements.

2. CLUSTER SAMPLE AND OBSERVATIONS

2.1. SZ Observations

The SPT is a 10 m diameter off-axis Gregorian telescope with a 1 deg$^2$ field of view, designed to operate at millimeter and submillimeter wavelengths (Carlstrom et al. 2011). In 2007–2011, the SPT surveyed 2500 deg$^2$ in three frequency bands centered at 95, 150, and 220 GHz. This survey is referred to as the SPT-SZ survey. The cluster sample used in this work is drawn from the two fields (23 deg$^2$ each) observed with the SPT in 2008, one centered at right ascension (R.A.) 5$^h$30$^m$, declination (decl.) $-55^\circ$ (J2000), and one at R.A. 23$^h$30$^m$, decl. $-55^\circ$. A nearly identical cluster sample was used in Vanderlinde et al. (2010, hereafter V10), Andersson et al. (2011, hereafter A11), and Benson et al. (2013, hereafter B13).

Observing procedures, data processing, and detection algorithms for these clusters are described in detail in V10 and Staniszewski et al. (2009) and are summarized here. Details of the data processing pipeline are also described in Schaffer et al. (2011).

Each field was observed by scanning the telescope back and forth in azimuth at 0.25 s$^{-1}$, and then stepping in elevation and repeating until the entire field was covered. This process covers a 100 deg$^2$ field in approximately two hours Field scans were repeated several hundred times until the noise in the co-added maps reached a completion depth of 18$\mu$Karcmin for 150 GHz. (See Staniszewski et al. 2009, V10, or Williamson et al. 2011 for a description of field depth measurements.) The timesreams of the individual detectors were filtered to remove sky signal that was spatially correlated across the focal plane and long timescale detector drift. The combination of these filters effectively removes signals with angular scales larger than $\lesssim$0.5. Data from individual detectors were combined using inverse-variance weighting, and the resulting maps were calibrated by comparison to the Wilkinson Microwave Anisotropy Probe (WMAP) 5 yr CMB temperature anisotropy power spectrum (Lueker et al. 2010).

2.2. Cluster Detection

Clusters are identified in the SPT maps using a matched filter (MF; Haehnelt & Tegmark 1996; Herranz et al. 2002a, 2002b; Melin et al. 2006). Specifics on this procedure can be found in Staniszewski et al. (2009) and V10 for single frequency cluster detection, and in Williamson et al. (2011) and Reichardt et al. (2013) for multi-frequency detection. To locate clusters, the maps are multiplied in Fourier space with a filter matched to the expected spatial signal-to-noise (S/N) profile of galaxy clusters. The MF, $\psi$, is given by

$$\psi(k_x, k_y) = \frac{B(k_x, k_y)S(|\vec{k}|)}{B(k_x, k_y)^2N_{astro}(|\vec{k}|)+N_{noise}(k_x, k_y)},$$  

(1)
where $B$ is the instrument response after timstream filtering, $S$ is the source template, and the noise has been divided into astrophysical ($N_{\text{astro}}$) and terrestrial ($N_{\text{ter}}$) components. $N_{\text{astro}}$ includes power from lensed primary CMB anisotropies, an SZ background from faint undetected clusters, and millimeter-wave emitting point sources. The noise power spectrum $N_{\text{ter}}$ includes atmospheric and instrumental noise, estimated from jackknife maps. The source template is a two-dimensional projection of an isothermal $\beta$-model, with $\beta$ set to 1 (Cavaliere & Fusco-Femiano 1976):

$$\Delta T = \Delta T_0 (1 + \theta^2 / \theta_c^2)^{-1},$$

where the central SZ temperature decrement $\Delta T_0$ and the core radius $\theta_c$ are free parameters.

Clusters are detected using a (negative) peak detection algorithm similar to SExtractor (Bertin & Arnouts 1996). The significance of a detection, $\xi$, is defined to be the highest S/N ratio across all $\theta_c$. In our analysis, we use the unbiased significance $\xi = \sqrt{\langle \xi \rangle^2 - 3}$, where $\langle \xi \rangle$ is the average detection significance of a cluster across many noise realizations (V10).

It is important to note here that maximizing the S/N in the MF maximizes the likelihood used in the MCMC method below (Section 3) at a fixed $\theta_c$. The two methods are exploring the same likelihood, the difference is that the MF is optimized for detecting clusters not parameterizing them, since it does not simultaneously explore all dimensions in the parameter space.

2.3. Optical and X-Ray Observations

The optical and X-ray observations for the clusters used in this work have previously been described in A11 and B13, which we briefly describe here. All 18 clusters have redshift measurements, 15 of which are spectroscopic, and 14 of the clusters have X-ray measurements.

Optical $griz$ imaging and photometric redshifts for these clusters were obtained from the Blanco Cosmology Survey (Desai et al. 2012), and from pointed observations using the Magellan telescopes (High et al. 2010). Of the 15 clusters with spectroscopic redshifts, 8 were obtained by the Low Dispersion Survey Spectrograph (LDSS3) on the Magellan Clay 6.5 m telescope (High et al. 2010), and 1 by the Inamori Magellan Areal Camera and Spectrograph (IMACS) on the Magellan Baade 6.5 m telescope (Brodwin et al. 2010). The final six cluster redshifts were measured with IMACS and GMOS on Gemini South (Ruel et al. 2013). X-ray follow-up observations were performed with Chandra ACIS-I and XMM-Newton EPIC (A11; B13).

3. MCMC ANALYSIS METHODS

In this work, we follow a Bayesian approach to parameter estimation, using a MCMC method to estimate the parameters of our cluster source model. The application of MCMC methods to the detection and characterization of compact astrophysical sources in noisy backgrounds was proposed by Hobson & McLachlan (2003), and several experiments have used MCMC methods for parameterizing SZ signals from galaxy clusters. Bonamente et al. (2004), Bonamente et al. (2006), and LaRoque et al. (2006) used MCMC methods to analyze SZ data from BIMA and OVRO, in conjunction with X-ray data from Chandra, and fit $\beta$-model profiles to galaxy clusters. Muchovej et al. (2007), Culverhouse et al. (2010), and Marrone et al. (2009) parameterized SZA clusters, and Halverson et al. (2009) parameterized the Bullet Cluster using APEX-SZ data, all using the $\beta$-model. Culverhouse et al. (2010), Marrone et al. (2009), and Marrone et al. (2012) additionally estimated cluster $Y_{\text{SZ}}$ values.

In recent years there has been a surge of interest in MCMC methods for parameterizing cluster from a number of experiments, including Bolocam (Sayers et al. 2013), AMI (AMI Consortium et al. 2012a, 2012b), the Planck Collaboration (Planck Collaboration et al. 2012; Planck and AMI Collaborations et al. 2013), and ACT and SZA (Reese et al. 2012). Here we estimate galaxy cluster $Y_{\text{SZ}}$ values and $Y_{\text{SZ}}-M$ scaling relations in addition to estimating $\beta$-model parameters.

3.1. Posterior Distribution for a Compact Source

We use a Metropolis–Hastings algorithm implementation of the MCMC method for parameter estimation (Metropolis et al. 1953; Hastings 1970). For the case of a compact object with source template $S(\mathcal{H})$ in a two-dimensional astronomical data set $D$ with Gaussian noise, the Bayesian likelihood has the form

$$P(D|\mathcal{H}) = \frac{\exp(-\frac{1}{2}(D-S(\mathcal{H}))C^{-1}(D-S(\mathcal{H})))}{(2\pi)^{N_{\text{pix}}/2}|C|^{1/2}},$$

where $C$ is the noise covariance matrix for the data set $D$, and $N_{\text{pix}}$ is the number of pixels in $D$ (Hobson & McLachlan 2003).

In this method, $C$ is composed of the combined $N_{\text{astro}}$ and $N_{\text{ter}}$ noise terms in the MF $\psi$ (Equation (1)).

We are interested in parameterizing galaxy clusters using the SZ effect, which is the spectral distortion they produce in the blackbody CMB spectrum. At two of the SPT’s observing frequencies (95 and 150 GHz), this distortion is manifested as a decrement in CMB power, while the net change in CMB power at 220 GHz is negligible.

Equation (3) is easily generalizable to the case of astronomical images in multiple frequency bands, where the unnormalized log likelihood may be calculated in the Fourier domain as

$$\log(P(D|\mathcal{H})) = -\frac{1}{2} \sum_{k,v_i} \left[ \tilde{N}_{v_i}^{-1}(k) \tilde{D}_{v_i}(k) - \tilde{S}_{v_i}^2(k) \right],$$

where $\tilde{k}$ is the two-dimensional Fourier space vector, $\tilde{D}_{v_i}(k)$ is the Fourier transform of the map for frequency $v_i$, $\tilde{S}_{v_i}^2(k)$ is the frequency-dependent Fourier transform of the cluster model for parameter set $\mathcal{H}$, and $N_{v_i}(k)$ is the frequency-dependent covariance matrix for the $v_i$ and $v_j$ frequency maps. Here $N_{v_i}(k)$ is simply the multiband extension of the covariance matrix $C$ in Equation (3). It is convenient to perform the likelihood calculations in Fourier space rather than physical space because $N_{v_i}(k)$ is diagonal in Fourier space, assuming stationary noise.

The frequency-dependent covariance matrix is computed as follows. For a given $k$, the two frequency matrix for CMB+Noise covariance is

$$N_{v_i,v_j}(k) = \begin{bmatrix} C(k)B_1^2(k) & C(k)B_1(k)B_2(k) & C(k)B_1(k)B_2(k) & C(k)B_2(k)^2 + N_2(k) \\ C(k)B_1(k)B_2(k) & C(k)B_1(k)B_2(k) & C(k)B_2(k)^2 + N_2(k) & C(k)B_2(k) \\ C(k)B_1(k)B_2(k) & C(k)B_1(k)B_2(k) & C(k)B_2(k)^2 + N_2(k) & C(k)B_2(k) \\ C(k)B_1(k)B_2(k) & C(k)B_1(k)B_2(k) & C(k)B_2(k)^2 + N_2(k) & C(k)B_2(k) \end{bmatrix},$$

where $C(k)$ is the CMB power at $k$, $B_1(k)$ and $N_1(k)$ are the beam and noise for the first frequency, and $B_2(k)$ and $N_2(k)$ are the beam and noise for the second frequency.
The covariance becomes more complicated when including point sources and the SZ background because the signal component has a different magnitude at different frequencies. If \( Q_1(\hat{k}) \) and \( Q_2(\hat{k}) \) are the point source or SZ covariance at two of our observing frequencies, then we have

\[
N_{n_{\nu,\nu'}}(\hat{k}) = \begin{bmatrix}
C(\hat{k})B_1(\hat{k})^2 + N_1(\hat{k}) + B_1(\hat{k})^2 Q_1(\hat{k}) \\
C(\hat{k})B_1(\hat{k})B_2(\hat{k}) + B_1(\hat{k})B_2(\hat{k})\sqrt{Q_1(\hat{k})Q_2(\hat{k})} \\
C(\hat{k})B_1(\hat{k})B_2(\hat{k}) + B_1(\hat{k})B_2(\hat{k})\sqrt{Q_1(\hat{k})Q_2(\hat{k})} \\
C(\hat{k})B_2(\hat{k})^2 + N_2(\hat{k}) + B_2(\hat{k})^2 Q_2(\hat{k})
\end{bmatrix}.
\] (6)

Before the log-likelihood is calculated, to account for the filtering of the data set \( D_\nu(\hat{k}) \) as described in Section 2.1, both the covariance \( N_{n_{\nu,\nu'}}(\hat{k}) \), and the source template \( \tilde{z}_i^\nu(\hat{k}) \), are multiplied by a Fourier space filter function that emulates the timestream filtering used to produce \( D_\nu(\hat{k}) \).

3.2. Implementation

Our MCMC is modeled after the generic Metropolis–Hastings method described in Hobson & McLachlan (2003), and is implemented in MATLAB38.

In this work, we use the MCMC method for cluster parameterization, not detection. Our testing found that it was more computationally costly and not more effective at cluster detection than the MF method. Throughout this work, our MCMC is run over a relatively small area of sky (512 pixels × 512 pixels, or \( \sim 2^\circ \times 2^\circ \)) centered on a cluster that has already been identified.

Cluster parameter recovery is tested in single- and multi-frequency simulations below (Section 5), but we use only 150 GHz when investigating scaling relations (for observed and simulated clusters) to match the SPT cluster analysis in B13, from which our sample is derived. We use the \( \beta \)-model source template given in Equation (2). T. E. Montroy et al. (2015, in preparation) demonstrate, using simulations and methods similar to those described in Section 5, that \( Y_{\text{SZ}} \) is recovered accurately with a \( \beta \)-model for either \( \beta \)-model or Arnaud profile (Arnaud et al. 2010) input clusters.

Clusters, as described by the \( \beta \)-model, are characterized by four parameters: their location on the sky in R.A. and decl., the magnitude of the SZ temperature decrement \( \Delta T_0 \), and the core radius \( \theta_c \). We apply priors in the form of uniform probability distributions in each parameter. Given that we are characterizing clusters that have already been detected by the MF, our position priors can be quite tight. We impose a simple square-box prior on R.A. and decl., centered at the MF cluster location and extending ±1/25. Our \( \Delta T_0 \) and \( \theta_c \) priors restrict these parameters to broadly reasonable values given the expected mass and redshift range of our cluster sample. Our SZ temperature decrement prior is \(-2.5 \text{ mK} \leq \Delta T_0 \leq 0.0 \text{ mK}\), and our radius prior is \(0.025 \leq \theta_c \leq 2.5\). \( \theta_c \) is not allowed to fall to zero for numerical reasons.

For a detailed examination of the SPT beam functions and noise properties, see Schaffer et al. (2011). Figure 2 of that work shows how the SPT beams scale with physical radius and \( \theta_c \). Figure 7 shows the signal-noise and noise PSDs for an SPT map, both from the raw map, and corrected for the beam and transfer functions.

Burn-in, as evaluated by stability of the likelihood values, is typically complete within several hundred steps. For the 12,000 simulated cluster realizations in Section 5, we cut the first 10⁴ steps, using the rest of the 10⁴ steps to characterize the probability surface. In the scaling relation analysis discussed in Section 6 many fewer clusters were analyzed, allowing the chain length to be extended to 10⁵ steps, from which we exclude the first 10⁴ steps in order to ensure convergence. We define recovered parameter values to be the median of the MCMC equilibrium distribution for each parameter, marginalizing over the other parameters. Uncertainties are given by the 68% confidence interval of the marginalized distribution for each parameter, centered on the median value. Figure 1 shows the parameter distributions for a typical cluster detected with the SPT.

4. SIMULATIONS

4.1. Simulated Thermal SZ Cluster Maps

We used two sets of simulations: one uses \( \beta \)-model clusters (defined by \( \Delta T_0 \) and \( \theta_c \)) to investigate cluster parameter recovery (Section 5), while the second uses cluster gas profiles inferred from dark matter light cone simulations to calibrate \( Y_{\text{SZ}}-M \) scaling relations (Section 6). The second set of simulations is described in detail in Shaw et al. (2010) and will be referred to as the S10 simulations for convenience. The thermal SZ (tSZ) cluster profiles used in each set are discussed in more detail in the relevant sections below.

4.2. Astrophysical Backgrounds

We use simulated maps of astrophysical backgrounds that include contributions from the CMB and extragalactic point sources. Simulated CMB anisotropies were generated based on realizations of the gravitationally lensed WMAP 5 yr ΛCDM CMB power spectrum.

The extragalactic point source population at 150 GHz consists of two classes of objects: “dusty” sources dominated by thermal dust emission from star formation bursts, and “radio” sources dominated by synchrotron emission. We use the source count model of Negrello et al. (2007) at 350 GHz, which is based on physical modeling by Granato et al. (2004) for dusty sources. Source counts at 150 GHz are estimated by assuming the flux densities scale as \( S_\nu \propto \nu^\alpha \), where \( \alpha = 3 \) for high-redshift protospheroidal galaxies, and \( \alpha = 2 \) for late-type galaxies. For radio sources we use the De Zotti et al. (2005) model at 150 GHz, which is in agreement with observed radio source populations at \( S < 100 \text{ mJy} \) (Vieira et al. 2010; Marriage et al. 2011; Planck Collaboration et al. 2011a; Marsden et al. 2013; Mocanu et al. 2013).

Point source population realizations were generated by sampling from Poisson distributions for each population in bins with fluxes from 0.01 mJy to 1000 mJy. Sources were randomly distributed across the map. Correlations between sources or with galaxy clusters were not modeled, following V10. These 150 GHz simulated point source populations were used for the scaling relation simulations of Section 6, but not for the multi-band pipeline checks of Section 5.

4.3. Simulated Observations

Ideally, we could emulate the SPT transfer function for the 95 GHz and 150 GHz frequency bands by producing

---

38 Mathworks Inc., Natick, MA 01760, USA.
synthetic timestreams from simulated maps convolved with the SPT beam, observing them using the SPT scan strategy, and convolving the resulting timestreams with detector time constants. We would then produce maps by processing the simulated timestreams as in Section 2.1. However, this full emulation of the SPT transfer function is computationally intensive. To simplify this process and produce a large number of sky maps, we model the transfer function as a two-dimensional Fourier filter. V10 show that this approximation introduces systematic errors in the recovered cluster $\xi$ values of less than 1%.

The instrumental and atmospheric noise in the SPT maps were estimated by creating difference maps, which were constructed to have no astrophysical signal. Each field consists of several hundred individual observations. We randomly multiply half of the observations by $-1$, and then coadd the full set of observations. We repeat this several hundred times, each time calculating the two-dimensional spatial power spectrum, which we average to estimate the instrumental and atmospheric noise in the coadded SPT map. This averaged noise spectrum is used to generate random map realizations of the SPT noise, which are added to the simulated maps.

5. PIPELINE CHECKS

5.1. Cluster Model

We use mock observations of clusters in simulated sky maps to evaluate the accuracy and bias of the recovered cluster parameters. We begin with simulated maps that contain the astrophysical signals described in Section 4.2. To this we add mock clusters with an assumed $\beta$-model profile, with known SZ decrements and radii, at specified locations. Simulated SPT observations are then performed on these maps. Three different cluster core radii ($0.25, 0.5, \text{and } 1.0$) are used, combined with eight values for peak Comptonization between 175 $\mu$K and 2 mK, spanning the range of values typically found for SPT-detected clusters with $\xi > 5$. These cluster profiles are convolved with the SPT transfer function, and then placed in the simulated maps. For each combination of $\beta$-model cluster parameters, we create five pairs of simulated maps (150 GHz and 95 GHz) by placing 100 copies of the cluster at random locations in five unique noise maps. This results in 500 noise realizations for each combination of cluster parameters, or 12,000 clusters total. As usual, $\sim 2^\circ \times 2^\circ$ cutouts are made around each cluster, and the MCMC is run on each patch.

In Sections 5.2 and 5.3, we test parameter recovery in the single-band (150 GHz) and multiband (95 GHz and 150 GHz) cases. The 220 GHz data contain no SZ information, but could in principle be used to remove primordial CMB anisotropy. However, the noise level of the 220 GHz maps, dominated by residual atmospheric emission, is larger than the intrinsic astrophysical confusion caused by CMB anisotropy. Therefore, we do not use the 220 GHz SPT measurements to fit the cluster model.

5.2. Position, Radius, and Amplitude

We first examine the recovered values of the four baseline cluster parameters: the R.A. and decl. position, $\Delta T_0$, and $\theta_c$. The cluster positions are measured accurately, and we find no bias in either position parameter, for both the single-band and multiband cases. For clusters near the SPT beam size ($\sim 1^\prime$ FWHM at 150 GHz) and selection threshold, the amplitude and shape of the cluster will not be well constrained, however the integrated signal within the SPT beam will be. A similar degeneracy has previously been noted in other cluster analyses (e.g., Benson et al. 2004; Planck Collaboration et al. 2011a, 2013a). In Figure 2, we show the recovered $\Delta T_0$ and $\theta_c$. 

![Figure 1](image-url)
distributions for a typical cluster in the SPT catalog (SPT-CL J0533-5005, ξ = 5.59, z = 0.8810), an SPT observed cluster with core radius θc < 1’. While the position is well-constrained, there is a significant degeneracy between the constraints on θc and ΔT0.

5.3. Integrated Comptonization

In general, the cylindrically projected integrated Comptonization of a cluster is calculated by integrating the source function, S(θ), out to a given angular aperture θint:

\[ Y_{SZ} = 2\pi \int_{0}^{\theta_{int}} S(\theta) \theta \, d\theta. \tag{7} \]

For much of this work, θint will be a constant angular aperture. We distinguish this estimator of \( Y_{SZ} \) from others by referring to it as \( Y_{SZ}^\theta \) hereafter.

In the case of a two-dimensional projection of a spherical \( \beta \)-model with \( \beta = 1 \) (Equation (2)), this integral can be solved analytically:

\[ Y_{SZ}^\theta = \frac{\pi \Delta T_0 \theta_c^2}{f_x T_{CMB}} \log \left[ 1 + \left( \frac{\theta_{int}}{\theta_c} \right)^2 \right], \tag{8} \]

where \( \theta_c \) is the core radius in arcminutes, \( \Delta T_0 \) is the central temperature decrement in units of \( K_{CMB} \), the equivalent CMB temperature fluctuation required to produce the observed power fluctuation, \( T_{CMB} \) is the CMB blackbody temperature of 2.725 K, and \( f_x \) is given by

\[ f_x = \left( x \frac{e^x + 1}{e^x - 1} - 4 \right) \left[ 1 + \delta(x, T_c) \right], \tag{9} \]

where \( x = h\nu/K_{CMB} \) and \( \delta(x, T_c) \) accounts for relativistic corrections to the SZ spectrum (Itoh et al. 1998; Nozawa et al. 2000). For details of the calculation of \( f_x \) for the SPT, see A11.

We use this equation to calculate \( Y_{SZ}^\theta \) for every step in the MCMC chain, and thus to produce a marginalized distribution of \( Y_{SZ}^\theta \) values. In these simulations, integration to a radius approximately corresponding to the 150 GHz SPT beam diameter (roughly the range 0.75 < θint < 1.25) produces \( Y_{SZ}^\theta \) distributions that are well constrained despite the degeneracy of \( \theta_c \) and \( \Delta T_0 \), with minimal error in recovered cluster \( Y_{SZ}^\theta \) values. Integration in this section is performed to \( \theta_{int} = 0.75 \), though other values are explored for scaling relations in Section 6 below. Note that we calculate \( Y_{SZ}^\theta \) from the marginalized distribution of the source model parameters, not by integrating the flux on the sky. It is also important to note that the likelihood is calculated (in frequency space) over the full \( 2^\circ \times 2^\circ \) patch of sky, \( \theta_{int} \) is simply the radius out to which the best fit \( \beta \)-model is integrated.

If the redshift of a cluster is known, it is also possible to integrate \( Y_{SZ} \) within an angular aperture corresponding to a specific physical radius, \( \rho \):

\[ \theta_{int} = \rho \, D_A^{-1}(z), \]

where \( D_A(z) \) is the angular diameter distance to the redshift \( z \). In Sections 6 and 7, we examine \( Y_{SZ} \) integrated within a constant physical radius, \( \rho \), for all clusters in a sample. We will refer to this quantity as \( Y_{SZ}^\rho \).

In Figure 2, we show a typical SPT cluster in which \( Y_{SZ}^\rho \) is well constrained despite the degeneracy between \( \Delta T_0 \) and \( \theta_c \). Figure 3 shows \( Y_{SZ}^\theta \) and \( \theta_c \) parameter distributions for 500 runs of a typical simulated cluster with a radius smaller than the SPT 150 GHz beam size (\( \theta_c = 0.5, \Delta T_0 = 300 \mu K, \xi = 6.2 \)). The 68% confidence interval for \( Y_{SZ} \) in these simulations is typically \( \sim 14\% \) of the central value. The cutoff at low \( \theta_c \) is due to the small, but non-zero, minimum priors on \( \theta_c \) and \( \Delta T_0 \), this is not a feature of the data likelihood. Despite only having an upper bound on \( \theta_c \), \( Y_{SZ} \) is still well constrained. We find that \( Y_{SZ} \) is well constrained and consistent regardless of whether the analysis is single-band (150 GHz) or multiband (95 GHz and 150 GHz).

In Figure 4, we show the ratio of the recovered to input \( Y_{SZ} \) as a function of core radius and cluster detection significance, \( \xi \), for 24 different combinations of \( \theta_c \) and \( \Delta T_0 \), each with 500 independent noise realizations. Despite a slight apparent bias for some \( \theta_c \) values, we find no significant bias as a function of the detection significance, and recover \( Y_{SZ} \) accurately to <2% in all cases. On average recovered \( Y_{SZ}^\rho \) values are 0.27% lower than input values, which is below the 0.49% error in the mean.
the line of sight. The resulting maps are projections of all the clusters in the lightcone simulation onto a simulated sky. Forty $10^4 \times 10^5$ maps were produced by this procedure, together with catalogs of cluster masses, redshifts, and positions.

6.2. $Y_{SZ}$–$M$ Scaling Relation Fitting Methods

We assume a scaling between $Y_{SZ}$ and $M$ of the form

$$Y_{SZ} = A_{SZ} \left( \frac{M_{\text{vir}}}{3 \times 10^{14} M_{\odot} h^{-1}} \right)^{B_{SZ}} \left( \frac{E(z)}{E(0.6)} \right)^{C_{SZ}},$$

parameterized by the normalization $A_{SZ}$, the mass scaling $B_{SZ}$, and the redshift evolution $C_{SZ}$, and where $E(z) \equiv H(z)/H_0$. For self-similar evolution, $B_{SZ} = 5/3$ and $C_{SZ} = 2/3$ (e.g., Kravtsov et al. 2006). The pivot points of the scaling relation were defined to match the approximate mean mass and redshift for the SPT cluster sample.

We fit the $Y_{SZ}^0$–$M_{\text{vir}}$ scaling relation by minimizing the fractional scatter, $S$, in $Y_{SZ}$, defined as

$$S = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left( \frac{Y_{\text{recov}}^n - Y_{\text{input}}^n}{Y_{\text{input}}^n} \right)^2},$$

where $Y_{\text{recov}}^n$ is the integrated Comptonization recovered by the MCMC for the $n$th cluster, $Y_{\text{input}}^n$ is the corresponding Comptonization calculated from the input catalog mass and the assumed scaling relation (Equation (10)), and we sum over $N$ simulated clusters. The scaling relation parameters $A_{SZ}$, $B_{SZ}$, and $C_{SZ}$ are varied using a grid search method, and the scatter $S$ is calculated at each point in the parameter space. The combination of parameters that minimizes $S$ is taken to be the best-fit set of parameters. This definition of fractional scatter is used to fit $Y_{SZ}^0$–$M_{\text{vir}}$ scaling relations in Section 6.3 and in Section 6.4.

6.3. Results for Simulated Thermal-SZ-only Maps

We run both the MCMC and MF methods on tSZ-only maps from the S10 dark matter lightcone simulations described in

Figure 3. The marginalized constraints on $Y_{SZ}$ and $\theta_c$ from 500 noise realizations of a typical simulated cluster with radius smaller than the SPT 150 GHz beam ($\theta_c = 0.5$, $\Delta T_0 = 300 \mu K$, $\xi = 6.2$). The contours show the 68% and 95% confidence regions. The cross marks the input $\theta_c$ and $Y_{SZ}$ values ($Y_{SZ}^{\text{input}} = 1.20 \times 10^{-4}$ arcmin$^2$). Despite only having an upper bound on $\theta_c$, $Y_{SZ}$ is well constrained.

Figure 4. Average ratios of recovered to input $Y_{SZ}$ for 24 $\beta$-model source profiles, generated from different combinations of $\theta_c$ and $\Delta T_0$. Each point is the mean recovered $Y_{SZ}$ for a simulated cluster with 500 independent noise realizations. The errorbars represent the error on the mean of these recovered $Y_{SZ}$ values.
Section 6.1. These simulated tSZ maps contain only tSZ signal, and no CMB, point sources, atmospheric noise, or instrumental noise.

We measure the SZ signal in these maps using the methods described in Sections 5.1 and 5.3 for clusters with $M_{\rm vir} > 4 \times 10^{14} M_\odot h^{-1}$, and redshift $0.3 < z < 1.2$. We then use the cluster virial masses and Equation (10) to find the best-fit scaling relation parameters by minimizing the fractional scatter in Equation (11). We do this for both the $Y_{\ SZ} - M_{\ vir}$ and $\zeta - M_{\ vir}$ scaling relations, which allows for direct comparison of these analysis methods. The redshift range corresponds to the redshift range of observed SPT clusters, and the mass criteria correspond to the mass of clusters at the lower SPT significance limit of the Reichardt et al. (2013) cluster catalog ($\xi = 4.5$), at the survey median redshift of $z = 0.6$.

As a baseline for the scatter in the measured $Y_{\ SZ} - M_{\ vir}$ scaling relations for these simulations, we examine the intrinsic scatter between $M_{\ vir}$ and $Y_{\ SZ}$, the contribution to the SZ flux from within the spherical virial radius for each cluster. We fit the $Y_{\ SZ} - M_{\ vir}$ scaling relation parameters using the same method as for measured $Y_{\ SZ}$ values, and find the fractional scatter in the best-fit scaling relation to be 16%.

We fit $Y_{\ Sz}^0 - M_{\ vir}$ relations for a range of angular apertures, $\theta$, with $Y_{\ SZ}^0$ defined in Equation (8). This shows the fractional scatter as a function of the integration angle for angles ranging from 0.25 to 3.0. We find that the fractional scatter in $Y_{\ SZ}^0$ does not vary significantly with angular aperture, with a broad minimum in the scatter at $\sim 0.75 - 1.0$ ($Y_{\ SZ}^{0.75}$). The exact location of the minimum scatter shifts between the tSZ-only and the full-noise S10 maps, but is near 0.75 in both cases (see Figure 5). For simplicity, and for comparison between the different simulated maps and observed clusters, we use the $Y_{\ SZ}^{0.75} - M_{\ vir}$ scaling relation as our nominal scaling relation. The $Y_{\ SZ}^{0.75} - M_{\ vir}$ scaling relation has 23% ± 2% fractional scatter in $Y_{\ SZ}$, which is slightly less than the 27% ± 2% scatter in the $\zeta - M_{\ vir}$ scaling relation for these clusters. The scatter in the tSZ-only simulations is primarily due to the intrinsic scatter in the mass to SZ observable scaling, scatter from the tSZ background is sub-dominant. (Note that the scatter here is fractional scatter, whereas previous SPT analyses in V10 and B13 quoted a log-normal scatter, at a level consistent with the values found in this work.)

Figure 6 shows $Y_{\ SZ}^{0.75}$ versus $M_{\ vir}$ for 1187 mass-selected clusters in the S10 simulated tSZ-only maps, where we only include clusters with $M_{\ vir} > 4 \times 10^{14} M_\odot h^{-1}$ in the redshift range $0.3 < z < 1.2$. Fractional scatter in $Y_{\ SZ}$ is 23% ± 2%. The solid line is the best-fit $Y_{\ SZ}^{0.75} - M_{\ vir}$ scaling relation found for this cluster sample.
to measure $Y_{SZ}$ within a fixed physical radius. In Figure 5, we plot the best-fit scatter for a range of integration radii between 0.1 and 1.0 Mpc. We find that the minimum fractional scatter in $Y_{SZ}$ within a fixed physical radius is higher than the minimum fractional scatter within a fixed angular aperture. For $Y_{SZ}$, the scatter is increased by the varying angular size of the chosen physical radius at different redshifts. The optimal physical radius corresponds roughly to the optimal angular aperture, at the median redshift of the cluster sample, $z = 0.6$. Clusters farther from the median redshift will have integration angles farther from the optimal angle, resulting in relatively higher scatter in $Y_{SZ}$ than in $Y_{θ}$.

As can be seen in Figure 5, we find a broad minimum in scatter at $\sim 0.3-0.4$ Mpc, with a minimum scatter of $27\% \pm 3\%$. This is comparable to the $\zeta - M_{vir}$ relation for these clusters, and slightly higher than the scatter in the $Y_{SZ}^{0.75} - M_{vir}$ scaling relation. The scaling relation parameters for $Y_{SZ}$ within 0.3 Mpc ($Y_{SZ}^{0.3 \text{ Mpc}}$; 0.3 Mpc being equivalent to 0.75 at $z = 0.6$) are given in Table 1.

### 6.4. Results for Full-noise Simulated Maps

We also fit $Y_{SZ}^{\rho} - M_{vir}$ scaling relations for the simulated clusters in the presence of other astrophysical and instrumental noise components (see Sections 4.2 and 4.3). The same cluster sample ($M_{vir} > 4 \times 10^{14} M_\odot h^{-1}$, and $0.3 < z < 1.2$) was analyzed in this set of simulations as in the simulated tSZ-only maps. We will refer to this set of simulations as the full-noise S10 maps. These simulations are important for understanding the cluster sample selection, which will be affected by the total noise.

The scaling relation fitting for the clusters from this set of simulations was performed as in Section 6.3. As in Section 6.3, the scatter is a weak function of angular aperture, with the minimum shifted to $\sim 0.5-0.75$. Figure 5 shows the fractional scatter as a function of angular aperture of integration.

For the $Y_{SZ}^{0.75} - M_{vir}$ scaling relation we find a fractional scatter in $Y_{SZ}$ of $27\% \pm 1\%$. Since the scatter here includes both intrinsic scatter and the measurement uncertainty, we expect it to be larger than the scatter in $Y_{θ}^{0.75}$ in Section 6.3. This level of scatter is comparable to the $27\% \pm 2\%$ scatter in $\zeta$ found in the $\zeta - M_{vir}$ scaling relation for these simulations.

### Table 1

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Integration Radius</th>
<th>$A_{SZ}$ ($\times 10^{-4}$)</th>
<th>$B_{SZ}$</th>
<th>$C_{SZ}$</th>
<th>Scatter</th>
<th>MF Scatter</th>
</tr>
</thead>
<tbody>
<tr>
<td>tSZ-only S10 sims</td>
<td>0.75 Mpc</td>
<td>1.44 ± 0.11</td>
<td>1.20 ± 0.11</td>
<td>1.63 ± 0.24</td>
<td>23% ± 2%</td>
<td>27% ± 2%</td>
</tr>
<tr>
<td></td>
<td>0.3 Mpc</td>
<td>1.53 ± 0.16</td>
<td>1.26 ± 0.17</td>
<td>1.13 ± 0.13</td>
<td>28% ± 2%</td>
<td></td>
</tr>
<tr>
<td>Full-noise S10 sims</td>
<td>0.75 Mpc</td>
<td>1.37 ± 0.10</td>
<td>1.04 ± 0.11</td>
<td>1.02 ± 0.20</td>
<td>27% ± 1%</td>
<td>27% ± 2%</td>
</tr>
<tr>
<td></td>
<td>0.3 Mpc</td>
<td>1.49 ± 0.18</td>
<td>1.12 ± 0.22</td>
<td>0.53 ± 0.25</td>
<td>34% ± 2%</td>
<td></td>
</tr>
<tr>
<td>B13 SPT observed clusters</td>
<td>0.75 Mpc</td>
<td>1.85 ± 0.36</td>
<td>1.77 ± 0.35</td>
<td>0.96 ± 0.50</td>
<td>21% ± 11%n</td>
<td>21% ± 9%a</td>
</tr>
<tr>
<td></td>
<td>0.3 Mpc</td>
<td>2.09 ± 0.35</td>
<td>1.43 ± 0.20</td>
<td>0.35 ± 0.28</td>
<td>26% ± 9%a</td>
<td></td>
</tr>
</tbody>
</table>

**Notes.** The tSZ-only maps contain only thermal SZ signal, while the full-noise S10 maps include tSZ, CMB, point sources, atmospheric noise, and realistic SPT instrumental noise. The values of scatter reported for the simulations are fractional scatter, while the values reported for the B13 clusters are intrinsic log-normal scatter. In the S10 simulations virial masses are used to fit the scaling relations, while for the B13 cluster sample the masses are $M_{vir}$. For comparison with the scatter in each $Y_{SZ} - M$ scaling relation we list the scatter in the corresponding MF derived $\zeta - M$ scaling relation for the same data set.

*a These values are intrinsic log-normal scatter.

Figure 7 shows $Y_{SZ}^{0.75}$ versus $M_{vir}$ for 1187 mass-selected clusters in the full-noise S10 simulations, which include CMB, point sources, astrophysical noise, and realistic SPT instrument noise. We include only clusters with $M_{vir} > 4 \times 10^{14} M_\odot h^{-1}$ in the redshift range $0.3 < z < 1.2$. Fractional scatter in $Y_{SZ}$ is $27\% \pm 1\%$. The solid line is the best-fit $Y_{SZ}^{0.75} - M_{vir}$ scaling relation found for this cluster sample.

Using these simulations we also calculate $Y_{SZ}^{\rho}$ for a range of $\rho$ values, as in Section 6.3, and fit $Y_{SZ}^{\rho} - M_{vir}$ scaling relations for each $\rho$. Figure 5 shows the fractional scatter as a function of the integration radius for a range of physical radii. We find a broad minimum in scatter at $\sim 0.2-0.3$ Mpc, with a minimum scatter of $33\% \pm 2\%$. The optimal integration radius shifts down...
here relative to the simulated tSZ-only maps because of the scale dependence of the noise sources added in the full-noise S10 maps, which dominate the scatter in these simulations. In particular, the noise induced by the simulated CMB increases with angular scale, leading to a preference for smaller integration radii. The scatter in $Y_{SZ}$ for these simulations is slightly higher than the scatter in both the $\xi$ and the $Y_{SZ}^{v, 75}$ mass scaling relations. The scaling relation parameters for the nominal $Y_{SZ}^{v, 3}$ mass scaling relation are given in Table 1. The optimal physical radius again corresponds roughly to the optimal angular aperture, at the median redshift of the cluster sample.

7. $Y_{SZ}$ FOR SPT OBSERVED CLUSTERS

7.1. $Y_{SZ}$–$M_{500}$ Scaling Relation Fitting Methods

In this section, we perform $Y_{SZ}^{0}$–$M_{500}$ scaling relation fitting for a sample of SPT observed clusters, using the same scaling relation as in the simulations (Equation (10)) and the X-ray determined cluster masses. We use only 150 GHz data in this section to be directly comparable to the results in Benson et al. (2013), which does not include 95 GHz data. In this section, we define cluster mass as $M_{500}$, the mass inside a spherical radius $r_{500}$, within which the average density is 500 times the critical density of the universe at the cluster’s redshift. To fit this scaling relation with clusters selected in the SPT-SZ survey, we have to account for the shape of the cluster mass function and the SPT survey selection, which was based on the SPT significance, $\xi$. This is similar to the procedure followed in previous SPT analyses (Reichardt et al. 2013; B13; V10), with the added complication that in this work we must express the SPT selection function in $Y_{SZ}$ instead of $\xi$.

The (unnormalized) probability of a mass $M$ given an integrated Comptonization $Y_{SZ}$ is given by

$$\mathcal{P}(M|Y_{SZ}) = P(Y_{SZ}|M)\, P(M),$$

where $P(Y_{SZ}|M)$ is the Gaussian probability distribution with which we have been working previously, and $P(M)$ is the mass function. The number of clusters is a steep function of cluster mass, which (combined with the measurement uncertainty in $Y_{SZ}$) results in relatively more low-mass than high-mass clusters at a given $Y_{SZ}$, an effect commonly referred to as Eddington bias.

For our cluster sample we use the 18 clusters from B13, 14 of which have X-ray derived masses (see Section 2.3), and all of which have $\xi > 5$. For this analysis we use only the 150 GHz data, the SPT band with the highest SZ sensitivity. For a list of cluster names, $\xi$ values, and redshifts for this sample, see Table 2.

To fit for scaling relations, we use a method similar to the one described in B13, which we modify to account for the cluster selection based on $Y_{SZ}$ instead of $\xi$. In B13, we used a version of the CosmoMC (Lewis & Bridle 2002) analysis package, modified to include the cluster abundance likelihood $P(Y_{SZ})$. The (unnormalized) probability of a mass $M$ given an integrated Comptonization $Y_{SZ}$ is given by

$$\mathcal{P}(M|Y_{SZ}) = P(Y_{SZ}|M)\, P(M),$$

where $P(Y_{SZ}|M)$ is the Gaussian probability distribution with which we have been working previously, and $P(M)$ is the mass function. The number of clusters is a steep function of cluster mass, which (combined with the measurement uncertainty in $Y_{SZ}$) results in relatively more low-mass than high-mass clusters at a given $Y_{SZ}$, an effect commonly referred to as Eddington bias.

For our cluster sample we use the 18 clusters from B13, 14 of which have X-ray derived masses (see Section 2.3), and all of which have $\xi > 5$. For this analysis we use only the 150 GHz data, the SPT band with the highest SZ sensitivity. For a list of cluster names, $\xi$ values, and redshifts for this sample, see Table 2.

To fit for scaling relations, we use a method similar to the one described in B13, which we modify to account for the cluster selection based on $Y_{SZ}$ instead of $\xi$. In B13, we used a version of the CosmoMC (Lewis & Bridle 2002) analysis package, modified to include the cluster abundance likelihood in the CosmoMC likelihood calculation. All fitting is performed assuming a standard flat $\Lambda$CDM cosmology, parameterized with the standard six-parameters ($\Omega_m, h, \Omega_b h^2, \Omega_c, n_s, \Delta_k^2$, and $\tau$), and using the WMAP 7 yr data set. The cosmological parameters are held constant throughout, we marginalize over only the cluster scaling relation parameters. At each step in the chain, a point in the joint cosmological and scaling relation parameter space is selected. The Code for Anisotropies in the Microwave Background (Lewis et al. 2000) is used to compute the matter power spectrum at 20 redshift bins between 0 < $z$ < 2.5, spaced logarithmically in 1 + $z$. The matter power spectrum,
cosmological parameters, and $Y_{SZ-M_{500}}$ and $Y_{X-M_{500}}$ scaling relation parameters are then input to the cluster likelihood function. $Y_X$ is defined as $Y_X \equiv M_X T_X$, where $M_X$ is the cluster gas mass within $r_{500}$, and $T_X$ is the core-excised X-ray spectroscopic temperature in an annulus between $0.15$ and $1.0 \times r_{500}$.

To calculate the cluster likelihood, first the matter power spectrum and cosmological parameters are used to calculate the cluster mass function, based on Tinker et al. (2008). Next, the mass function is converted to the predicted cluster abundance in our observable space, $N(Y_{SZ}, Y_X, z)$. This conversion is accomplished using our standard $Y_{SZ-M_{500}}$ scaling relation (Equation (10)), with flat unbounded priors on all parameters, and the $Y_{X-M_{500}}$ scaling relation from B13:

$$
\frac{M_X^{500}}{10^{14} M_\odot h^{-1}} = (A_X h^{3/2}) \left( \frac{Y_X}{3 \times 10^{44} \text{keV}} \right)^{B_X} E(z)^{C_X},
$$

parameterized by the normalization factor $A_X$, the mass scaling $B_X$, the redshift evolution $C_X$, and the log-normal intrinsic scatter $D_X$. This scaling relation is based on the relation used in Vikhlinin et al. (2009a).

The priors on the $Y_{X-M_{500}}$ scaling relation parameters were $A_X = 5.77 \pm 0.56$, $B_X = 0.57 \pm 0.03$, $C_X = -0.40 \pm 0.20$, and $D_X = 0.12 \pm 0.08$. All these priors are Gaussian. The slope and normalization priors were motivated by Vikhlinin et al. (2009a), and the priors on the redshift evolution and scatter were motivated by the range observed in several different sets of simulations that included varying astrophysics (Krautsgrov et al. 2006; Fabjan et al. 2011). These priors are identical to those used in Benson et al. (2013), and a more detailed description motivating them is given there.

The predicted cluster density as a function of $Y_{SZ}, Y_X$, and $z$ can be written as follows:

$$
\frac{dN(Y_{SZ}, Y_X, z)}{dY_{SZ} \, dY_X \, dz} = \int P(Y_{SZ}, Y_X \mid M, z, p) \, P(M, z \mid p) \, \Phi(Y_{SZ}) \, dM,
$$

where $p$ is the set of cosmological and scaling relation parameters, and $\Phi(Y_{SZ})$ is the selection function in $Y_{SZ}$. This predicted cluster density function differs from B13 in that the selection function must be transformed from a Heaviside step function at $\xi = 5$ into a function of $Y_{SZ}$. We assume that $Y_{SZ}$ and $\xi$ can be related with a log-normal distributed scaling relation, and that the selection in B13 can therefore be well approximated by an error-function in $Y_{SZ}$. We then define our selection function as

$$
\Phi(Y_{SZ}) = \frac{1}{2} \text{erf} \left( \frac{Y_{SZ} - Y_{SZ}^\phi(z)}{\sqrt{2} Y_{SZ}^\phi(z) D} \right) + \frac{1}{2},
$$

where the selection threshold, $Y_{SZ}^\phi(z)$, is defined as the $Y_{SZ}$ value corresponding to $\xi = 5$ at the redshift $z$. We estimate $Y_{SZ}^\phi(z)$ by fitting a $Y_{SZ-\xi}$ scaling relation of the form

$$
Y_{SZ} = A \xi^B E(z)^C,
$$

using the catalog of SPT observed clusters given in R13. The width of the selection error-function is given by the scatter in the $Y_{SZ-\xi}$ scaling relation, $D$.

We evaluate Equation (13) on a $200 \times 200 \times 30$ grid in $(Y_{SZ}, Y_X, z)$ space, and convolve with a Gaussian error term in $Y_{SZ}$ to account for the measurement noise. The width of the Gaussian is given by the uncertainty in $Y_{SZ}$ as a function of $Y_{SZ}, \delta Y_{SZ}(Y_{SZ})$, as determined by the cluster parameterization MCSC (see Section 5.3).

The likelihood function of the observed cluster sample is defined by the Poisson probability:

$$
\log (\mathcal{L}(p)) = \sum_i \log \left( \frac{dN(Y_{SZ}, Y_X, z_i \mid p)}{dY_{SZ} \, dY_X \, dz} \right) - \int \frac{dN(Y_{SZ}, Y_X, z \mid p)}{dY_{SZ} \, dY_X \, dz} \, dY_{SZ} \, dY_X \, dz, \quad (16)
$$

where the summation is over the SPT clusters in our catalog. Note also that this is the unnormalized log-likelihood.

There is a complication, in that $Y_X$ is dependent on the cosmological parameters. $Y_X \equiv M_X T_X$, where $M_X$ is the gas mass within $r_{500}$, and $T_X$ is the core-excised X-ray temperature in an annulus between $0.15 \times r_{500}$ and $1.0 \times r_{500}$. To maintain consistency with the cosmological parameters, we recalculate $Y_X$ for each cluster at every step in CosmoMC, given the current $Y_{SZ-M_{500}}$ relation and $r_{500}$. In the likelihood, we add $\sum_i \log (Y_{X_i})$ to the right hand side of Equation (16) to account for the recalculation of $Y_X$. For a detailed explanation of this correction term, see Appendix B of B13.

To account for measurement error in $Y_X$ and $z$ for each cluster, we marginalize over the relevant parameter, weighted by a Gaussian likelihood determined by its uncertainty. For the few clusters without observed $Y_X$ data, we instead weight the marginalized parameter by a uniform distribution over the allowed parameter range.

The likelihood of this set of cosmological and scaling relation parameters is then used by CosmoMC in the acceptance/rejection computation. Only the $Y_{SZ-M_{500}}$ scaling relation parameters are of interest to us in this analysis. The cosmological and $Y_{X-M_{500}}$ scaling relation parameters were used as a cross-check to verify that the results were in agreement with the analysis performed on these clusters in B13, but will not be presented here. All parameters differed from the values presented in B13 by $\ll 1\sigma$.

### 7.2. $Y_{SZ-M_{500}}$ Scaling Relation Results

We use CosmoMC to fit $Y_{SZ-M_{500}}$ scaling relations for a range of angular apertures, and find a broad minimum in scatter in the range $0.5-0.75$, with a minimum intrinsic log-normal scatter of $21\% \pm 11\%$. For these scaling relation parameters we apply flat, unbounded priors. The scatter in the $\xi-M_{500}$ scaling relation for these clusters is comparable, at $21\% \pm 9\%$. The scaling relation parameters for the $Y_{SZ-M_{500}}$ scaling relation are given in Table 1.

We also fit mass scaling relations for $Y_{SZ}$ integrated within a range of physical radii, $r$, from $0.1$ Mpc to $0.5$ Mpc. We find a broad minimum in scatter in the range $0.2-0.3$ Mpc, with a minimum intrinsic log-normal scatter of $23\% \pm 5\%$. This is comparable to the scatter in both the $\xi$ and $Y_{SZ}$ mass scaling relations. The parameters for the nominal $Y_{SZ}^{0.75}$ mass scaling relation (0.3 Mpc corresponds to $0.75$ at the survey median redshift of $z = 0.6$) are listed in Table 1.

### 7.3. Cluster Masses

To calculate the masses of the clusters, the $Y_{SZ}^{0.75}$ CosmoMC chains were used. The probability density function for the mass
was computed on an evenly spaced mass grid for each step in the CosmoMC chains. These probability density functions were then summed to obtain a mass estimate fully marginalized over all scaling relation and cosmological parameters. This was done for CosmoMC chains containing only all scaling relation and cosmological parameters. This was done then summed to obtain a mass estimate fully marginalized over the CosmoMC chains. These probability density functions were computed on an evenly spaced mass grid for each step in Figure 8.

Figure 8 shows the cluster masses calculated from the $Y_{SZ} - M_{500}$ scaling relation versus the masses calculated from the $Y_X - M_{500}$ scaling relation for the B13 cluster sample. The solid line is the reference line $M_{500}^{SZ} = M_{500}^X$.

7.4. $Y_{SZ}(r_{500})$

The self-similar model of cluster formation assumes that clusters scale in well-defined ways based on their mass, typically defined within physical radii proportional to the critical density of the universe at the cluster’s redshift (e.g., Kravtsov & Borgani 2012; Kaiser 1984). For this reason, studies of the scaling relations of clusters typically compare physical observables defined by this physical radius, usually $r_{500}$. In this section we will calculate $Y_{SZ}(r_{500})$, denoted $Y^{SZ}_{500}$, for comparison with other published parameters for the clusters in B13.

We investigated a method for estimating $r_{500}$ from SZ data, as a way to measure $Y^{SZ}_{500}$ solely from SZ data. This method proved to be problematic, however, because it required estimating $M_{500}$ from a fixed angular aperture, and calculating $r_{500}$ from that estimate. This results in the scatter in the $Y^{0.75}_{SZ} - M_{500}$ scaling relation feeding back into the calculation of $Y^{SZ}_{500}$. Instead, we use the X-ray determined $r_{500}$ in our calculations of $Y^{SZ}_{500}$.

In Table 2, we give the measured $Y^{SZ}_{500}$ values for our cluster sample. We note that, as defined in Equation (7), the MCMC fits for a cylindrically projected measure of $Y^{SZ}_{500}$ rather than the spherical de-projected value often used in other $Y_{SZ} - M$ scaling relation results (e.g., A11; Arnaud et al. 2010). $Y_{SZ}^{500}$ values are given in $M_{\odot}$ keV here, for comparison with $A_{11}$.

$A_{11}$ describe a template fitting method of estimating $Y_{SZ}^{500}$, which uses an SZ source template motivated from X-ray measurements of each cluster. The profile is assumed to match the product of the best-fit gas density profile to the X-ray measurements of each cluster, and the universal temperature profile of Vikhlinin et al. (2006). These profiles are multiplied together to produce the radial pressure profile, and projected onto the sky using a line-of-sight integral through the cluster. $A_{11}$ then construct a spatial filter using Equation (1), and this X-ray derived source model. The X-ray determined cluster position is used to place priors on the cluster location to prevent maximization bias in the recovered $Y_{SZ}^{500}$ values. $Y_{SZ}^{500}$ is calculated by integrating the source model over a solid angle corresponding to $r_{500}$, as in Equation (7).

In Figure 9, we plot the $Y_{SZ}^{500}$ estimated by the MCMC method against the $Y_{SZ}^{500}$ estimated by the template fitting method in $A_{11}$. The best fit relation between the two is $Y_{SZ}^{500}$(MCMC) = (0.98 ± 0.09) $Y_{SZ}^{500}$(A11), where the uncertainty is the range for which $\Delta \chi^2 < 1$ (68% confidence limit) compared to the best-fit. (For a treatment of the calculation of $\chi^2$ with uncertainty in both variables, see, for example, Numerical Recipes in C++, Section 15.3, Press et al. 2002.) We see that these two methods of calculating $Y_{SZ}^{500}$ are consistent, that is, the best-fit scaling relation is consistent with equality between $Y_{SZ}^{500}$(MCMC) and $Y_{SZ}^{500}$(A11). The scatter about the expected one-to-one line here is dominated by differences in cluster model shape between the two methods (X-ray derived SZ profile versus $\beta$-model).

We also verify that our $Y_{SZ}^{500}$ values for these clusters are in agreement with the $Y_X$ values presented in B13, given the expected $Y_{SZ} - Y_X$ scaling. Figure 10 shows the $Y_{SZ}^{500}$ values of
our catalog of SPT observed clusters plotted against their $Y_X$ values from B13.

We can make a prediction of the relationship between $Y_{SZ}$ and $Y_X$ based on the universal pressure profile from Arnaud et al. (2010), based on X-ray measurements of a representative sample of local, massive clusters. Even though $Y_{SZ}$ and $Y_X$ are effectively measures of the cluster pressure, they depend on the details of the shape of the profile differently, which can still vary somewhat between clusters. Assuming the Arnaud et al. (2010) pressure profile, we predict a relationship of $Y_{SZ}^{500} = 1.08 Y_X$ (solid), the best-fit line (green dashed), and the uncertainty in the fit defined as the range for which $\Delta \chi^2 < 1$ compared to the best fit (red dot-dashed). The best-fit normalization is $A = 1.17 \pm 0.12$, consistent with the expected scaling between $Y_{SZ}^{500}$ and $Y_X$.

8. CONCLUSIONS

We describe and implement a method of constraining $Y_{SZ}$ generalizable to any cluster profile, and we show that this method accurately recovers $Y_{SZ}$ in simulations. We compare $Y_{SZ}$ to SPT cluster detection significance, focusing on scatter with mass. Finally, we apply this method to clusters detected in the SPT-SZ survey, and compare the estimated $Y_{SZ}$ values to $Y_{SZ}$ estimated by a template fitting method, and to $Y_X$.

We apply our method to clusters in simulated tSZ-only maps and measure $Y_{SZ}^{\rho}$, the integrated Comptonization within a constant angular aperture. We find that $Y_{SZ}$ is measured with the lowest fractional scatter in an aperture comparable to the SPT beam size ($\sim 1'$ FWHM at 150 GHz). We fit $Y_{SZ}^{\rho} - M_{\rho}$ scaling relations for a range of angular apertures and find a minimum fractional scatter of $23\% \pm 2\%$ in $Y_{SZ}$, at a fixed mass, with the minimum occurring for an angular aperture of 0.75'. We also calculate $Y_{SZ}$ within a range of physical radii, $\rho$, and find a minimum scatter in $Y_{SZ}$ at an integration radius of 0.3 Mpc, which corresponds roughly to 0.75 at the survey median redshift ($z = 0.6$), with a fractional scatter of $28\% \pm 2\%$ at a fixed mass. Using the same simulated clusters, we also fit a $\xi-M_{\rho}$ relation, where $\xi$ is the MF SZ detection significance measured by SPT, and find a fractional scatter of $27\% \pm 2\%$.

We also analyze clusters in simulations including tSZ, CMB, point sources, atmospheric noise, and realistic SPT instrumental noise. In these full-noise simulations, the $Y_{SZ}^{0.75} - M_{\rho}$ scaling relation has $27\% \pm 1\%$ scatter, the $Y_{SZ}^{0.3 \text{ Mpc}} - M_{\rho}$ scaling relation has $34\% \pm 2\%$ scatter, and the $\xi-M_{\rho}$ scaling relation has $27\% \pm 2\%$ scatter. These simulations demonstrate that scatter in $Y_{SZ}^{0.75}$ is comparable to the scatter in $\xi$.

To investigate $Y_{SZ}$ scaling relations in SPT observed clusters, we fit $Y_{SZ}^{\rho} - M_{\rho}$ and $Y_{SZ}^{\rho} - M_{500}$ scaling relations to the sample of 18 SPT clusters described and examined in Benson et al. (2013). Of these, 14 have X-ray observations and measured $Y_X$ values, which we use to estimate the cluster $M_{500}$ masses. We fit the scaling relations using a version of CosmoMC, similar to the one described in Benson et al. (2013), modified to account for the cluster selection based on $Y_{SZ}$ instead of SPT significance.

For these clusters, the $Y_{SZ}^{0.75} - M_{500}$ scaling relation is found to have $21\% \pm 11\%$ intrinsic log-normal scatter in $Y_{SZ}$, at a fixed mass, the $Y_{SZ}^{0.3 \text{ Mpc}} - M_{500}$ scaling relation has $26\% \pm 9\%$ scatter, and the $\xi-M_{500}$ relation has $21\% \pm 9\%$ scatter.

We also calculate a cylindrically projected $Y_{SZ}^{\rho}$, the integrated Comptonization within $r_{500}$, for the clusters in the Benson et al. (2013) sample. We compare the $Y_{SZ}^{500}$ values recovered by our Markov Chain Monte Carlo method to those calculated for the same clusters by the template fitting method described in A11 and find the two methods to be consistent. We further compare the MCMC derived $Y_{SZ}^{500}$ values to the $Y_X$ values for these clusters from Benson et al. (2013) and find that they are consistent with the expected scaling between $Y_{SZ}$ and $Y_X$, based on the universal pressure profile of Arnaud et al. (2010).

We have demonstrated, with both simulations with realistic SPT noise and SPT observed clusters, that $Y_{SZ}$ is most accurately determined in an aperture comparable to the SPT beam size. We have used this information in measuring $Y_{SZ}$ for the catalog of clusters observed with the SPT in the 2008 and 2009 seasons (Reichardt et al. 2013). The SPT-SZ survey of 2500 deg$^2$ was completed in 2011 November, and has detected ~500 clusters with a median redshift of ~0.5 and a median mass of $M_{500} \sim 2.3 \times 10^{14} M_\odot$ h$^{-1}$. The methods and results presented here will inform the measurement and use of $Y_{SZ}$ for the clusters detected in the full SPT-SZ survey.