

Mass calibration of the Sunyaev-Zel'dovich effect using APEX-SZ galaxy clusters

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Aarti Nagarajan
aus
Chennai, India

Bonn, 05.02.2018

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

1. Gutachter: Prof. Dr. Frank Bertoldi
2. Gutachter: Prof. Dr. Thomas H. Reiprich

Tag der Promotion: 28.06.2018
Erscheinungsjahr: 2018

*For amma, appa & Anu.
In loving memory of my paati, Janaki.*

Abstract

Clusters of galaxies are the most massive, gravitationally bound structures in the Universe. Their abundance across cosmic time and space is sensitive to cosmological parameters. Of particular interest to modern cosmology is the nature of the mysterious “dark energy” and the total mass of all neutrino species. To place tight constraints on these from cluster observations, it is crucial to accurately relate the cluster observables to the cluster masses, which are the link to theoretical modelling in cosmology. This work aims at measuring accurately the relationships between cluster properties and mass using a well-defined cluster sample.

A conspicuous observable component of a cluster is the diffuse hot ($\gtrsim 10^7$ K) ionized medium filling up the space between the galaxies. This intra-cluster medium emits X-rays and causes a well-understood distortion in the cosmic microwave background radiation spectrum known as the Sunyaev-Zel'dovich (SZ) effect. In this work, I calibrate the cluster SZ effect against the cluster total gravitational mass, making use of spatially resolved SZ observations of 39 galaxy clusters, obtained with the APEX telescope. These observations are riddled with atmospheric noise, which requires strong filtering to retrieve the cluster SZ signal that thereby is also attenuated and spatially distorted.

In the first part of the thesis, I discuss various ways of estimating the true SZ signal from the filtered APEX observations, in order to obtain robust estimates of the SZ signal for all the clusters in the sample.

In the second part, I focus on the SZ effect to mass scaling relation. I compare the SZ measurements with direct mass estimates obtained from gravitational lensing, a process by which the images of background galaxies are distorted by the gravitational potential of a cluster that deflects the light from its original path. For a reliable cosmological use of clusters, their mass-observable scalings must be obtained from a sample that is representative of the actual cluster population or whose composition and completeness (as a fraction of the underlying cluster population) are at least well-understood, such that resulting biases can be corrected for. For this reason, I restrict the analysis to a sub-sample of the APEX clusters, selected from an X-ray all-sky survey, and qualified as complete (>90%) inside a well-defined range of X-ray luminosities. I implement a novel Bayesian method to account for the sample selection biases, the measurement uncertainties, the shape of the cluster mass function and scatters in the true mass-observable relations.

In particular, I allow for a possible covariance between the X-ray luminosity and the SZ effect at fixed mass, which so far has been neglected in similar works. The resulting calibration of the SZ to mass scaling relation shows sensitivity to this covariance. Neglecting the covariance causes the normalization of the SZ to mass relation to be biased high by one to two standard deviations, and the slope low by one standard deviation, even when the SZ effect plays no role in the sample selection. Based on different mass-observable scaling relations, I estimate the impact of such a systematic effect on forthcoming cosmological analyses by predicting the number of cluster detections in a near-future SZ survey experiment. The prediction from the trusted calibration yields 5000, whereas the scaling relation that ignores the X-ray luminosity and SZ covariance at fixed mass predicts the cluster count too high by a factor of four. This highlights the need for considering the intrinsic covariance of cluster properties in measuring scaling relations. For the higher-precision data and larger cluster samples anticipated from on-going and near-future cluster cosmology experiments, biases in the mass-observable calibrations from covariances of cluster properties could dominate the cosmological error budget if not considered with care.

List of Publications

Chapters 3, 5, 6 and Section 2.5 of Chapter 2 of this thesis are part of the paper submitted to journal Monthly Notices of the Royal Academic Society (MNRAS):

“Weak-lensing mass calibration of the Sunyaev-Zel’dovich effect using APEX-SZ galaxy clusters”

A. Nagarajan, F. Pacaud, M. Sommer, M. Klein, K. Basu, F. Bertoldi, A. T. Lee, P. A. R. Ade, A. N. Bender, D. Ferrusca, N. W. Halverson, C. Horellou, B. R. Johnson, J. Kennedy, R. Kneissl, K. M. Menten, C.L. Reichardt, C. Tucker, B. Westbrook, Submitted to MNRAS.

Additional publications that were not incorporated into this thesis:

“Galaxy cluster scaling relations measured with APEX-SZ”,

A. N. Bender, J. Kennedy, P. A. R. Ade, K. Basu, F. Bertoldi, S. Burkutean, J. Clarke, D. Dahlin, M. Dobbs, D. Ferrusca, D. Flanigan, N. W. Halverson, W. L. Holzapfel, C. Horellou, B. R. Johnson, Z. D. Kermish, M. Klein, R. Kneissl, T. Lanting, A. T. Lee, J. Mehl, K. M. Menten, D. Muders, **A. Nagarajan**, F. Pacaud, C. L. Reichardt, P. L. Richards, R. Schaaf, D. Schwan, M. W. Sommer, H. Spieler, C. Tucker, and B. Westbrook (2016). MNRAS, DOI: 10.1093/mnras/stw1158

Contents

Abstract	v
List of Publications	vii
1 Motivation and outline	1
2 Clusters of galaxies	5
2.1 Formation of galaxy clusters in Λ CDM cosmologies	7
2.1.1 Cosmological distances and volume: implications for counting clusters	8
2.1.2 Evolution of structures through cosmic time	10
2.2 Galaxy cluster observables in multi-wavelength	16
2.2.1 X-rays	16
2.2.2 Microwaves (Sunyaev-Zel'dovich effect)	18
2.2.3 Optical and near-infrared	21
2.3 Linking cluster observables to mass: scaling relations	25
2.4 Current cosmological constraints from galaxy clusters and prospects for the future . .	28
2.5 Scaling relations and sample selection biases	30
2.6 Summary	32
3 Cluster sample: APEX-SZ experiment and multi-wavelength follow-up	33
3.1 APEX-SZ program	33
3.2 APEX-SZ cluster sample	34
3.2.1 The eDXL cluster sample	34
3.2.2 Other APEX-SZ clusters	36
3.2.3 X-ray observables and parameter estimation	37
3.3 Optical follow-up observations and lensing masses	38
3.4 APEX-SZ data analysis performed with bolometer analyser (BoA)	43
3.4.1 Calibration	43
3.4.2 Time stream processing	43
3.4.3 Point source transfer function	44
3.5 Post-processed data products	46
3.5.1 Co-added reduced source maps	46
3.5.2 Jack-knived noise maps	48
3.5.3 Point source transfer function (PST)	48
3.6 Summary	48
4 Methods: measuring integrated Comptonization from APEX-SZ	51
4.1 Motivation	51

4.2	Method	52
4.2.1	Parametric models	52
4.2.2	Generating mock APEX-SZ Compton-y images	53
4.2.3	Bayesian fitting approach	56
4.2.4	Fitting methods	56
4.3	Results	62
4.3.1	Testing the Fourier Domain (FD) method	62
4.3.2	Testing the Radial Binning (RB) method	68
4.3.3	Comparison of the two methods	72
4.4	Conclusion & discussion	79
5	Results: measuring integrated Comptonization from APEX-SZ	83
5.1	Generalised Navarro-Frenk-White profile fit to APEX-SZ	83
5.1.1	Propagation of uncertainties in R_{500} into the SZ modelling	85
5.1.2	Centers for the gNFW model	88
5.2	Comparison with Bender et al.	89
5.3	Summary and discussion	91
6	Weak-lensing mass calibration of the Sunyaev-Zel'dovich effect using APEX-SZ galaxy clusters	93
6.1	Introduction	94
6.2	Mass proxy measurements	96
6.3	Method	96
6.3.1	Statistical model	97
6.3.2	Application to the eDXL sample	99
6.3.3	Tests with simulations	103
6.4	Results	105
6.4.1	Including correlated intrinsic scatters in Y_{SZ} and L_x at fixed mass	105
6.4.2	Including uncorrelated intrinsic scatter in the weak-lensing masses	110
6.4.3	Correlated intrinsic scatter: interpretation from residuals	111
6.5	Robustness and limitations of the analysis	113
6.5.1	Redshift evolution of scaling relations	113
6.5.2	Treatment of completeness of the eDXL sample	115
6.5.3	Consistency check of measured masses and model prediction	115
6.5.4	Additional covariances in the scatters of mass observables	115
6.5.5	Impact of mis-centring gNFW profile	115
6.5.6	Systematic uncertainty in weak-lensing mass estimates	116
6.5.7	Treatment of outliers	116
6.6	Discussion	118
6.6.1	Comparison to literature	118
6.6.2	Origin of correlation in intrinsic scatters of L_x and Y_{SZ} at fixed mass	120
6.6.3	Impact on cluster based cosmological studies	122
6.7	Summary & Conclusions	123
7	Summary & discussion	125
7.1	Summary and significance	125
7.2	Scope of this work	126

7.3	Limitations of this work	127
Bibliography		129
A Appendix: Mass calibration and statistical models		135
A.1	Derivation of the normalised likelihood for the eDXL sample	135
A.2	Probabilities	136
A.2.1	Measurement probabilities	136
A.2.2	Scaling model probabilities	137
A.2.3	Intrinsic scatter in weak-lensing mass	137
A.3	Tests with simulated data	138
A.3.1	Mock data with correlated intrinsic scatters	138
A.3.2	Mock data with intrinsic scatter in the lensing mass	142
List of Figures		145
List of Tables		155
Declaration		159
Acknowledgements		161

CHAPTER 1

Motivation and outline

In late 1990s, using Supernovae type I as standard distance candles, it was discovered that the Universe is undergoing accelerated expansion (Riess et al., 1998; Perlmutter et al., 1999). In a Friedmann Universe, this late-time acceleration can only be caused by an unknown constituent of negative pressure, which is usually termed *dark energy*. Current constraints on the energy density contents of the Universe inform that the dark energy makes up nearly 70% of the total energy budget. The nature of this energy can be probed via its equation of state that relates its energy density to its pressure. In this respect, growth of structures are an incredible tool for probing the nature of the dark energy as the structure formation and its growth are sensitive to the dark energy equation of state. The largest structures formed in the Universe are galaxy clusters and super-clusters. Numerous recent cosmological analyses have demonstrated the use of galaxy clusters in constraining the content and geometry of the Universe (e.g., Vikhlinin, Burenin et al., 2009; Rozo et al., 2010; A. B. Mantz et al., 2015; de Haan et al., 2016; Schellenberger and Reiprich, 2017b) and they confirm the existence of a late-time accelerated expansion of the Universe.

Near-future galaxy cluster survey experiments (Benson, Ade et al., 2014; Merloni et al., 2012; Abazajian et al., 2016) will probe the growth and evolution of massive structures through their abundance in mass and redshift much more precisely as they discover many more galaxy clusters. Specifically, these experiments will map the abundance in a proxy observable such as X-ray luminosity or the Sunyaev-Zel'dovich (SZ) effect (introduced in Chapter 2: Sunyaev and Zel'dovich, 1970). It is vital to link these cluster observables to mass for the cosmological analysis. One of the ways of achieving the link is through the calibration of scaling laws of cluster observables and mass over a statistically significant sample of galaxy clusters. Galaxy clusters are immensely diverse in their dynamical state and their formation and evolution are affected by non-gravitational physics. Additionally, due to physical limitations set by the sensitivity and optics of the survey probes, the cluster population sampled in any survey tend to be biased one way or another. For these reasons, calibration of scaling laws from observations of a sample of galaxy clusters is a difficult task and requires simultaneous treatment of the systematic effects in measurements of cluster observables and mass proxies, the underlying mass function, the influence of physical processes that deviate from expectations of a gravity only dynamics for these cluster observables and the survey selection biases of the sample.

The work presented in this thesis is part of the APEX-SZ project and collaboration. The APEX-SZ experiment (Dobbs et al., 2006; Schwan et al., 2011) was a bolometer camera that observed at 150

GHz and mapped the SZ effect (the decrement in the *cosmic microwave background*, a relic radiation from the early Universe) in the direction of 45 galaxy clusters. 39 of these clusters were followed up in optical and X-rays by the collaboration. The goal of this work is to study the relationship between cluster mass and the integrated Comptonization of the SZ effect. The integrated Comptonization produced by galaxy clusters is a proxy for the thermal energy of the gaseous component residing in these massive structures. Measuring the mass scaling of the integrated Comptonization provides an important link between the thermal gas medium and cluster mass, which is important for cosmology. To calibrate this observable to mass, weak-lensing mass estimates from the dedicated optical follow-up by Klein et al. (in preparation) are used. The challenges in dealing with the sample selection biases due to the observable limited selection and the correlations in the intrinsic scatters of cluster observables are addressed in this work. For this purpose, an X-ray selected sub-sample of 27 clusters with a well-defined selection function is used for the measurement of the scaling behaviour. A Bayesian approach is employed to mitigate these biases in measuring the scaling relations. The biases are modelled by simultaneous treatment of the underlying mass function, the observable limited sample selection, intrinsic covariance of cluster observables at fixed mass, scaling relations and measurement uncertainties. Only a few authors have attempted to measure scaling relations by accounting for all these factors together (e.g., A. Mantz, Allen, Ebeling et al., 2010; A. Mantz, Allen, Morris et al., 2016).

Outline

A brief outline of this thesis is as follows:

- Chapter 2 gives an introduction to clusters of galaxies and presents an overview of their use as a cosmological probe. It introduces the key challenges and current limitations for cluster cosmological analyses.
- Chapter 3 introduces the APEX-SZ project and the galaxy cluster sample. A focus is placed on the X-ray selected sample with well-defined selection, which is of prime importance for measuring the mass calibration of the SZ effect. The relevant data and data reduction methods are summarised. These data form the basis for the work done in rest of the thesis.
- Chapter 4 focuses on methods for estimating the integrated Comptonization from filtered APEX-SZ maps in an unbiased manner. It presents different approaches to fitting parametric models to the post-processed data. It gives a quantitative and qualitative assessment of the performance of these techniques in the context of the APEX-SZ observations. A determination is made on the choice of the method for measuring the cluster observable from the SZ observations based on the analyses presented in this chapter.
- Chapter 5 provides the estimates on integrated Comptonization from APEX-SZ data. The modelling assumptions are then varied and discussed to assess the robustness of our measurements.
- Chapter 6 presents our final measurement of the scaling relation between integrated Comptonization and mass by anchoring the masses to weak-lensing measurements. This chapter provides a complete description of the Bayesian methods used to address the sample selection biases in scaling relations. A detailed discussion on the scaling relation and implications of the results are presented.

-
- Chapter 7 is a conclusion chapter that discusses the significance and limitations of this work.

The work presented in this thesis was developed in collaboration with members of the APEX-SZ project. Chapter 3 is presented mainly to summarise and introduce the data and their analyses from the collaborative work. These data form the basis for the main work carried out in this thesis. Chapters 4, 5, 6 present my original contributions to the work. In these chapters, I include a disclaimer note at the beginning of the sections that discuss work done primarily by the collaborators.

CHAPTER 2

Clusters of galaxies

Overview

Galaxy clusters are massive virialised structures and occupy the top of the hierarchical structure formation model. They are visible in wide range of the Electromagnetic (EM) spectrum. They contain matter in the form of stars and galaxies, cold molecular gas, hot thermal gas, non-thermal plasma, and dark matter. Only about 20% of its constituents are traceable in the EM spectrum. The baryon component is dominantly made up of the thermal gas, whereas the stellar contribution to the baryon component is relatively small $\sim 12\%$ (Lin and Mohr, 2004; Mulroy et al., 2014; Andreon, 2015).

All these components are maintained in a bound quasi-equilibrium state by the gravitational potential of the cluster, which is dominated by dark matter. Typical mass range for these virialised massive structures are in the range of 10^{13} – $10^{15} M_{\odot}$. Low mass galaxy clusters (with masses between 10^{13} and $10^{14} M_{\odot}$) are termed as galaxy groups. The distinction between clusters and groups is somewhat arbitrary. Galaxy groups tend to be more sensitive to non-gravitational effects and therefore, may deviate from the scaling relations established for massive

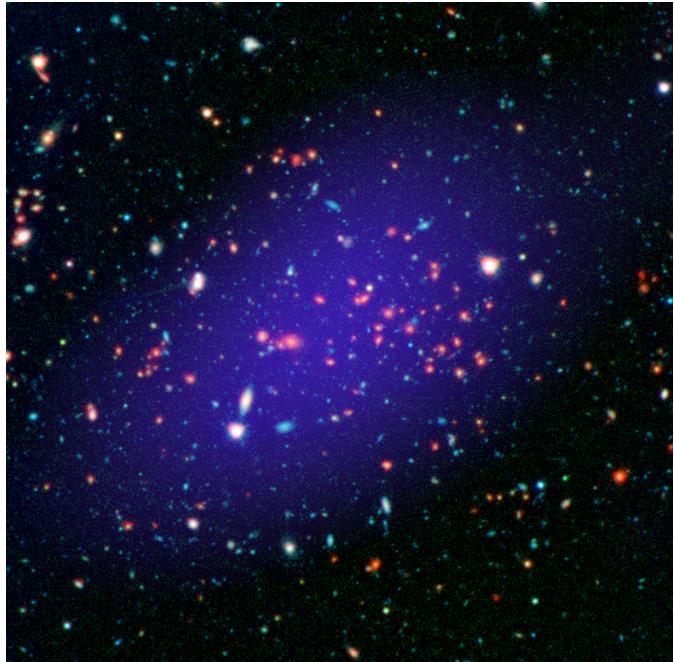


Figure 2.1: MOO J1142+1527, a galaxy cluster at 8.5 billion light years away discovered using the *Wide-field Infrared Survey Explorer (WISE)*. The image shows infrared observations from the *Spitzer Space Telescope* combined with near-infrared and visible light seen by the *Gemini Observatory*. The cluster of red galaxies in the center of the image forms the core part of the cluster. The blue halo shown is the Sunyaev-Zel'dovich effect tracing the thermal gas medium within the galaxy cluster observed with the *Combined Array for Research in Millimeter-wave Astronomy (CARMA)*. Credits: NASA/JPL-Caltech/Gemini/CARMA, originally adapted from Gonzalez et al. (2015).



Figure 2.2: Multi-wavelength composite image of the Abell 520 galaxy cluster. The multi-components are overlaid on the optical image using false colors. X-ray emission from gas medium (Green); Dark matter (Blue); Starlight (Orange). Credits: <http://chandra.harvard.edu/photo/2012/a520/>. Originally adapted from Jee et al. (2012).

galaxy clusters.

One of the most massive galaxy cluster discovered in infrared wavelength at $z > 1.1$ (when the Universe was approximately half its current age) and confirmed by the presence of thermal gas associated with it is shown in Figure 2.1 (Gonzalez et al., 2015). The appearance of galaxy clusters in different wavelengths is characterized by the matter constituents emitting radiation at these wavelengths. In optical and infra-red wavelengths, galaxy clusters appear as a concentration of galaxies and probe the light emitted by stars. In X-rays and mm wavelengths, they trace the diffuse gas medium residing in the cluster. This complementarity in probing of matter constituents in wide range of the EM spectrum allows astronomers to determine the distribution of these matter constituents in a galaxy cluster.

Although, it is not uncommon to find roundish and virialised galaxy clusters, a fraction of them can be found disturbed from their equilibrium state by violent mergers. One such galaxy cluster known as Abell 520 or “Train wreck cluster” is shown in Figure 2.2. This is an extremely disturbed cluster with its galaxies thrown away from where the majority of dark matter and the intra-cluster gas medium (ICM) reside. This illustrates that even though galaxy clusters are large enough to be a representation of matter contents in the Universe, they can have incredibly complex structures due to mergers.

In sections below, a review of the theoretical set up for the structure formation in Λ cold dark matter (Λ CDM) cosmological models is given. This is aimed at explaining how such massive structures could come to exist in the Universe and how they tie into the geometry and energy density budget of

the Universe.

2.1 Formation of galaxy clusters in Λ CDM cosmologies

The basis of Λ CDM model is that the general relativity of gravity, characterized by the Einstein's field equations, describe the gravitational force acting at cosmological scales. An important assumption is that the dark matter component is 'cold', which makes its momentum negligible, and that a cosmological constant Λ is part of the homogeneous and isotropic universe solution to the Einstein's field equations. According to this Λ CDM concordance model, structure formation driven prominently by gravitational forces is a hierarchical bottom-up process. By this, smaller structures form first and gradually build to accumulate more mass as the Universe grew over time. Such a formation was seeded by the overdensities in the primordial matter density fluctuations that were tiny in the early Universe. Evidence of such primordial fluctuations was first discovered imprinted on the *cosmic microwave radiation* (CMB) relic from early times of the Universe when radiation and matter decoupled from each other (Smoot et al., 1992). Many such overdensities have collapsed into massive structures we see today. Census of galaxy clusters in a volume, mass and redshift provide a tool for probing the geometry and content of the Universe. Here the structure formation and relevance of galaxy clusters to cosmology is presented.

First, some basic premises of a Λ CDM model are presented. The isotropic and homogeneous solution to the Einstein Field equations give the following two relations known as the Friedmann equations:

$$\frac{\ddot{a}}{a} = \frac{-4\pi G}{3} \left(\rho + 3\frac{p}{c^2} \right) + \frac{\Lambda c^2}{3} \quad (2.1)$$

and

$$\left(\frac{\dot{a}}{a} \right)^2 + \frac{Kc^2}{a^2} = \frac{1}{3}(8\pi G\rho + \Lambda c^2), \quad (2.2)$$

where a is the scale factor which determines the physical length scale of such a universe, ρ is the energy density of matter and radiation, p is the pressure of the radiation and matter, K is the spatial curvature and Λ is the cosmological constant. It becomes clear that the geometry of a universe is decided by the energy densities.

If the matter in the Universe behave as a perfect fluid, with an average equation of state parameter $w = p/\rho c^2$, it is clear that solving the previous system of equations requires first to determine the value of w .

The energy density ρ is usually decomposed into a number of components, depending on their equation of state parameter: pressureless matter ($w = 0$) often referred to as *dust* which combines dark and baryonic matter, and radiation ($w = 1/3$) which consists of massless and massive relativistic particles. The cosmological constant is often interpreted as a vacuum energy component ρ_Λ . By making some simple transformations, $\rho \rightarrow \rho + \rho_\Lambda = \rho + \frac{\Lambda c^2}{8\pi G}$ and $p \rightarrow p + p_\Lambda \rightarrow p - \frac{\Lambda c^4}{8\pi G}$, the dark energy component in the Λ CDM has an equation of state, $w = -1$. Generally speaking, the different components of the energy density are related to the scale factor by a power law, i.e., $\rho_i \propto a(t)^{-3(1+w_i)}$. Considering a universe is made up of only matter (i.e., $\Lambda = 0$ and $\rho = \rho_m$), the spatial curvature K only vanishes when the energy density is precisely $\rho_m = \rho_c$, such that, $\rho_c = 3(\frac{\dot{a}}{a})^2/8\pi G$. This is the critical density of the Universe. This critical energy density ρ_c is used to characterize the different

types of geometries that define a universe. Substituting $\frac{\dot{a}}{a}$ with the characteristic density ρ_c in the equation (2.2), we obtain

$$\rho_c = \rho + \frac{\Lambda c^2}{8\pi G} - \frac{3Kc^2}{8\pi Ga^2} = \rho_m + \rho_r + \rho_\Lambda + \rho_k, \quad (2.3)$$

where $\rho_k := -\frac{3Kc^2}{8\pi Ga^2}$ is applied. From the relation above and defining density parameters $\Omega_i = \rho_i/\rho_c$ for i 's corresponding to each type of energy density in the Universe, the spatial curvature density of the Universe at any given time can be determined:

$$\Omega_k = 1 - \Omega_r + \Omega_m + \Omega_\Lambda. \quad (2.4)$$

Assuming that $\Lambda = 0$, the curvature of the space defines, qualitatively, the fate of the universe. For $K = 0$, the total energy density is equal to the critical density. This corresponds to a *flat* universe. The universe will expand forever such that the rate of expansion will eventually halt at an infinite time. This is the EdS universe discussed above. If $K > 0$, then $\Omega_k < 0$, which is interpreted as a universe having more energy density than the critical density. Such a universe without any dark energy can be considered as *closed* and this particular universe will expand to a maximum scale and eventually start shrinking leading to its collapse. For $K < 0$, $\Omega_k > 0$, making the universe *open* as it has less energy density than the critical density. This universe will expand forever. Current observational evidences (e.g., Komatsu et al., 2011) point to a nearly flat Universe.

2.1.1 Cosmological distances and volume: implications for counting clusters

The Hubble parameter is defined as $H := \frac{\dot{a}}{a}$, which is a measure of the expansion rate of the Universe. In an expanding universe, the redshift experienced by light traveling along the shortest path between two comoving coordinates relates to the ratio of scale factor of the universe at two epochs, time of emission t_e and time of observation t_o . The relationship of the redshift (z) to the scale factors is $1+z = \frac{a(t_0)}{a(t_e)}$. By convention, $a(t_0)$ is unity for a present day observer.

From Einstein's field equations for a homogeneous and isotropic universe, we get

$$H(z)^2 = H_0^2 E(z)^2, \quad (2.5)$$

where $E(z)^2 = \Omega_R(1+z)^4 + \Omega_k(1+z)^2 + \Omega_m(1+z)^3 + \Omega_\Lambda$ and H_0 is the Hubble parameter at present time. A characteristic distance that is often used is the Hubble distance defined as $D_H = c/H_0$. The comoving distance in the Friedmann-Robertson-Walker geometry is the distance between two points measured at present cosmic time (t_0).

The line-of-sight comoving distance is

$$D_C = \int_{t_e}^{t_0} c \frac{dt'}{a(t')} = \int_0^r dr/(1-Kr^2) = D_H \int_{z_o}^{z_e} dz'/E(z'), \quad (2.6)$$

where t_e is the cosmic time when photons left the one point and t_0 is the time when the photons reach the observer at another point. There are different ways of measuring distance in cosmology. For an object on the sky which has a physical size of l , the angle subtended by this object today would depend on the distance between the object and the observer. This distance is the angular diameter

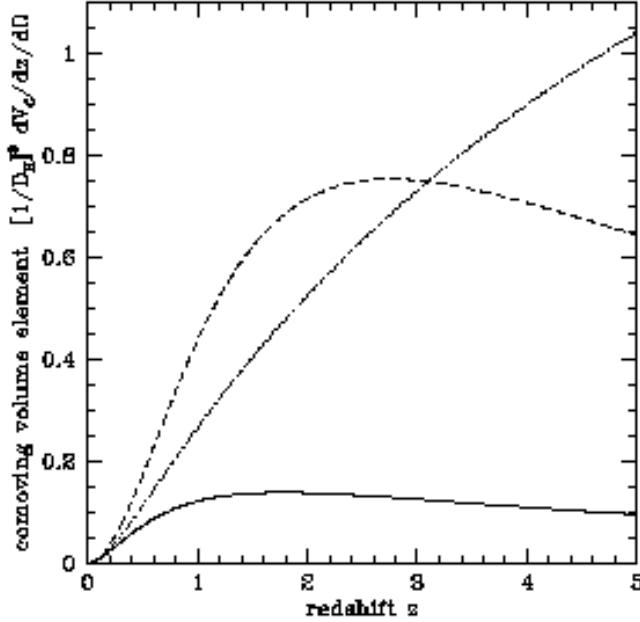


Figure 2.3: The dimensionless comoving volume elements for three different cosmologies as a function of redshift. Solid, dashed and dotted curves are for $(\Omega_m, \Omega_\Lambda)$ set to $(1.0, 0.0)$, $(0.2, 0.8)$ and $(0.05, 0.0)$ cosmologies respectively. Credit: Hogg (1999)

distance D_A . In a flat universe, this relates to the comoving distance as (P. J. E. Peebles, 1993)

$$D_A = \frac{l}{\theta} = D_C/(1+z). \quad (2.7)$$

This serves to relate the angular size of galaxy clusters on the sky to their actual physical size. Another useful distance is the luminosity distance, D_L . This relates the bolometric luminosity (total power emitted by the source) of an astrophysical object to the bolometric flux (total power received per unit area), S , measured by a distant observer. The luminosity distance scales with the comoving distance:

$$D_L = \sqrt{\frac{L}{4\pi S}} = D_C(1+z). \quad (2.8)$$

The luminosity distance is useful in determining how bright a galaxy cluster at a distance from us would appear on the sky. Both of these distances are important in taking a census of galaxy clusters. Finally, to obtain a number density of galaxy clusters in the Universe, a useful definition of a volume element is the comoving volume element. The number density within such a volume element is known as the comoving number density. This number density for non-evolving systems remain constant in redshift. The volume element dV_C is given by the relation

$$dV_C = D_H \frac{(1+z)^2 D_A^2}{E(z)} d\Omega dz, \quad (2.9)$$

where $d\Omega$ is the differential solid angle. Figure 2.3 shows the dimensionless comoving volume element as a function of redshift for three different cosmologies. The one that corresponds closest to a flat universe with a constant non-zero Λ is shown in dashed curve. From this figure, it sticks out that the comoving volume element is highly sensitive to the cosmologies. This necessarily implies that the inferences on the census of galaxy clusters are heavily dependent on the cosmology defining the comoving volume elements.

2.1.2 Evolution of structures through cosmic time

At the early times when matter decoupled from radiation, $z \sim 1100$, the initial density fluctuations were tiny enough to be modelled using a linear perturbation scheme (i.e., $\delta = (\rho - \bar{\rho})/\bar{\rho} \ll 1$). The density variance is characterised by $\sigma^2 = \langle \delta^2 \rangle$. The $\delta(\mathbf{x})$ at initial times, in the most general case, can be seen as superposition of different modes of plane waves. In terms of the Fourier modes the fluctuations are given as

$$\delta_{\mathbf{k}} = \int \delta(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}} d^3\mathbf{x}, \quad (2.10)$$

where \mathbf{k} is the wavenumber vector of a Fourier mode plane wave.

Linear growth factor

At early times, the fluctuations could be treated using the linear perturbation formalism. The linearised equation of motion for density fluctuations is given by

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\rho_m(z)\delta, \quad (2.11)$$

where the matter is treated as pressureless. The Hubble term in the equation acts against the gravitational force of the matter fluctuation. A presence of a pressure exerting component would provide an additional force acting against gravity. A solution to the differential equation for the evolution of the density perturbation from a homogeneous scenario over time is given by the growth factor, D_+ . The complete solution includes a linearly independent solution that behaves as a decaying factor. The second solution over large time scale becomes negligible. The growth factor, D_+ , is proportional to $\frac{H(a)}{H_0} \int_{a_i}^{a_0} \frac{da'}{[\Omega_{m,0}/a' + \Omega_\Lambda a'^2 - (\Omega_m + \Omega_\Lambda - 1)]^{3/2}}$. At an epoch t_0 and at a comoving distance \mathbf{x} , the density perturbation is related to initial perturbation by $\delta(\mathbf{x}, t_0) = D_+(t_0 - t_i)\delta(\mathbf{x}, t_i)$. The position of the density fluctuation does not change over time in the linear regime. For an Einstein-de Sitter universe, $D_+(t) = (t/t_0)^{(2/3)} = a(t)$.

Non-linear structure formation

A simplistic view of cluster formation model can be understood using the spherical symmetric collapse scenario (Gunn and Gott, 1972). As the initial density fluctuations grow with time by the growth factor, at some point the fluctuation is large enough to start worrying about the non-linear terms. To have an intuitive understanding of the formation process of collapsed structures, this spherical symmetric collapse model is instructive. For a spherically symmetric overdensity perturbation, there is an eventual shift in the local geometry that starts deviating from the background global geometry.

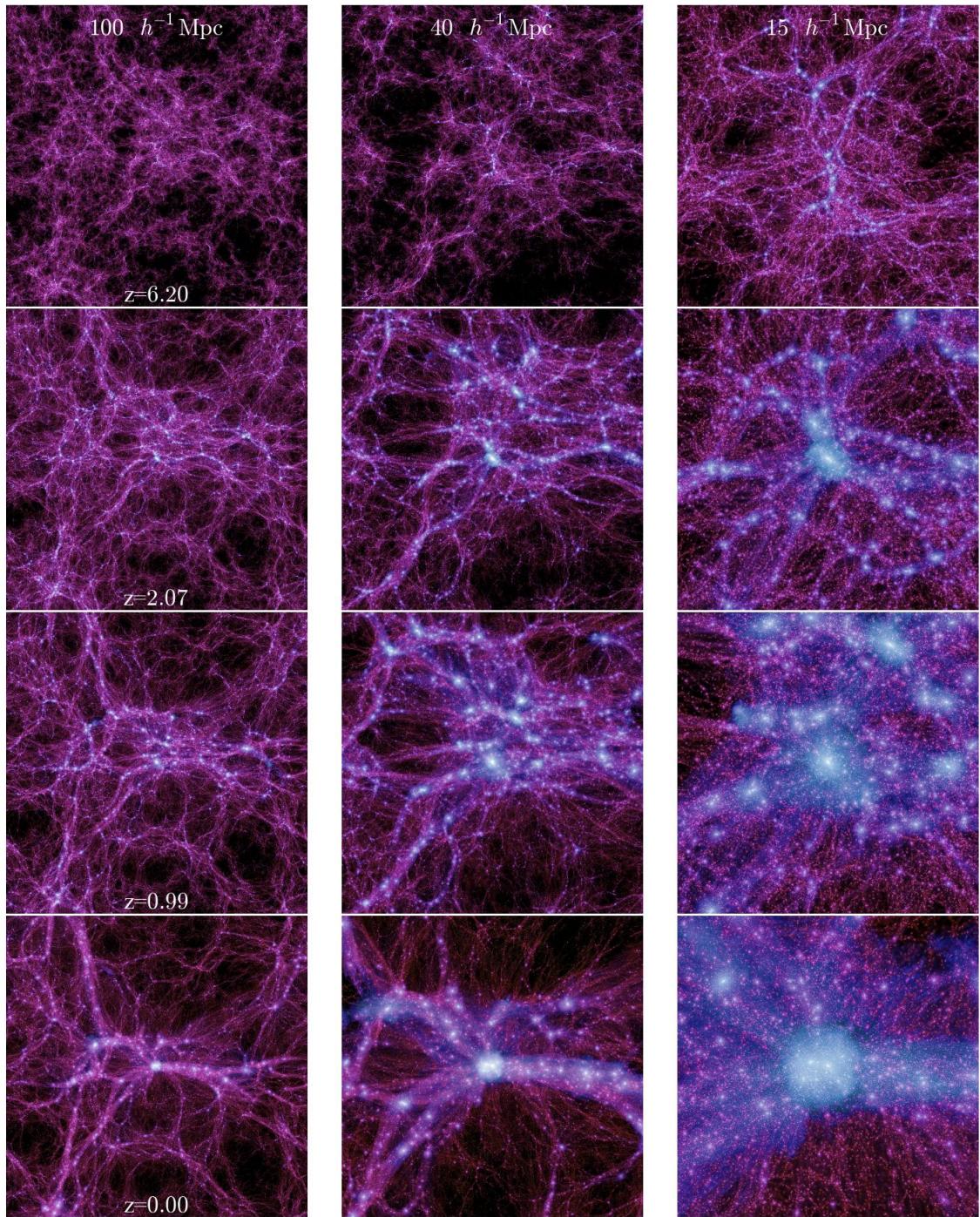


Figure 2.4: Time evolution of a massive halo in Millennium II simulations. Image credit: (Boylan-Kolchin et al., 2009)

For a critically overdense region, it slows down its own expansion and eventually reach a point when the local geometry is similar to a matter dominated “closed” universe.

The equation of motion of a shell enclosing total mass M within a radius r_{sh} is given by

$$\ddot{r}_{\text{sh}} = -\frac{GM}{r_{\text{sh}}^2} - \frac{1+3w}{2}\Omega_{\Lambda}H_0^2(1+z)^{3(1+w)}r_{\text{sh}}, \quad (2.12)$$

as long as the dark energy component is non-clustering (Percival, 2005). This equation of motion is analogous to the equation of motion for a homogeneous and isotropic universe.

For a flat universe with $\Omega_{\Lambda} = 0$, the above equation can be solved parametrically. The solution gives $t(\theta) = B(\theta - \sin \theta)$ and $r_{\text{sh}}(\theta) = A(1 - \cos \theta)$. For a mass shell whose specific energy leads to a bounded state (positive curvature), the growing shell is expected to reach a maximum radius, $R_{\text{max}} = 2A$, at a time $t_{\text{max}} = B\pi$ and then start contracting. From equation (2.1.2) and above relations, we get $A^3 = GM^2$.

If the local geometry is analogous to the Einstein-de Sitter universe in which the sphere itself is embedded in, the radius of the shell grows at decreasing rate yielding $R_{\text{EdS}} \propto t^{\frac{2}{3}}$. For an overdense perturbation, the expansion of the shell would be at a slower rate and slightly less than R_{EdS} initially. Then the density contrast of the local sphere with respect to the EdS density in initial times is given by the relation $\delta(t) = \frac{3}{20} \left(\frac{6t}{B}\right)^{2/3} (1 - \epsilon)$, where ϵ is related to the initial local curvature density. So it can be expected that for a sphere with a geometry analogous to a universe with density greater than than the critical value of δ_c set by the EdS would behave as a “closed” universe whose fate is to collapse eventually into a single point. Time taken to collapse from the initial time is $t_{\text{collapse}} = 2t_{\text{max}} = 2B\pi$. The linear extrapolation of the perturbation in EdS to a time t_{max} is $\delta_{\text{lin}} = 1.06$ and at the time of collapse is $\delta_c \sim 1.69$. This is a critical overdensity such that those perturbation whose linear contrast reaches a value of 1.69 at a point in time would collapse to form structures. This value of critical density is insensitive to difference in background cosmologies (Percival, 2005).

In a realistic scenario, the presence of some small-scale inhomogeneities within the sphere causes fluctuations in gravitational potential as soon as the sphere begins to collapse. For a collisionless matter, the time variations in gravitational potential experienced by the mass particles causes re-distribution of the energies of individual mass particles leading to a *violent relaxation* (Lynden-Bell, 1967) scenario. This process re-distributes the velocities of the mass particles to a Maxwell-Boltzmann distribution in a time scale comparable to the free-fall time (shell-crossing time). The system reaches a virial state by forming a quasi-equilibrium halo of a finite radius R_{vir} .

For the flat matter dominated universe, the overdensity or contrast at virial radius is $18\pi^2 \approx 178$. The overdensity at virial radius relative to a Λ universe background was calculated by Bryan and Norman (1998) as

$$\Delta_c = 18\pi^2 + 82x - 39x^2, \quad (2.13)$$

where $x = \Omega_m(z) - 1$ and $\Omega_m(z)$ is the matter density of the Universe at the time of virialisation. It is common practice to define masses in terms of an overdensity Δ such that the mean density in the halo is $\Delta \rho_{\text{cr}}$.

The above scheme is useful for having an intuitive handle on non-linear structure formation. To solve such non-linear systems, numerical simulations have made much progress in the past few decades (Borgani and A. Kravtsov, 2011). Figure 2.4 shows snapshots of a universe in a numerical simulation

at different redshifts. It illustrates the formation of a massive galaxy cluster (dark matter) halo from a small density perturbation to a non-linearly collapsed structure. It indicates that structures in the Universe form a cosmic web of filamentary structures¹ and massive virialised systems occupy the nodes connecting these web of structures.

Matter density profile of halos

The analytical scheme of spherical collapse in an Einstein-de Sitter universe suggest the dark matter profile within an isothermal equilibrium structure should follow the isothermal profile, i.e., $\rho \propto r^{-2}$. The equilibrium dark matter density profile studied in N-body simulations suggest that the matter density distribution have a universal structure as per the formula (Navarro, Frenk and White, 1995; Navarro, Frenk and White, 1996; Navarro, Frenk and White, 1997):

$$\rho(r) = \frac{\rho_0}{(r/r_s)[1 + (r/r_s)]^2}, \quad (2.14)$$

where r_s and ρ_0 are the characteristic scale radius and the density concentration that vary from one halo to another.

Defining mass of galaxy clusters

To link theoretical modelling of mass to observations, defining the boundary for mass using the spherical overdensity Δ is beneficial. For observing clusters, regions of high density contrast are easier to trace. Therefore, the choice of Δ to be 200 or 500 is often preferred where the contrast is higher than the virial choice. This choice also helps in eliminating the complications in dealing with the infalling matter near the cluster outskirts (the physics near cluster outskirts are reviewed by Reiprich, Basu et al., 2013). By convention, the mass of a galaxy cluster is defined by the spherical overdensity as given below

$$M_\Delta = \Delta \frac{4\pi}{3} \rho_c(z) R_\Delta^3. \quad (2.15)$$

Using the NFW profile in Equation (2.14) and the above mass definition,

$$M_\Delta = \int_0^{R_\Delta} \rho(r) 4\pi r^2 dr = 4\pi \rho_0 r_s^3 \left[\ln(1 + c_\Delta) - \left(\frac{c_\Delta}{1 + c_\Delta} \right) \right], \quad (2.16)$$

where c_Δ is the concentration parameter defined as $c_\Delta := \frac{R_\Delta}{r_s}$. It can be shown that

$$\rho_0 = \Delta \rho_c c_\Delta^3 / [3[\ln(1 + c_\Delta) - c_\Delta/(1 + c_\Delta)]] .$$

Halo mass function

In the hierarchical cluster formation scheme, galaxy cluster abundance can be predicted if the statistics of the initial primordial fluctuations are understood. Press and Schechter (1974) gave an analytical

¹ Bond, Kofman and Pogosyan (1996) explain the coherence of filamentary network of structures depends on the shape of the initial perturbation power spectrum, n_s .

formalism for the prediction of galaxy cluster abundance by assuming a Gaussian random field for initial perturbations and the spherical collapse of structures. Other analytical extensions of the model were suggested for improving consistency with N -body simulations (Sheth and Tormen, 1999). However, analytic models do not capture the realistic halo formation and since then N -body simulations have been used to fit various parameterised mass functions (Jenkins et al., 2001; Sheth, Mo and Tormen, 2001; Tinker et al., 2008; Angulo et al., 2012). Among these earlier works, Tinker et al. (2008) set the standard for providing the mass function in terms of spherical overdensity masses (i.e., $\Delta = 200, 500$), which are much easier to link to measurements of masses of galaxy clusters.

The initial scalar perturbations are assumed to have Gaussian random fields and evidence from CMB measurements find no significant evidence for deviations from gaussianity (Planck Collaboration, Ade, P. A. R. and Aghanim, N., 2016). The statistics of such fluctuations then can be easily described in Fourier space. For Gaussian statistics, the fluctuations are fully characterized by the amplitude of fluctuations (σ) and the two-point correlation or an auto-covariance function $\xi(r)$. An instantaneous primordial power spectrum such that there is no characteristic scale can be generally given as a power-law function of wavenumber $P_{\text{prim}}(k) \propto k^{n_s}$, where the spectral index n_s evolves and grows with the Universe. The matter power spectrum at any given instant can be understood by the following equation

$$P_m(k, a, \theta_\Omega) = D_+^2(a, \theta_\Omega) T^2(k, \theta_\Omega) P_{\text{prim}}(k), \quad (2.17)$$

where θ_Ω is set of cosmological parameters that describe the background cosmology, $T(k, \theta_\Omega)$ is the transfer function which corrects for the evolution before recombination in the radiation dominated era. The differential form of the cluster number density is

$$\frac{dn}{dM} = \frac{\bar{\rho}_m}{M} \left| \frac{d \ln \sigma}{dM} \right| f(\sigma, z), \quad (2.18)$$

where $n(M, z)$ is the population density in units of number per unit comoving volume. $\bar{\rho}_m$ is the comoving mean matter density $\Omega_m(z)\rho_{\text{cf}}(z)$. $\sigma = \sigma(M, z)$ in the above equation is the variance of the smoothed linear perturbations in matter density (cold dark matter) and baryons on mass scale $M = (4\pi/3)\rho_c R^3$:

$$\sigma^2(M, z) = \int \frac{d^3 k}{2\pi} W^2(kR) P_m(k, z) = \xi(0, z), \quad (2.19)$$

where $W(kR)$ is the Fourier transform of a spherical top-hat filter of radius R . $\sigma(M)$ is a monotonic decreasing function of the smoothing radius. $f(\sigma)$ is a parameterised model function of the form $A[(\sigma/b)^{-a} + 1] e^{-c/\sigma^2}$ (Tinker et al., 2008). This function exponentially cuts off the mass distribution due to the rarity of peak fluctuations being at the tail of a Gaussian distribution.

Setting the smoothing radius at $R = 8h_{70}^{-1}\text{Mpc}$ and evaluating the relation in Equation (2.19) at redshift zero gives σ_8 . This quantity registers as an amplitude for the linear matter power spectrum today. The convention of adopting $8h_{70}^{-1}\text{Mpc}$ arises from measuring fluctuations in number of galaxies within this radius $\delta n_{\text{gal}}/n_{\text{gal}} \sim 1$ (Davis and P. J. E. Peebles, 1983).

Until more recently, the impact of baryons on the halo abundance was sparsely studied. With the help of high resolution and large volume hydrodynamic simulations called *Magneticum*, it is now made possible (Bocquet et al., 2016). The mass function including baryons reported from that work is shown in Figure 2.5. For mass ranges 10^{14} – $10^{15} M_\odot$ and redshift range 0.13–0.8, (Bocquet et al., 2016) find a negligible impact on the mass function due to baryons and is in good agreement (< 10% and this difference reduces with mass in the considered range) with the mass function provided by

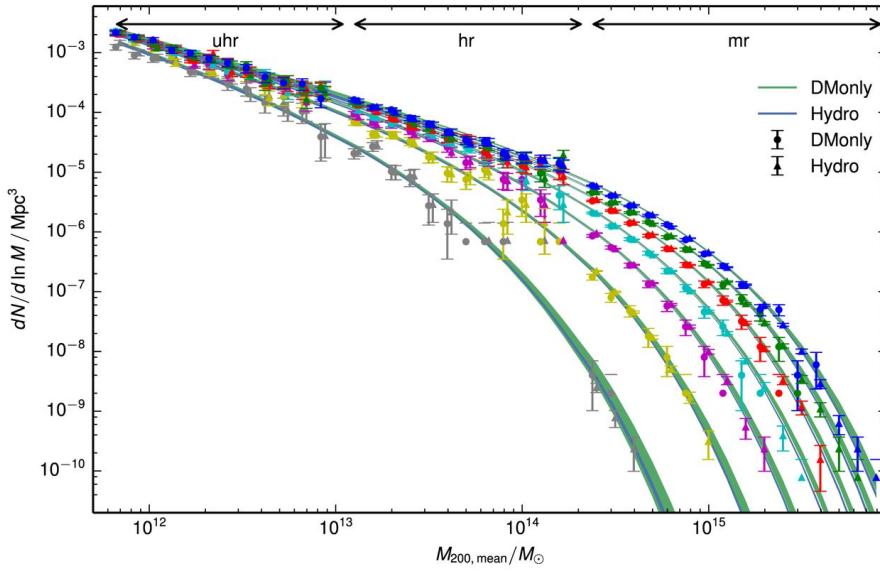


Figure 2.5: Halo mass function for spherical overdense masses from hydrodynamical and dark matter only simulations. The lower curves are at relatively higher redshifts (with the lowest one corresponding to $z = 2.0$) than the top most curve (corresponding to $z = 0$). The mass function is a monotonic decreasing function of mass and redshift. This qualitative behaviour is observed in both dark matter only and hydrodynamic simulations. Credit: Bocquet et al. (2016).

Tinker et al. (2008). We note this, since, we use the Tinker mass function in Chapter 6 to model the cluster population distribution for our measurements of the statistical relationships of cluster masses to the cluster observables.

Massive neutrinos and dark energy

For the most part of the history of the Universe and from the time neutrinos decoupled from matter and radiation, neutrinos were relativistic free-streaming particles. Eventually, with the expansion of the Universe, they cooled down such that their energy density is dominantly given by their mass and negligible momentum. In the time that the massive neutrinos were relativistic, they behaved as a non-clustering mass component (at scales smaller than the free-streaming scale). This led to a suppression in the growth of structures at small scales. By studying the abundances of structures in different mass scales and redshift, it can provide constraints on the total energy density of neutrinos, hence, their sum of masses. Much of these constraints in the current cluster cosmological studies is driven by the difference in the amplitude of the matter power spectrum estimated at early times from the CMB and at present day from the cluster abundance measurements (e.g., A. B. Mantz et al., 2015; Abazajian et al., 2016). Future galaxy cluster surveys probing the abundance and its evolution can be a powerful probe for the detection of the sum of neutrino masses in combination with other probes (Wang et al., 2005).

The dark energy enters in cluster abundance through the comoving volume geometry and in the growth rate. In terms of the growth rate, it acts against the gravitational potential and slows down the growth of structures. Therefore, galaxy cluster abundance and its evolution is a sensitive probe

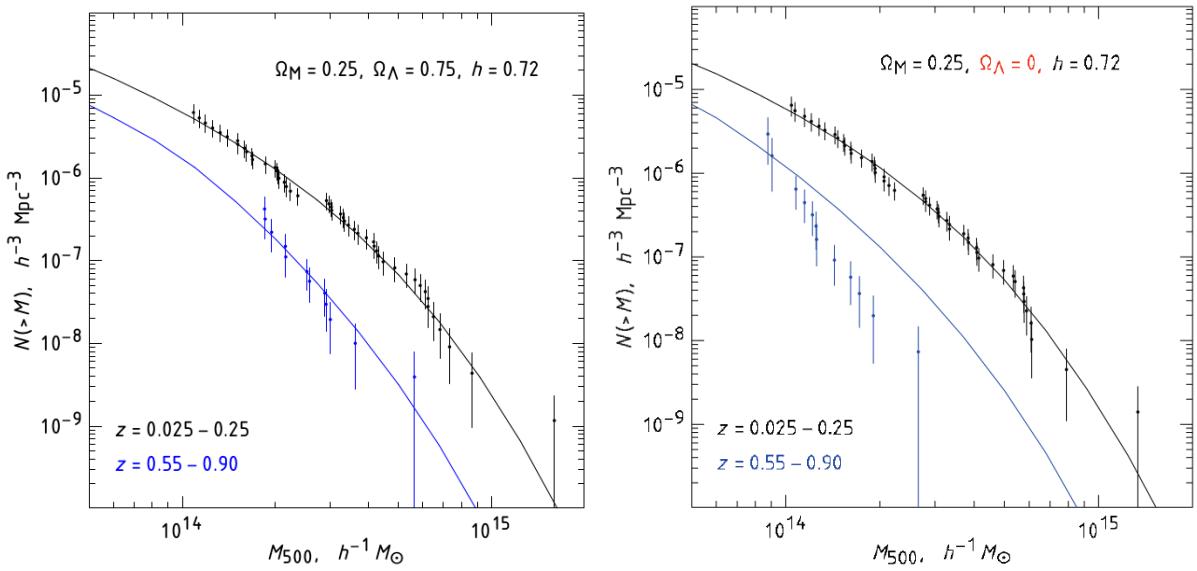


Figure 2.6: Measured mass function (data points) and model prediction (solid curves). *Left:* The model prediction of the mass function for a non-zero cosmological constant and lower matter density universe is shown along with the measured mass function for two redshift ranges. The normalisation of the mass function was adjusted to match the zero redshift measurements. *Right:* The measured mass function and model prediction for a low matter density universe without dark energy is shown. The mismatch between the measured abundance and prediction in the higher redshift bin demonstrates the sensitivity of growth of structures to cosmological parameters. Credit: Vikhlinin, A. V. Kravtsov et al. (2009).

of dark energy content of the Universe. This was demonstrated by Vikhlinin, A. V. Kravtsov et al. (2009) using measured census of galaxy clusters. Figure 2.6 from their work illustrates the cluster abundance in different redshifts can determine the geometry and content of the Universe.

2.2 Galaxy cluster observables in multi-wavelength

So far in this chapter, mass of galaxy clusters were treated as if they could be known precisely. The reality is far more precarious. As already mentioned in the beginning of this chapter, most of the matter in a galaxy cluster is made up of dark matter. The highest fraction of the ordinary matter resides in the intra-cluster gas medium. Direct observations of dark matter are not possible in any wavelength. Consequently, the cluster masses have to be indirectly inferred through the clues offered by observations of galaxy clusters in the EM spectrum. Here, a short review of galaxy clusters properties in multi-wavelength is presented.

2.2.1 X-rays

The temperatures of the ICM in massive galaxy clusters are driven to their virial temperatures in the range $10^7 - 10^8$ K. At such high temperatures the gas medium is found ionised. Felten et al. (1966) first sought to explain the diffuse X-ray emission from Coma cluster with the thermal electrons emitting bremsstrahlung radiation. The free electrons experience deceleration when they interact with

positive ions in the plasma and emit radiation (free-free emission by *Bremsstrahlung* process). The emissivity of such a process for an optically thin plasma is given by the relation (Sarazin, 1988)

$$\epsilon^{\text{ff}}(\nu) \propto Z^2 n_e n_i g_{\text{ff}}(Z, T_e, \nu) T_e^{-1/2} \exp\left(-\frac{h\nu}{k_B T_e}\right), \quad (2.20)$$

where k_B is the Boltzmann constant, Z is the metal abundance, T_e is the temperature of the thermal gas medium, g_{ff} is the gaunt factor, n_e and n_i are electron density and ion density respectively. The gas medium is optically thin with low density. For ICM with temperatures larger than 2 keV, the primary radiative process is the Bremsstrahlung emission. Apart from these, the gas may also emit X-rays by recombination (free-bound) and bound-bound line emission from the metals in the ICM. The emissivity is exponentially cut-off at energy scales higher than the thermal energy of electrons. Thus, the emission spectra has a characteristic cut-off related to the temperature of the medium. If one were to observe the full spectra, the temperature of the ICM could be measured through the X-ray emission. The bolometric emissivity, ϵ^{ff} , is obtained by integrating $\epsilon^{\text{ff}}(\nu)$ over all frequencies giving $\epsilon^{\text{ff}} \propto T^{1/2} n_e^2$. For a sufficiently hot cluster medium, the emissivity integrated within a band of energies that are much lower than the thermal energy scale is sensitive to n_e^2 with negligible temperature dependence. The surface brightness of the emission is

$$S_X(R) = \frac{D_A^2}{4\pi D_L^2} \int_{-\infty}^{+\infty} \epsilon dl, \quad (2.21)$$

where R is the projected radius distance on the sky from the center of the medium, dl is the differential line-of-sight element and ϵ is the total emissivity. The surface brightness is sensitive to the square density of the gas medium and it also undergoes dimming by the luminosity distance. The hydrostatic equilibrium for the ICM that is spherically symmetric is described by

$$\frac{1}{\rho_{\text{gas}}} \frac{dP}{dr} = -\frac{GM_{\text{tot}}(< r)}{r^2}, \quad (2.22)$$

where ρ_{gas} is the gas density. The mass can be determined from ICM temperature and density estimations by using the hydrostatic equilibrium condition

$$M_{\text{tot}}(< r) = -\frac{k_B T_e r}{G \mu m_p} \left(\frac{d \ln \rho_{\text{gas}}}{d \ln r} + \frac{d \ln T_e}{d \ln r} \right), \quad (2.23)$$

where μ is the mean molecular weight. The hydrostatic mass has a strong dependence on the temperature profile. The hydrostatic masses are expected to be biased by 10-20 % (e.g., Nagai, A. V. Kravtsov and Vikhlinin, 2007; Meneghetti et al., 2010; Rasia et al., 2012).

The *ROSAT all sky survey* (RASS) using the X-ray imaging *ROSAT* satellite has been vital in constructing several X-ray selected galaxy cluster catalogues. Several samples have been constructed using different flux cuts or luminosity cuts (e.g., Ebeling, Edge, Böhringer et al., 1998; Böhringer, Voges et al., 2000; Böhringer, Schuecker et al., 2004; Reiprich and Böhringer, 2002) One of the largest complete sample of galaxy clusters with ~ 900 of them discovered in X-ray is given as REFLEX II sample (Böhringer, Chon, Collins et al., 2013), which covers the southern sky. In the northern sky, NORAS II has about ~ 800 clusters (Böhringer, Chon, Retzlaff et al., 2017). The largest cluster catalogue in X-ray is the CLASSIX (Böhringer, Chon and Kronberg, 2016), which combines these two catalogues. The median redshift of the samples is ~ 0.1 and maximum redshift is ~ 0.5 .

X-ray emission suffers from dimming effect, which makes it difficult to reach sources at very high redshifts. Few X-ray serendipitous discoveries occurred with ROSAT deep pointed observations in a smaller area of the sky. The sample from this 400 degrees square region is constructed by Burenin et al. (2007) which could reach sources upto $z \sim 0.9$. In the near future, eROSITA (Merloni et al., 2012) will be sensitive enough to discover ~ 100000 galaxy clusters and groups with the goal of constraining the dark energy.

2.2.2 Microwaves (Sunyaev-Zel'dovich effect)

The CMB radiation is a relic of the early Universe. The hot thermal electrons in the ICM produce secondary anisotropies in the CMB (Sunyaev and Zel'dovich, 1970; Sunyaev and Zel'dovich, 1972). As the CMB pass through the hot intra-cluster medium, its black body spectral signature is distorted by the inverse Compton scattering. These distortions are very tiny and difficult to measure. It was first predicted to occur in galaxy clusters by Sunyaev and Zel'dovich (1970). The CMB photons when encountering hot electrons of the ICM in its path gain energy and this re-distributes the energy of the photons while preserving the number of photons. This leads to a deficit in the CMB intensity at low energies and an increment at higher energies.

The first marginal detection of this effect in known galaxy clusters was reported in Gull and Northover (1976) at 10.6 GHz. After nearly a decade, subsequent strong observational evidence of the SZ effect were seen (see review by Birkinshaw, 1999, and references therein).

The details of the derivation of the SZ effect can be found in various references (Sunyaev and Zel'dovich, 1972; Birkinshaw, 1999). The CMB has a black body spectrum as shown in Figure 2.7. For a black body, its temperature (T_{CMB}) and intensity (I_ν) are related. The SZE distortion, ΔT_{SZE} , of the cosmic microwave background temperature T_{CMB} , is given by

$$\frac{\Delta T_{\text{SZE}}}{T_{\text{CMB}}} = \frac{\Delta I_\nu}{I_\nu} \frac{d \ln T_{\text{CMB}}}{d \ln I_\nu} = f(x) \int \sigma_T n_e \frac{k_B T_e}{(m_e c^2)} dl = f(x)y, \quad (2.24)$$

where l is the line of sight variable, σ_T is the Thomson scattering cross-section for electrons, m_e is the electron mass, k_B is the Boltzmann constant, and c is the speed of light. T_e is the electron temperature of the X-ray emitting plasma and $f(x)$ gives the spectral shape of the effect, given by

$$f(x) \equiv \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) (1 + \delta_{\text{SZE}}(x, T_e)), \quad (2.25)$$

where x is the dimensionless frequency related to the frequency by $x = h\nu/k_B T_{\text{CMB}}$. $\delta_{\text{SZE}}(x, T_e)$ is a correction due to relativistic effects (e.g., Itoh, Kohyama and Nozawa, 1998). The frequency independent measure y is the line-of-sight Compton parameter, proportional to the electron pressure integrated along the line of sight as

$$y = \frac{\sigma_T}{m_e c^2} \int P_e(l) dl, \quad (2.26)$$

where $P_e = n_e k_B T_e$ is the electron pressure. For an optically thin hot plasma, the distortion amplitude y is extremely small ($\sim 10^{-4}$). In Figure 2.7, the distorted spectra are shown for different values of y .

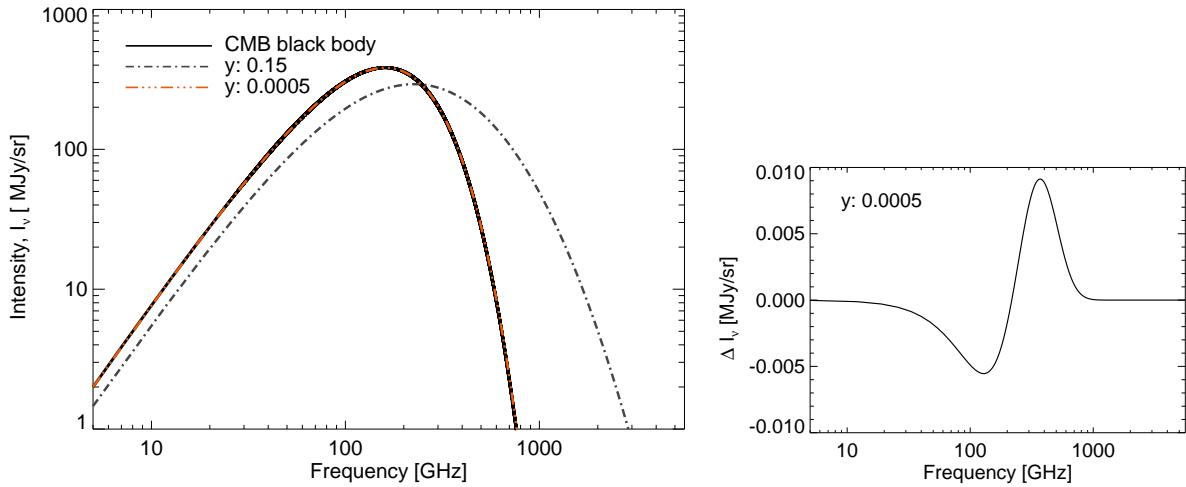


Figure 2.7: The distortion caused by the hot ICM in the CMB spectrum. The CMB spectrum is a black body (black solid). To demonstrate the distortion of the spectrum, an unrealistically large Compton parameter ($y = 0.15$) is chosen and this spectrum is represented by the grey dotted-dashed curve. The typical distortion amplitude for galaxy clusters are estimated in the order of 10^{-4} (orange dashed curve). In order to distinguish the distorted spectrum from black body, (right) the differential intensity is shown.

The integrated Compton parameter, denoted Y , is defined by

$$Y = \int y d\Omega, \quad (2.27)$$

where the integration is over solid angle Ω in a given aperture, resulting in a cylindrically integrated quantity which we shall refer to as Y_{cyl} . Given an azimuthally symmetric radial model, Y_{cyl} can be converted to the spherical counterpart Y_{sph} , representing the integrated Comptonization in a sphere of corresponding radius. The SZ Comptonization in terms of its physical units (or extent) is given by $Y_{\text{SZ}} = D_A^2 Y_{\text{sph}}$, where D_A is the angular diameter distance of the cluster determined by cosmology and redshift.

Unlike X-ray luminosity, the SZ effect does not suffer from the dimming effect with redshift. This is mainly because the effect is a fractional change in the CMB brightness, which itself grows with redshift by $(1 + z)^4$. This cancels out the dimming factor and effectively makes the SZ effect redshift independent. Owing to this, this effect is a powerful in discovering galaxy clusters upto high redshifts for a given mass. However, the SZ effect does suffer from projection effect as it probes the total line-of-sight thermal electron pressure. Below is a short review of galaxy clusters surveys mapping the SZ effect in large fraction of the sky.

The potential of the SZ effect in galaxy clusters as a cosmological probe was reviewed by Carlstrom, Holder and Reese (2002). In the past few years, many experiments have undertaken survey of SZ effect in large area of the sky such as *the Atacama Cosmology Telescope* (ACT: Marriage et al., 2011; Hasselfield et al., 2013), *South Pole Telescope* (SPT: Vanderlinde et al., 2010; Reichardt, Stalder et al., 2013; Bleem et al., 2015), *Planck satellite* (Planck: Planck Collaboration, Ade, Aghanim, Arnaud et al., 2011; Planck Collaboration, Ade, Aghanim and al., 2016b). The largest cluster samples are of

¹ X-ray clusters taken from Meta-catalogue by Piffaretti et al. (2011).

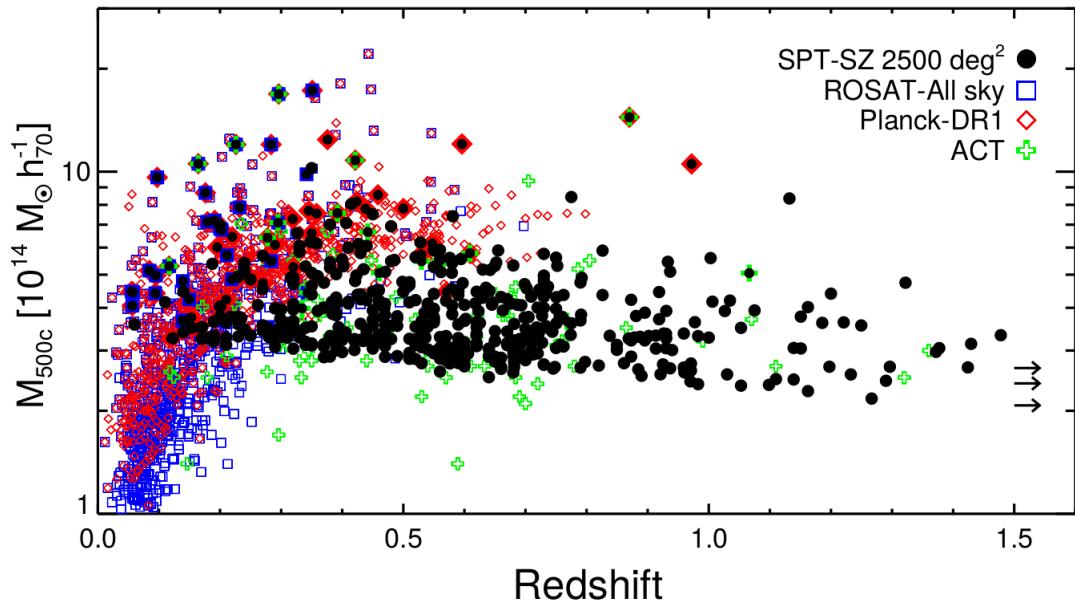


Figure 2.8: Cluster catalogs from SZ and X-ray selections. The 516 optically confirmed clusters from SPT catalog, 91 clusters from ACT survey, 809 SZ-selected clusters from Planck, 740 X-ray clusters¹ from ROSAT observations are shown here. The SPT selection is approximately mass selected, whereas the X-ray selection suffers from dimming effect with redshift, the Planck selection suffers from dilution of SZ signal due to the beam. Credit: Bleem et al. (2015).

a size comprising $\sim 500 - 1000$ clusters in the current SZ selected catalogs (Planck Collaboration, Ade, Aghanim and al., 2016b; Bleem et al., 2015). In particular, the SPT has demonstrated that SZ experiments with sufficient resolution produce galaxy cluster catalogs that are approximately mass-limited upto high redshifts (Bleem et al., 2015). A summary of current galaxy cluster catalogues from SZ and X-ray is represented in Figure 2.8. The representation of the samples in terms of M_{500c} is possible here by using a statistical baseline model connecting cluster observables (discussed in Section 2.2.1 and here) to mass. Such a need for a baseline model and the reason for expecting such a baseline link between observables and cluster masses is described in Section 2.3. The above mentioned surveys form the Stage I and II of the CMB experiments. In the near future, CMB experiments such as , SPT-3G, AdvACT will ramp up the number of galaxy cluster in SZ samples to $\sim 5000 - 10000$ clusters (Benson, Ade et al., 2014; Henderson et al., 2016). The Stage-IV CMB experiments are predicted to increase the survey sample size in SZ to 100000 within next few years (Abazajian et al., 2016).

Apart from survey experiments, few experiments performed targeted observations of numerous galaxy clusters in micro-wavelengths to better understand the pressure distribution in individual clusters (e.g., Romero et al., 2016; Basu et al., 2010; J. Sayers et al., 2013) and statistical link between global SZ observable and mass (e.g., Czakon et al., 2015; Marrone et al., 2012). These include the APEX-SZ experiment (Dobbs et al., 2006; Schwan et al., 2011) which is discussed in detail in Chapter 3. In this thesis, the observations and results from APEX-SZ will be used to understand the relationship of the SZ effect to cluster masses in a statistical manner. The motivation for the statistical link is discussed in Section 2.3.

Kinetic SZ effect

This is an additional distortion that occurs when the cluster is moving with a bulk velocity with respect to the CMB rest frame. This effect is generally 10 times fainter than the thermal SZ effect in galaxy clusters:

$$\frac{\Delta T_{\text{SZE}}}{T_{\text{CMB}}} = -\tau_e \frac{v_{\text{pec}}}{c}, \quad (2.28)$$

where τ_e is the optical depth of the free electron population. The correlation length of the bulk velocity is assumed to be larger than the galaxy velocity correlation. In the non-relativistic regime, the kSZ effect simply distorts the thermal spectrum in terms of the amplitude. The spectrum is still described by Planck spectrum but with smaller or larger temperature depending on the positive or negative direction of the peculiar velocity (Sunyaev and Zel'dovich, 1972; Phillips, 1995; Birkinshaw, 1999). More recently, J. Sayers, Mroczkowski et al. (2013) reported the first significant direct detection of the kSZe signal in a single cluster. The first detection of the kSZe signal was reported by Hand et al. (2012) using the pairwise momentum statistics from CMB observations with Atacama Cosmology Telescope (ACT). The expectation is that on an average pairs of massive collapsed structures should move towards each other and such a momentum would relate to the distance between the structures. Using a similar approach, Soergel et al. (2016) used the South Pole Telescope CMB data to detect the pairwise kSZe signal with a significance of 4.2σ . Planck Collaboration and et al. (2016b) also reported a $1.8 - 2.5\sigma$ detection of pairwise kSZe signal. One of the key advantages of kSZe is that it is a powerful tool for directly probing the electron momentum. Unlike the thermal SZe which is only useful in tracing hot thermal electrons, the kSZe can probe the total line-of-sight electron distribution that is moving with the bulk motion irrespective of their temperature distribution. This makes it a powerful probe of optical depth of electron population in clusters of galaxies and galaxies. This could also potentially help in probing baryons that are residing in cosmic web and help solve the long found “missing baryons”² problem (Ferraro et al., 2016; Planck Collaboration and et al., 2016b).

2.2.3 Optical and near-infrared

At optical wavelengths, galaxies and stars within galaxy clusters are visible. It makes probing galaxy distribution and properties of the cluster through starlight possible. Galaxy clusters also modify the image of galaxies that are far behind galaxy clusters through gravitational lensing. This gravitational lensing depends on the mass distribution of a galaxy cluster. This provides a way to map the mass distribution arising from all matter within a galaxy cluster. These two aspects are reviewed here.

2.2.3.1 Probing starlight

The earliest identification of galaxy cluster systems occurred in optical wavelength and were characterized by concentration of galaxies (Abell, 1958; Zwicky et al., 1961; Abell, Corwin and Olowin, 1989). The overdensity of galaxies are typically characterized by the term *richness*. However, detecting galaxy clusters in optical is subject to projection effects. The projected distribution of galaxies may be contaminated with galaxies that do not belong to the same virialised structure leading to

² A census of baryons in the Universe from observations tells us that the fraction of the baryons predicted by the BBN is still unaccounted for Shull, B. D. Smith and Danforth (2012).

some false-positives in the identification of clusters. The virial structures can be confirmed from velocity dispersion measurements of galaxies. The distribution of velocities of galaxies in a relaxed system is expected to follow a Gaussian distribution. By measuring the velocity distribution, possible superposition of unrelated structures could be ruled out as a galaxy cluster. Zwicky (1933) measured the velocity dispersion for Coma galaxy cluster and concluded that the measured dispersion suggest the total mass of the cluster must be larger than the estimated mass from accounting for the visible galaxy population. This was the first evidence of dark matter. Velocity dispersions are useful in measuring mass of galaxy clusters.

Modern optical detections and observations of galaxy clusters use multi-colour band motivated by the expectation that the galaxy population in clusters are dominated by red elliptical galaxies in the redshift range $z < 1.0$ (e.g., Bower, Lucey and Ellis, 1992; Hennig et al., 2017). Historically, optical search of galaxy clusters have produced large cluster samples (e.g., Gladders and Yee, 2000; Koester et al., 2007; Wen, Han and Liu, 2012; Rykoff et al., 2016). The size of the cluster samples constructed from optical search are the largest with the over 10 000 candidates (Koester et al., 2007; Wen, Han and Liu, 2012; Rykoff et al., 2016). Such large samples are now being used for cosmological studies (e.g. Rozo et al., 2010), but their full constraining potential is not reached yet, owing to remaining uncertainties and systematics in the selection and mass- observable calibration.

2.2.3.2 Probing matter distribution: Weak lensing effect

A complete review on weak gravitational lensing theory and its application to galaxy clusters as lens systems is given by Bartelmann and Maturi (2017). A light ray leaving a source and traveling along a path experience the gravitational potential of massive structures along the way which deflect the light from its geodesic path in absence of these massive structures. Galaxy clusters can be approximated as a thin gravitational lens system. This lens system is represented in Figure 2.9, where a galaxy cluster behaves as a convex lens system. Three distances characterise the lens system and these are the angular diameter distances: distance between a source and observer D_s , distance of the deflecting lens from observer D_l , distance between the lens and source D_{ls} .

The lens equation for a deflector with gravitational lensing potential ψ is:

$$\vec{\beta} = \vec{\theta} - \vec{\nabla}\psi, \quad (2.29)$$

where $\vec{\beta}, \vec{\theta}$ have the same meaning as in Figure 2.9. $\vec{\nabla}\psi$ dictates the value of the reduced deflection angle $\vec{\alpha}$. The reduced angle $\vec{\alpha}$ and the deflection angle at the lens $\vec{\alpha}$ are related by the characteristic distances, $\vec{\alpha} := \frac{D_{ls}}{D_s} \vec{\alpha}(\vec{\theta})$. For a sufficiently small background source, the above equation can be linearised in the differential using Taylor expansion as:

$$\delta\vec{\beta} = A\delta\vec{\theta}, \quad (2.30)$$

where A is a Jacobian matrix and its elements are given by

$$A_{ij} = \frac{\partial\beta_i}{\partial\theta_j} = \delta_{ij} - \frac{\partial^2\psi}{\partial\theta_i\partial\theta_j}. \quad (2.31)$$

The Jacobian can be decomposed into an isotropic and anisotropic components. The isotropic term is

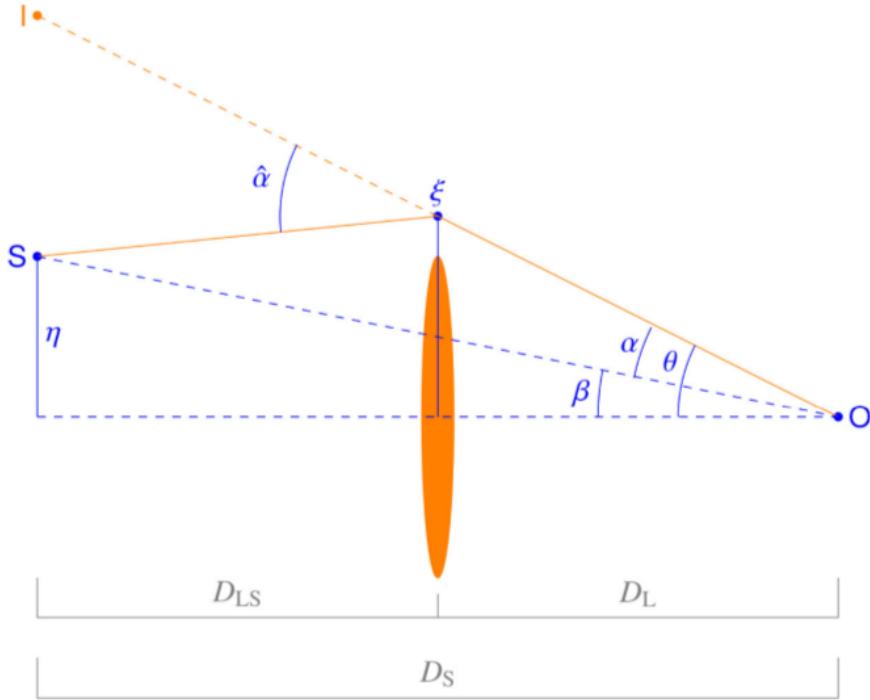


Figure 2.9: A sketch of a gravitational lens system. β is the angle between the source S and the optical axis. The angle between the image and the optical axis at the observer is θ . The deflection angle of the light ray at the lens is $\hat{\alpha}$. α is the reduced deflection angle of the image at the observer. Credit: Bartelmann and Maturi (2017), originally adapted from Narayan and Bartelmann (1999).

given by the trace of the matrix. The anisotropic term is a shear matrix defined as

$$\Gamma := -\left(A - \frac{1}{2}tr(A)I\right), \quad (2.32)$$

where $tr(A)$ is $2 - \kappa$ and κ is $\vec{\nabla}^2\psi$. κ is the convergence.

In the weak lensing regime the matrix A is invertible as its determinant relatively close to 1. The matrix A^{-1} can be explicitly written out as

$$\begin{bmatrix} 1 - \kappa + \gamma_1 & \gamma_2 \\ \gamma_2 & 1 - \kappa - \gamma_1 \end{bmatrix},$$

where $\gamma_1 := \Gamma_{11} = \frac{1}{2}(\partial^2\psi/\partial\theta_1^2 - \partial^2\psi/\partial\theta_2^2)$, $\gamma_2 := \Gamma_{12} = \Gamma_{21} = \partial^2\psi/\partial\theta_1\partial\theta_2$.

The above relations suggest that the image of the background source undergoes magnification and deformation. The *magnification factor* is given by

$$\mu = \det A^{-1} = \frac{1}{(1 - \kappa)^2 - \gamma_1^2 - \gamma_2^2} \approx 1 - 2\kappa, \quad (2.33)$$

where the second relation comes from retaining the linear terms from Taylor expansion.

A circular galaxy behind the lens would appear enlarged and elliptical due to the convergence and shear. The reduced shear or the ellipticity of the image of the galaxy is given by

$$g = \gamma/(1 - \kappa). \quad (2.34)$$

By observing sufficient number of background galaxies, the average ellipticity of the galaxies can be related to the lensing potential of the cluster. This is true as long as the galaxies inherently have random orientations and ellipticities. The first detection of the weak-lensing effect by galaxy clusters was reported by Tyson, Valdes and Wenk (1990). Few authors constructed different samples of the order of 20 – 50 clusters for dedicated lensing measurements of some already known massive clusters. The *Canadian Cluster Comparison Project* (CCCP) constructed a sample of 50 clusters using a heterogeneous selection of X-ray luminous massive clusters based on their strong lensing properties (Hoekstra, 2007; Hoekstra, Mahdavi et al., 2012; Hoekstra, Herbonnet et al., 2015). These clusters were studied for obtaining lensing measurements of mass from Canadian-France-Hawaii Telescope observations (Hoekstra, Mahdavi et al., 2012; Hoekstra, Herbonnet et al., 2015). Another group followed-up a X-ray luminosity limited sample of 50 clusters from RASS catalogues (Ebeling, Edge, Böhringer et al., 1998; Ebeling, Edge, Allen et al., 2000; Böhringer, Schuecker et al., 2004) called *Local Cluster Substructure Survey* (LoCuSS) (Okabe, G. P. Smith et al., 2013; Martino et al., 2014; G. P. Smith et al., 2016). Using the Suprime-Cam mounted on the 8.2-m Subaru telescope, these authors measured the lensing masses of these clusters (Okabe, Takada et al., 2010; Okabe, G. P. Smith et al., 2013; Okabe and G. P. Smith, 2016). The *weighting the giants* (WtG) programme measured lensing masses for 51 X-ray selected clusters using *Subaru* and CFHT observations and the analysis is collectively described in von der Linden et al. (2014), P. L. Kelly et al. (2014), D. E. Applegate et al. (2014). Another sample of clusters called *the Cluster Lensing and Supernova Survey with Hubble* (CLASH Postman et al., 2012), made up 25 clusters selected by X-rays and lensing magnification, were studied using lensing by Umetsu, Medezinski et al. (2014). From all of the above mentioned samples, only LoCuSS sample has some well-defined selection criteria for the population of galaxy clusters. Most of these work tried to optimise the sample such that the lensing signal could be measured with sufficient or high signal to noise.

The previously noted literature primarily studied clusters at relatively low redshifts (median $z \sim 0.3$). The serendipitous sample of clusters from the deep ROSAT pointed 400d survey (Burenin et al., 2007) which are on average lower mass and with median redshift of 0.5 were also studied for lensing by Israel et al. (2012) and Doria et al. (in preparation). The advent of the SZ detected clusters has required lensing observations of some high redshift clusters, which has also been studied for a few number of structures (median $z = 0.9$; Schrabback, D. Applegate et al., 2016). The weak lensing of background galaxies become increasingly challenging to measure for high redshifts.

Simulations show that lensing measurements are biased by only by a few percent (Meneghetti et al., 2010; Becker and A. V. Kravtsov, 2011; Rasia et al., 2012). They directly probe the dark matter distribution and provide a good anchor to cluster masses. However, the weak lensing being sensitive to all matter distribution along the line-of-sight, the masses inferred from lensing observations suffer uncertainties at 20% level by triaxiality and projections of large scale structures (Becker and A. V. Kravtsov, 2011; Gruen et al., 2015). Many of these lensing measurements have been used to study the statistical relationship of cluster observables and mass (e.g., Schrabback, D. Applegate et al., 2016; Hoekstra, Herbonnet et al., 2015; G. P. Smith et al., 2016; Penna-Lima et al., 2016; Marrone

et al., 2012). It can be seen that the use of lensing masses for absolute mass calibration of cluster observables has been possible only in recent years. In this thesis, lensing measurements by Klein et al. (in preparation) of an X-ray selected sample of 27 clusters (eDXL sample with median $z \approx 0.3$, introduced in Chapter 3) will be used for measuring the relationship of cluster masses to cluster observables. A summary of the lensing analysis for this sample is presented in Chapter 3.

2.3 Linking cluster observables to mass: scaling relations

Detailed multi-wavelength studies of galaxy clusters are ideal for measuring individual cluster masses. These are especially important for cluster systems that deviate from relaxed dynamical state. Several authors have pursued performing detailed multi-wavelength analyses to trace the baryon and dark matter distribution in few clusters (e.g., Umetsu, Sereno et al., 2015; Jee et al., 2012). Only a fraction of galaxy clusters known today has been studied in galaxy weak-lensing as already described in Section 2.2.3.2. But it stands that for precision cosmology, we need mass measurements for a large sample of clusters and the size of the sample will grow to ~ 100000 within next few years. Individual and detailed studies are expensive and not feasible for such large number.

In that, the self-similarity of structures becomes handy in providing a baseline model that link the cluster global properties to the cluster mass (Giodini et al., 2013). Below, is a description of the theoretical formalism for self-similar galaxy cluster observable relations to cluster mass.

Self-similar model

Kaiser (1986) showed that structures seeded by scale free initial conditions, and growing hierarchically under the sole action of gravity, follow self-similar relations between cluster mass and their global thermodynamic properties. According to the virial theorem, the total kinetic energy K and the total potential energy U of a virialised system are related as

$$2K = -U. \quad (2.35)$$

For a monatomic gas with isothermal temperature T , the average kinetic energy per particle is $\langle K \rangle = \frac{3}{2}k_B T$. Then, the total kinetic energy of the monatomic gas is $\frac{3}{2}Nk_B T$. It is possible to express the relationship of the kinetic energy to the total gas mass $M_{\text{gas},\Delta}$, i.e., $K \propto Nk_B T \propto M_{\text{gas},\Delta}k_B T$. The gravitational potential energy for a self-gravitating sphere of radius R_Δ and constant density ρ is given by $U \propto \int_0^{R_\Delta} \frac{4\pi}{3} \rho r^3 dr \propto \frac{M_\Delta^2}{R_\Delta}$.

Using the virial relation, the mass of a cluster and temperature of the gas medium are related as $T \propto M_\Delta/R_\Delta$, where $M_{\text{gas},\Delta} := f_{\text{gas}}M_\Delta$ and f_{gas} is the gas mass fraction. Expressing R_Δ in terms of mean density of the cluster such that $M_\Delta = \frac{4\pi}{3}\Delta\rho_c(z)R_\Delta^3$, and substituting for R_Δ in the mass and temperature proportionality gives the relation

$$T \propto M_\Delta^{2/3} \rho_c(z)^{1/3}. \quad (2.36)$$

Given that $\rho_c(z) = H_0^2 E(z)^2$, the above relation can be written out as

$$M_\Delta E(z) \propto T^{\frac{3}{2}}. \quad (2.37)$$

The X-ray bolometric luminosities from the bremsstrahlung emitting gas medium is $\propto \int n_e n_p T^{1/2} dV$. The densities of electrons and ions are proportional to the cluster density ρ due to self-similarity. Then

$$L_{\text{bol}} \propto \rho^2 R_\Delta^3 T^{1/2} \propto \rho M_\Delta E(z)^{1/3} M_\Delta^{1/3} \propto M_\Delta^{4/3} E(z)^{7/3}, \quad (2.38)$$

where the $M-T$ relation and the substitution of R_Δ with M_Δ is used.

The integrated Compton parameter Y_{SZ} , which is a product of density and temperature of electron in gas medium, is expected to have tight correlation with mass. $Y \propto \int n_e T dV$, which can be written as $Y \propto \rho R_\Delta^3 T \propto M_\Delta T$ and combining with Equation 2.37, we get

$$Y \propto M_\Delta^{\frac{5}{3}} E(z)^{2/3}. \quad (2.39)$$

Y_X is the X-ray analogue of the integrated Compton parameter proposed by A. V. Kravtsov, Vikhlinin and Nagai (2006) as a robust X-ray cluster observable that is least affected by the dynamical state of the ICM gas. It is constructed from the product $Y_X \equiv M_{\text{gas}} T$.

Empirical measurements of scaling relations

A review on scaling relations is given by Giodini et al. (2013). The above self-similar relations are theoretical expectations for gravity only cluster formation process. Deviations from the above relations are expected due to the diversity in the dynamical state of galaxy clusters and non-gravitational astrophysical process affecting the formation. The cluster observables such as L_X , Y_{SZ} are expected to scatter from the power-law relationship with mass due to the cluster-to-cluster differences.

Two mass-observable relations become very important for near future cluster surveys. The Sunyaev-Zel'dovich observables to mass for the SZ surveys and X-ray luminosities to mass for eROSITA. Many empirical measurements of scaling behaviour between the observables and mass has been studied using observations and, as well as, simulations (e.g., Giodini et al., 2013). X-ray luminosity-mass relations have been measured by multiple authors and they find that the scaling relation deviate from the self-similar relations in their slope, and a large intrinsic scatter of $\sim 40\%$ (Reiprich and Böhringer, 2002; Vikhlinin, A. V. Kravtsov et al., 2009; Giles, Maughan, Dahle et al., 2017; Reichert et al., 2011; Pratt et al., 2009; A. Mantz, Allen, Ebeling et al., 2010). X-ray luminosity is sensitive to the ICM density and variation in gas mass fraction within a cluster and with mass of the structures could lead to deviations from self-similar slope. Additionally, the cores of the ICM can be affected by energy injection processes such as AGN feedback and black hole formation. The presence of denser cores in some clusters leads to more efficient cooling and give rise to the cool-core clusters whose temperatures near the core is cooler than away from the core. Such cool-cores are thought to be responsible for the large scatter in X-ray luminosity relations (Pratt et al., 2009). Some authors excise the core for their luminosity measurements and re-produce self-similar scaling (e.g., Maughan et al., 2012).

A good mass proxy for cluster mass is one that is least affected by cluster-to-cluster differences. A. V. Kravtsov, Vikhlinin and Nagai (2006) proposed the proxy of thermal energy $Y_X := M_{\text{gas}} T$, which shows small scatter. $Y_X - M$ relation is measured to be following self-similar scaling (Vikhlinin, Burenin et al., 2009). Simulations predict the Y parameter from the SZ effect to scatter $\sim 6\text{--}12\%$ (Stanek, Rasia et al., 2010; Yang, Bhattacharya and Ricker, 2010; Angulo et al., 2012; Sembolini

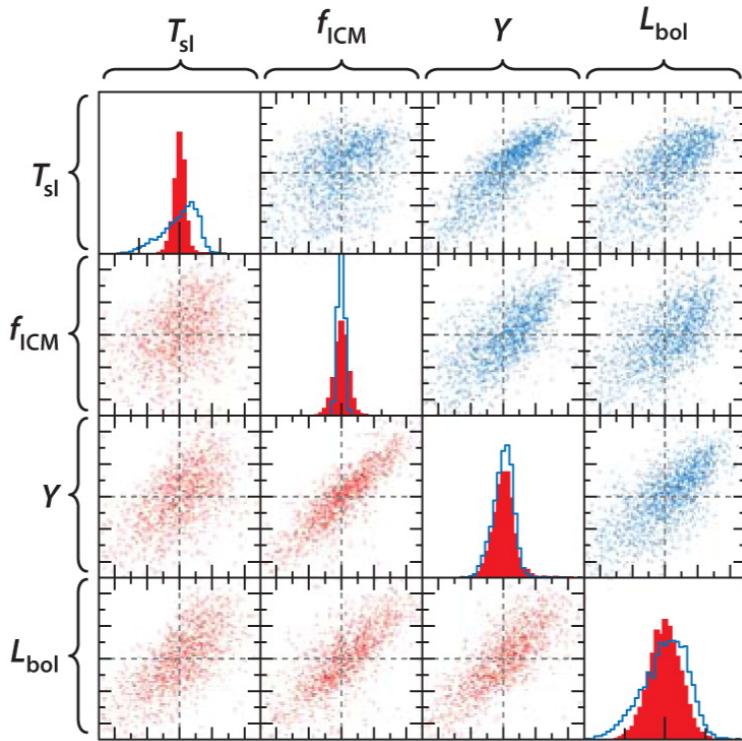


Figure 2.10: Covariance of cluster global properties at fixed mass from ~ 4500 halos in the Millennium Gas Simulations. Off-diagonal terms show the pair-wise correlation between two cluster properties at fixed mass. Two physical processes were considered. A gravity only process (blue) and a pre-heating scenario (red), where the gas was considered to have an entropy floor of 200 keVcm^2 at $z = 4$. The diagonals show the natural logarithmic deviations from the mean mass-scaling relation of each property. Credit: Allen, Evrard and A. B. Mantz (2011), originally adapted from Stanek, Rasia et al. (2010).

et al., 2013). Few authors have measured $Y-M$ scaling relation using X-ray masses (e.g., Andersson et al., 2011; Czakon et al., 2015; Bonamente et al., 2008) and others using lensing masses (Hoekstra, Mahdavi et al., 2012; Marrone et al., 2012; Hoekstra, Herbonnet et al., 2015; Sereno and Ettori, 2015; A. Mantz, Allen, Morris et al., 2016). Most of these work observe self-similar scaling slope with a couple of exceptions (A. Mantz, Allen, Morris et al., 2016; Czakon et al., 2015). In general, the intrinsic scatter found in the $Y-M$ relation is in the range 10–40%.

In the most general case, different cluster properties could be correlated for individual systems. It is becoming increasingly clear that for accurate understanding of the measured scaling behaviours, these intrinsic correlations are important for painting the complete picture (Allen, Evrard and A. B. Mantz, 2011; A. Mantz, Allen, Morris et al., 2016). But these correlations are challenging to measure from current data and, therefore, only few authors have attempted to include these correlations in their analysis of scaling behaviours (A. Mantz, Allen, Morris et al., 2016; A. Mantz, Allen, Ebeling et al., 2010). Due to challenges encountered in numerical simulations in producing detailed gas physics during cluster formation and evolution, only few authors have attempted to give predictions for these intrinsic correlations of cluster properties (Stanek, Rasia et al., 2010; Truong et al., 2016). A result from Stanek, Rasia et al. (2010) is shown in Figure 2.10, which measures the intrinsic correlations of four cluster properties, including the L_{bol} and Y_{SZ} . Their simulation used two simple physical

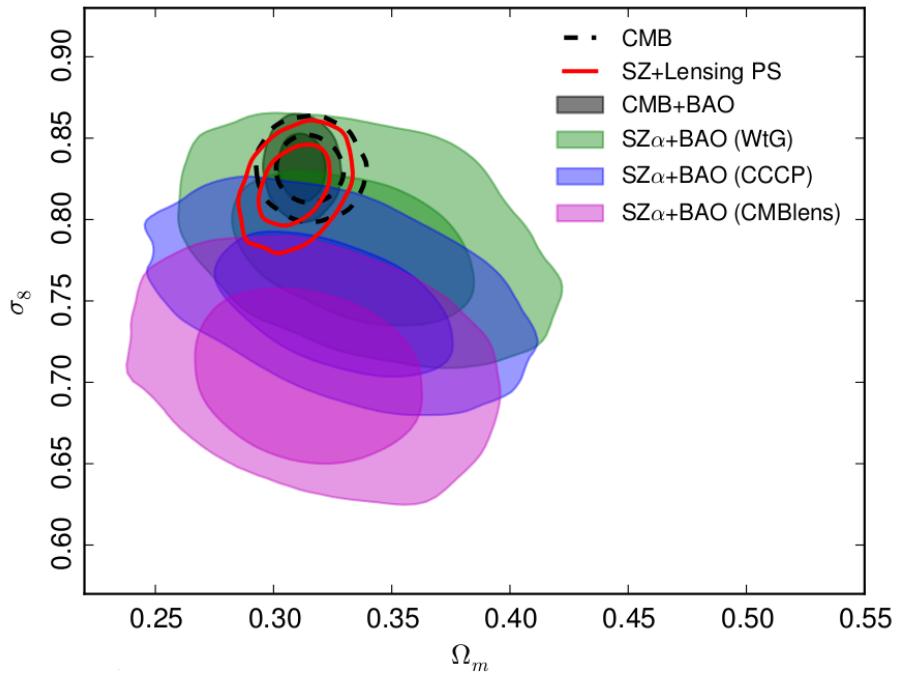


Figure 2.11: Systematic effects in cosmological constraints from galaxy clusters due to inaccuracies in the mass calibration of cluster observables. Green, blue and violet contours are 68% and 95% confidence regions of the constraints on the parameters obtained by using mass calibration from WtG, CCCP and CMB lensing estimates respectively. The systematic shift in the confidence levels from using the different mass calibrations demonstrate that the cluster abundance constraints on cosmology are currently limited by the uncertainties in the accuracy of mass-observable calibration. Credit: Planck Collaboration, Ade, Aghanim and al. (2016a).

process, a gravity only heating of gas and another scenario in which the gas medium was heated before collapse by an early energy injection at very early phases of cluster formation. Both these processes reveal that the thermodynamic cluster observables are strongly correlated at fixed mass in the range 0.5–0.8.

2.4 Current cosmological constraints from galaxy clusters and prospects for the future

Current constraints on cosmological parameters from cluster abundance measurements are limited by the systematic effects in the mass calibration of cluster observables.

This is demonstrated in Figure 2.11 taken from Planck Collaboration, Ade, Aghanim and al. (2016a). The figure shows constraints on σ_8 and Ω_m from *Planck* SZ selected galaxy clusters. Three different mass calibrations of the cluster observables provided by independent lensing mass measurements were used. Three independent measurements predict different mass scaling. The confidence levels on the cosmological parameters shift systematically based on which calibration is trusted. As can be seen from the same figure, there is some discordance between the CMB constraints on cosmology and the cluster constraints using the CMB lensing mass calibration. Whereas, adopting the calibration

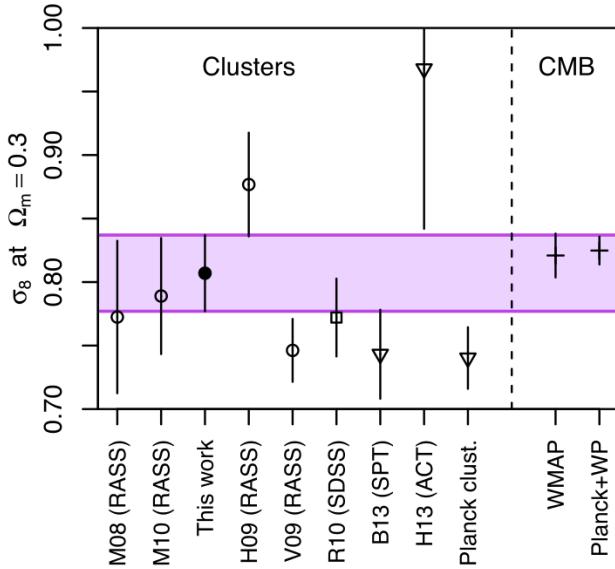


Figure 2.12: The constraint on σ_8 from different work using galaxy clusters are represented. The first five constraints are from X-ray selected clusters. The constraints from using an optically selected cluster catalog is represented as SDSS. SPT, ACT and Planck denote constraints from SZ selected samples. Credit: A. B. Mantz et al. (2015)

from WtG reconciles the galaxy cluster constraints with the CMB constraints. The discrepancy in the cosmological constraints from cluster abundance and CMB was first noted by Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al. (2014). Other possible ways to mitigate this discrepancy has also been tried (e.g., Salvati, Douspis and Aghanim, 2017) and this remains an open question.

The mass calibration limitation is not isolated to *Planck*. Other cluster cosmological analyses also rely on such priors on mass calibration of cluster observables (e.g., Benson, de Haan et al., 2013; de Haan et al., 2016). In Figure 2.12, a summary of some of the current constraints on σ_8 from cluster abundance from different work is shown to illustrate this. The figure illustrates constraints from X-ray selected (RASS), optical selected (SDSS) and SZ (SPT, Planck, ACT) samples. There is scatter in the accuracy of the measured value of σ_8 between these work. It can be observed that there is discordance among work using RASS or among SZ samples. The purple shaded region shows the constraints obtained from A. B. Mantz et al. (2015), which used the lensing estimates of cluster masses from WtG. Most of these work with three exceptions (Hasselfield et al., 2013; A. B. Mantz et al., 2015; Rozo et al., 2010) relied on X-ray masses or hydrostatic masses. For example, Benson, de Haan et al. (2013) used a prior on mass calibration from Vikhlinin, Burenin et al. (2009). In Figure 2.12, the constraints from Planck clusters are given using X-ray masses. It was already discussed in previous paragraph that the constraints from Planck clusters could prefer higher abundance if the WtG calibration is used instead.

This shows that for a reliable inference on cosmology with galaxy clusters, the accuracy of the mass calibration is extremely important. Therefore, mass scaling of galaxy cluster observables requires adequate handling of systematic effects in their absolute calibration.

In the next section, I present one of the key issues that tend to produce a biased absolute mass-observable scaling.

2.5 Scaling relations and sample selection biases

In the current and future cluster cosmology, scaling relations continue to play an important role. As already discussed in Section 2.4, it is currently one of the limiting factors for using clusters for precision cosmology. The accuracy of the mass calibration of cluster observables is vital for the experiments that are designed to detect neutrino masses and to constrain dark energy using cluster abundance and its evolution. Several factors have an impact on measuring the mass-observable scaling relations. For an absolute calibration, the mass proxies obtained from hydrostatic masses, dynamical masses, and/or weak-lensing masses have to be measured accurately with sufficient understanding of their systematic uncertainties. On the other hand, the measurement of scaling properties require a statistical framework as discussed in Allen, Evrard and A. B. Mantz (2011) due to the mass function, non-thermal physics affecting cluster observables, measurement uncertainties affecting cluster observable and mass estimations and sample selection effects.

Any sample of galaxy clusters constructed from detecting clusters in an observable is subject to selection biases. If a sample was constructed by a simple criteria, the biases in the sample and the sources of these bias can be characterized. To understand the kind of biases to expect in a typical observable limited sample, I review below some of the sample selection biases with the help of a toy model.

In Figure 2.13, a toy model of cluster population is used to demonstrate how sample selection biases a sample such as eDXL (introduced in Chapter 3). It must be noted that this applies to any sample that is selected with a threshold cut in an observable. For simplicity, the observable is chosen to be the X-ray luminosity of clusters and the measurement scatters shall be ignored. After assuming a power-law relation of the luminosities with cluster mass, their one-to-one relation is distorted by cluster-to-cluster differences. This is encapsulated by allowing for an intrinsic scatter at fixed mass in cluster observables that distort the distribution in the observable-mass plane. Such a distribution of sources using the Tinker mass function (Tinker et al., 2008), a scaling law and 40% intrinsic scatter in X-ray luminosity is shown in the upper panel of Figure 2.13. A selection threshold is chosen and applied to mimic the sample selection of a typical luminosity limited samples such as, LoCuSS (Okabe, G. P. Smith et al., 2013), REXCESS (Pratt et al., 2009), eDXL (introduced in Chapter 3). For the purpose of the demonstration exercise, the redshift is kept fixed at 0.3 for all sources. On applying the threshold and selecting 40 clusters (black plus points) form the cluster sample. This sample is biased towards low mass clusters that have been up-scattered from the scaling relation. Therefore, it can be said that the sample is biased towards low mass and intrinsically bright clusters. If a scaling relation is measured naively to the black data points, the measured relation would be biased towards a relation that has a shallower slope and higher normalisation. Such a bias in scaling of mass-observable, where the sample was selected on the observable, was identified by several authors and mitigation of such biases was recommended in measuring their scaling analysis (Pacaud, Pierre et al., 2007; Stanek, Evrard et al., 2006; Vikhlinin, Burenin et al., 2009; Pratt et al., 2009; A. Mantz, Allen, Ebeling et al., 2010).

Now, correlation of cluster observables on an individual cluster level as predicted from numerical simulations (e.g., Stanek, Rasia et al., 2010; Angulo et al., 2012; Truong et al., 2016) are considered.

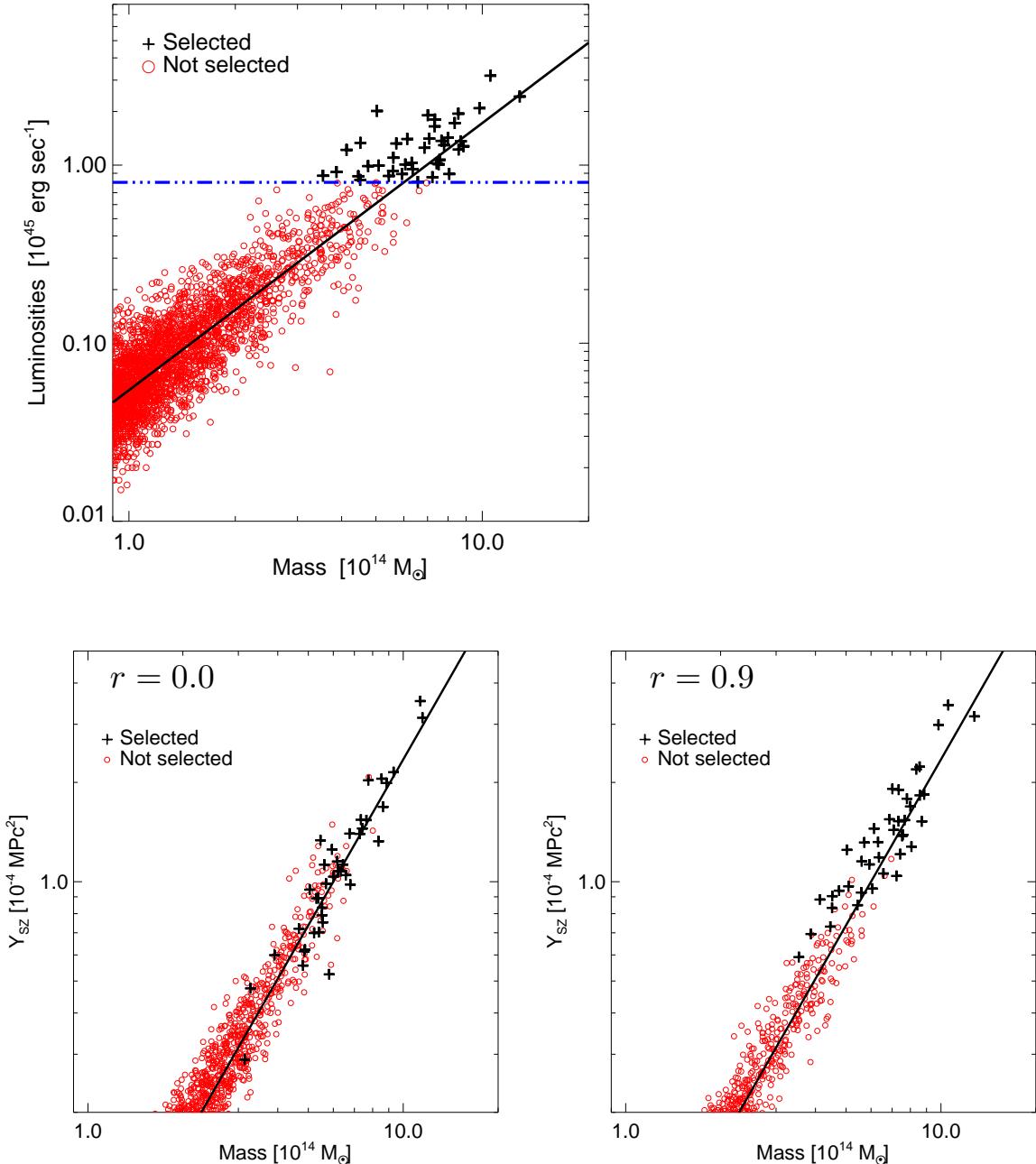


Figure 2.13: *Top:* A mock sample of clusters with X-ray luminosities vs. true mass generated using the Tinker mass function and a $L_x - M$ scaling relation (black solid line) are shown. As predicted by the mass function, more massive objects are rarer than low mass systems. The intrinsic scatter in X-ray luminosity mixes the populations of different mass bins. On applying a threshold cut in luminosity (blue dashed line), this selects the intrinsically bright sources at low mass end. Together, they produce a sample that is biased towards the low mass and intrinsically bright objects. *Bottom:* The sample generated above was used to also generate the Y_{SZ} for each source when (*left*) luminosity and Y_{SZ} are uncorrelated in their scatters at fixed mass, and (*right*) luminosity and Y_{SZ} are strongly correlated in their scatters at fixed mass.

For example, given that the SZE probes the pressure (product of electron density and temperature) of the gas medium, and the X-ray luminosity in the soft energy band is sensitive to the square of electron density, they might be correlated. These recent simulations predict a range of values between 0.5–0.9. For the toy model, this is considered to be 0.9 and cluster observable Y_{SZ} is also generated along with L_x using bivariate Gaussian distribution. Another set of generated sample was also considered with zero correlation between the intrinsic scatters. The two examples are shown in the lower panels of Figure 2.13. In the zero correlation case, there are clusters that can be found to be both up-scattered and down-scattered from the assumed scaling relation. Whereas, in the strongly correlated scatters case, at the lower mass end of the selected sample, there is a bias towards up-scattered population of clusters in the follow-up observables. This biases the follow-up observables in similar fashion as the original selection observable. This demonstrates that even when the selection is not on the SZE observable, a selection in X-ray luminosity and a significant correlation in these cluster observables at fixed mass can lead to a biased sample in SZE observable. In such an instance, ignoring the correlation would assume a relation for the SZE observable mass relation to be biased towards a higher normalisation and shallower slope scaling relation. This has been reviewed in Allen, Evrard and A. B. Mantz (2011), however, few authors have considered this effect in measuring mass scaling properties of an observable which was not directly selected on (e.g., A. Mantz, Allen, Ebeling et al., 2010; A. Mantz, Allen, Morris et al., 2016; Dietrich et al., 2017). In particular, while considering correlation in intrinsic scatter of Y and L_x , so far all authors have either completely ignored this effect (Marrone et al., 2012) or have considered only a strong prior on the correlation (Sereno, Covone et al., 2017; Czakon et al., 2015).

2.6 Summary

Within Λ CDM cosmological models, the existence and formation of massive structures like galaxy clusters is natural. The abundance of galaxy clusters in mass and redshift is an effective probe of the geometry of the Universe and the nature of dark energy. Their use as a cosmological tool requires an accurate and precise link of their observables to their mass. Scaling relations are the most effective way of linking cluster observables to mass for huge number of galaxy clusters. Current catalogues of galaxy clusters in X-rays and SZ used for measuring their abundance include less than 1000 clusters. Current constraints from these measurements have revealed the accuracy of mass-observable relation to be a limiting factor. Future experiments will increase the sample size to 100000 (Abazajian et al., 2016; Merloni et al., 2012). Cosmology with such huge number of clusters will improve measurements of the abundance more precisely, but will be more acutely limited by the accuracy of the mass calibration of observables. In section 2.5, a combination of sources of systematic effect in measuring scaling calibrations was identified. Few authors have considered all these sample selection effects together for measuring the scaling relations. In this work, the goal is to mitigate some of the issues identified in this chapter for our final calibration of the SZ effect to mass scaling.

CHAPTER 3

Cluster sample: APEX-SZ experiment and multi-wavelength follow-up

Overview

Note: Sections 3.2 to 3.4 of this chapter are re-produced with minor modifications from a recently submitted paper to the journal Monthly Notices of the Royal Academic Society titled “Weak-lensing mass calibration of the Sunyaev-Zel’dovich effect using APEX-SZ clusters” authored by A. Nagarajan, F. Pacaud, M. Sommer, M. Klein, K. Basu, F. Bertoldi, A. T. Lee, P. A. R. Ade, A. N. Bender, D. Ferrusca, N. W. Halverson, C. Horellou, B. R. Johnson, J. Kennedy, R. Kneissl, K. M. Menten, C.L. Reichardt, C. Tucker, B. Westbrook

The APEX-SZ experiment was a collaborative undertaking between institutes and universities in Europe (Max Planck Institute for Radioastronomy, Argelander Institute for Astronomy and Onsala Space Observatory) and North America (University of California at Berkeley, University of Colorado at Boulder, and McGill University at Montreal). The APEX-SZ instrument was a Stage-I CMB camera that performed targeted observations of ~45 galaxy clusters. This chapter introduces the relevant target sample of clusters, observations in multi-wavelength, and the relevant data analysis. The observations and data analysis in multi-wavelength were undertaken by several members of the APEX-SZ collaboration. A majority of this chapter discusses contributions from the APEX-SZ collaboration.

The optical follow-up and lensing analysis presented in Section 3.3, X-ray observable estimation from ROSAT discussed in Section 3.2.3, and the APEX-SZ data reduction described in Section 3.4 were primarily done by M. Klein, F. Pacaud, and M. Sommer respectively. This chapter forms the basis for rest of this thesis work.

3.1 APEX-SZ program

The APEX-SZ instrument (Dobbs et al., 2006; Schwan et al., 2011) was a milli-metre-wave transition-edge-sensor (TES) bolometer receiver designed to observe the Sunyaev-Zel’dovich (SZ) effect in the galaxy clusters. The instrument was used to make targeted observation of a sample of galaxy

clusters and map the Sunyaev-Zel'dovich effect produced by the hot gas medium residing in these galaxy clusters. The instrument was installed at the APEX telescope (Güsten et al., 2006) in Chile between 2005 and 2010, where the atmosphere is extremely dry enough to observe mm wavelength astrophysical signals. The goal of the project included studying the global properties of a sample of galaxy clusters, cross-calibrate multi-wavelength mass proxies for the targeted sample, and undertake detailed analysis of individual clusters like multi-wavelength analysis of radial profile of density and temperature of electrons in the intra-cluster medium.

3.2 APEX-SZ cluster sample

The APEX-SZ cluster targets were initially selected in an *ad hoc* manner, focusing on well-studied or seemingly interesting clusters with bright X-ray emission and hot ($T_e \gtrsim 5$ KeV) temperatures to ensure highly significant detections. To make a robust scaling relation analysis possible, later APEX-SZ observations were dedicated to follow-up a complete sample of 30 clusters, selected from the ROSAT All-Sky Survey (RASS) catalogues by applying well-defined cutoffs in the ROSAT luminosity-redshift plane. This sub-sample is essentially an extension of the REFLEX-DXL sample (Zhang et al. (2006)), and will, henceforth, be referred to as the extended Distant X-ray luminous galaxy clusters (eDXL) sample. In the following we describe the selection and characteristics of the eDXL sample. For completeness, a summary of the APEX-SZ clusters not belonging to the eDXL sample are given in Section 3.2.2.

3.2.1 The eDXL cluster sample

The sample was constructed as an extension of the volume complete DXL sample (Zhang et al., 2006), which consisted of the 13 clusters in the southern hemisphere with $0.27 < z < 0.31$ and ROSAT luminosities in the [0.1 – 2.4] keV band $L_X > 10^{45}$ erg/s. Taking advantage of the updated and deeper REFLEX-II catalogue enabled us to lower the luminosity cutoffs in the DXL redshift range to increase the mass coverage, and include some higher redshift clusters (up to $z = 0.54$). The precise luminosity cuts for each redshift range were set to maximize the overlap with earlier APEX-SZ observations, while staying above the nominal flux limit of the parent REFLEX-II catalogue (Böhringer, Chon, Collins et al., 2013).

The REFLEX II nominal flux limit, transposed onto the luminosity-redshift plane, is indicated in Figure 3.1. At this limit, the completeness of the parent sample is approximately 80% (as inferred from Figure 11 of the paper by Böhringer, Chon, Collins et al. (2013)). We also show in the same figure the location of the 90% completeness curve. Most of the clusters falls above this curve, ensuring a high completeness. As explained in more details in Section 6.5.2, our own scaling relation model permits to estimate a global completeness $\sim 90\%$ over our luminosity - redshift selection. In the low redshift range ($0.15 < z < 0.22$), all of the X-ray brightest clusters from REFLEX (Böhringer, Schuecker et al., 2004) and NORAS (Böhringer, Voges et al., 2000) catalogues are part of the APEX-SZ target list. This enables us to extend our sample selection to lower redshifts *a posteriori*, but requires the inclusion of NORAS to reach a meaningful number addition of five clusters. The high luminosity and redshift cuts were set to exclude other bright sources not observed with APEX-SZ. This luminosity cut is well above the nominal flux limit of both REFLEX and NORAS catalogues ensuring an effectively volume complete selection.

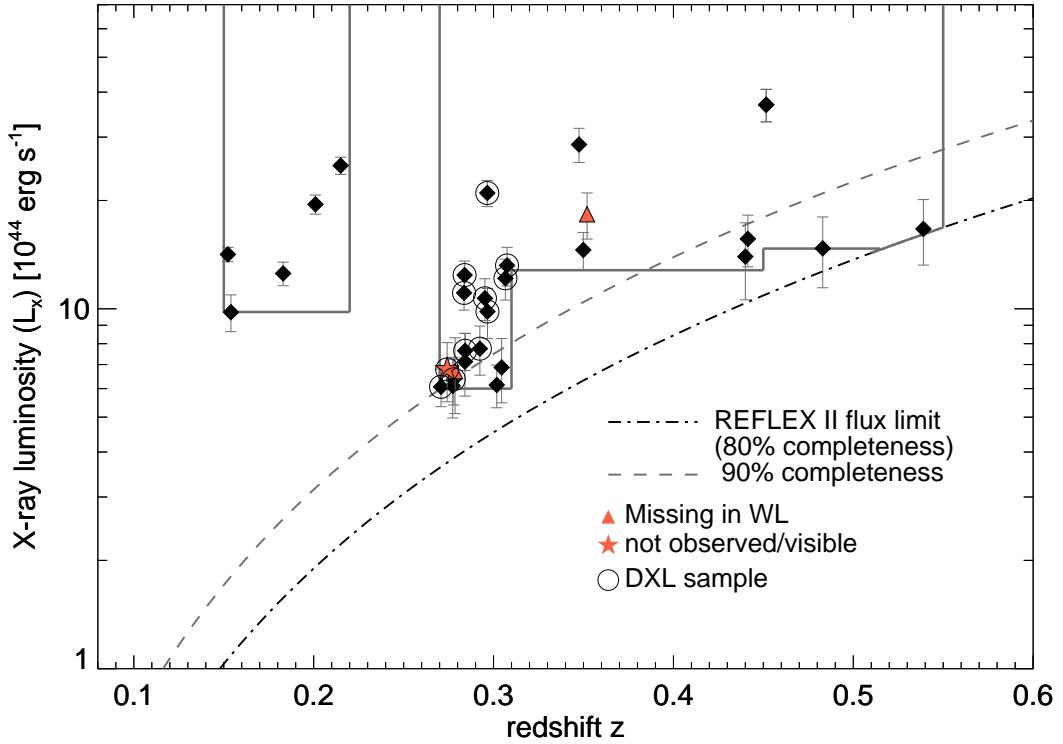


Figure 3.1: The extended DXL (eDXL) sample selection. The sample is selected in the luminosity-redshift plane from ROSAT catalogues. Above redshift equals 0.27, the sample was selected from REFLEX II catalogue. The low redshift sample is selected from NORAS and REFLEX. In total, 30 galaxy clusters are selected. The grey rectangular outlines enclose the sample selection plane. The details on optical follow-up observations of this sample are given in Section 3.3. As indicated in this figure, three galaxy clusters are missing in the follow-up program. The dotted-dashed curve represent the nominal flux limit (1.8×10^{-12} erg per sec) of the REFLEX II catalogue in the luminosity-redshift plane. This curve corresponds to approximately 80% completeness. The dashed curve is the luminosity curve for higher flux limit (3×10^{-12} erg per sec), which represent the 90% completeness of the REFLEX II catalogue. At low redshift, this curve corresponds to 80% completeness of the REFLEX catalogue.

Table 3.1: Luminosity cuts of the eDXL sample in different redshift bins. The selection is represented graphically in Figure 3.1. The luminosities quoted here are computed in the energy range [0.1 – 2.4] keV. In the final column, we give the number of clusters in each redshift range. In bracket, we mention the actual number of clusters that were completely followed-up in our multi-wavelength observations.

Redshift bin	Luminosity cut L^{\min} [10^{44} erg s $^{-1}$]	Parent sample	Number of clusters
$0.15 < z < 0.22$	9.78	REFLEX & NORAS	5 (5)
$0.27 < z < 0.31$	6.0	REFLEX II	17 (15)
$0.31 < z < 0.45$	12.8	REFLEX II	5 (4)
$0.45 < z < 0.55$	14.7	REFLEX II	3 (3)

With 30 galaxy clusters in total, the extended DXL selection more than doubles the number of clusters from the initial DXL sample. It was designed to provide both a good leverage on the slope of scaling relations at $z \sim 0.3$ and a large redshift coverage. This should permit breaking the degeneracy between the inferred slope and redshift evolution of scaling relations. The exact, redshift-dependent, luminosity thresholds used for the selection are given in Table 3.1. A graphical representation of the corresponding parameter space and the selected clusters is provided in Figure 3.1.

XMM observations are available for all 30 galaxy clusters in the eDXL sample. However, one of them could not be observed from the APEX site due to its too low declination. For two others, the lensing data were not of sufficient quality to provide any mass information due to bad weather conditions and poor seeing.

In this work, only those 27 cluster with complete follow-up data are included in the complete eDXL sample. Since the exclusion of the two clusters in this down-selection was random, i.e. does not depend on the cluster physical properties, we assume that the selection function of the sample remains unaffected.

3.2.2 Other APEX-SZ clusters

The complete APEX-SZ sample does not have a well-defined selection. In addition to the eDXL sub-sample, it contains a number of high redshift clusters and a few massive local clusters whose inclusion in our complete selection would have required the observation of many more targets to reach a complete sample. In total, 12 additional APEX-SZ clusters have complete multi-wavelength follow-up in X-rays (either by the *XMM-Newton* or the *Chandra* satellite) and were followed up with optical observations. The latter follow-up is summarized in section 3.3. For completeness, we provide the global observable measurements of these 12 additional clusters along with our eDXL clusters.

3.2.3 X-ray observables and parameter estimation

Note: The X-ray parameter estimation from ROSAT survey data was performed by Dr. Florian Pacaud for the purpose of homogeneously measuring the X-ray luminosities for all the clusters and these are used in Chapter 6. I present a summary of the analysis here as it has relevance to this work.

The procedure to consistently recompute the ROSAT X-ray luminosities for all the eDXL clusters derives from the REFLEX-II recipes described in Böhringer, Chon, Collins et al. (2013). The measurements rely on ROSAT PSPC photon and exposure maps in the [0.5 – 2] keV, where the signal-to-noise ratio is highest. However, the final luminosities quoted and used in this work correspond to the full [0.1 – 2.4] keV band, as is customary for ROSAT sources. The conversion between the two bands make use of the redshift and temperature dependent K-correction tables provided by Böhringer, Schuecker et al. (2004) which show little variation over a wide temperature range.

The process can be split into the following main steps:

1. The X-ray centroid for each cluster in the sample was calculated from the ROSAT photon map within a 3' aperture, iteratively updating the centre of the aperture until convergence.
2. The local background for each cluster was computed inside an annulus covering the radial range 20'-41.3'. To account for the possible contamination by surrounding AGNs, this annulus was split into 12 sectors azimuthally. The background count-rate in each sector was estimated and contaminated areas were rejected using an iterative 2.3σ clipping. The mean background was finally computed from the remaining sectors. Such a procedure is justified by the low AGN density in the ROSAT maps.
3. A growth curve analysis as prescribed in Böhringer, Chon, Collins et al. (2013) was used to estimate the integrated net aperture count-rate of the source in a suitable radius. The integration radius, R_x , is first defined as the radius above which all changes in the integrated flux stay within the 1-sigma error range at that radius. The corresponding integrated source count-rate, CR, is then estimated by fitting a straight line to the plateau at larger radii, as shown in Figure 3.2.
4. Finally, we estimated the value of $L_{x,500}$ in the [0.1-2.4] keV band corresponding to the measured $CR(< R_x)$. For this, we first use the $L_{x,500} - T_x$ relation of Pratt et al. (2009),

$$\frac{T_x}{1 \text{ keV}} = 3.31 \left(\frac{L_{x,500}}{10^{44} \text{ erg s}^{-1}} \right)^{0.332} h_{70}^{0.666}, \quad (3.1)$$

to estimate the temperature dependent K-correction suitable for any given $L_{x,500}$, and convert it to the [0.5 – 2] keV ROSAT count-rate in R_{500} , CR_{500} . $L_{x,500}$ is the X-ray luminosity within R_{500} . T_x is the X-ray temperature. Following the results of Reichert et al. (2011), we assumed the redshift dependence of the $L_{x,500} - T_x$ relation to be negligible. Then, we use the Reichert et al. (2011) Mass-Luminosity relation expressed as

$$\frac{R_{500}}{1 \text{ Mpc}} = 0.957 \left(\frac{L_x}{10^{44} \text{ erg s}^{-1}} \right)^{0.207} E(z)^{-1} h_{70}^{0.586}, \quad (3.2)$$

to estimate the radius R_{500} within which CR_{500} should be measured. Lastly, a fixed beta-model

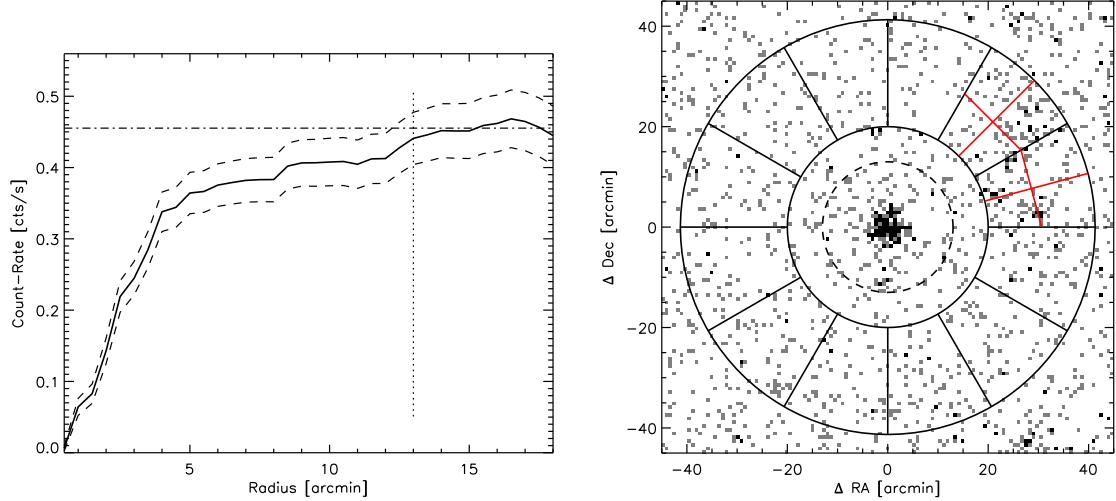


Figure 3.2: *Left:* Result of the growth curve analysis for the Bullet cluster. The net aperture count-rate of the cluster is plotted against the radial distance from the centre (solid line) together with its 1σ uncertainty (dashed lines). The integrated count-rate flattens, i.e. shows fluctuations lower than the 1σ error range, after the radius R_x , indicated by the vertical dotted line. A constant count-rate (CR), indicated by the horizontal dot-dashed line, is fitted to this plateau region and serves as the main source photometry indicator. *Right:* ROSAT image of the cluster is shown with the R_x indicated by dashed circle. The outer circles enclosing the region for background estimation are indicated by solid circles. The excluded sectors from the background estimation is marked in red. *Image credit:* F. Pacaud.

was assumed with $\beta = 2/3$ and $R_c = R_{500}/7$ to estimate the extrapolation factor from CR_{500} to $CR(< R_x)$. The full conversion process is performed for a grid of $L_{x,500}$ and the correct value is obtained after interpolation over the estimated CR.

The above procedure provides a complementary and independent information on cluster centroids for the baryonic component emitting X-rays. The X-ray luminosities measurements obtained from the above procedure are given in Table 3.2.

3.3 Optical follow-up observations and lensing masses

Note: The optical follow-up and the lensing estimation provided here was done by Dr. Matthias Klein and an earlier stage of this work was presented in his Ph.D. thesis (Klein, 2014). The summary here is an updated analysis that is provided due to relevance of these measurements in the final scaling relation study performed in Chapter 6.

A complete follow-up measurements of weak-lensing masses for the APEX-SZ sample was achieved by a combination of using the archival observation data and dedicated follow-up observations of clusters that were previously not observed with a desired quality for lensing analysis. The follow-up observations were carried out between January 2010 and February 2012 using the Wide Field Imager (WFI) at the 2.2 MPG/ESO telescope at La Silla, Chile.

The observations were done in the B, V and R bands, with exposure times depending on the cluster

redshift, reaching 12, 4.5 and 15 kilo-seconds per band, respectively, at $z=0.3$. A total of 21 clusters have quality weak-lensing data with WFI. 16 clusters had Suprime-Cam data from the Subaru telescope with imaging in at least three bands and sufficient quality for a weak-lensing analysis. For an additional 6 clusters a combination of WFI and Suprime Cam data was used with at least one photometric band supplied by the other instrument. For three clusters, sufficient amounts of data were available to perform independent weak-lensing analysis using both instruments separately.

For all clusters three band-photometry for background selection were used and a sub-arcsecond seeing for the shape measurement band was required. All colours were matched to the nearest colours available in COSMOS photo-z catalogues (Ilbert et al., 2009), which were used as reference catalogues for background galaxy selection for all targets.

For the optical follow-up data described above, the Schrabback, Erben et al. (2007) implementation of the KSB+ algorithm (Kaiser, Squires and Broadhurst (1995) and Erben et al. (2001)) to measure the shapes of individual galaxies was used. Distortions of the point spread function (PSF) could be well modeled and corrected for, using polynomials of orders up to five.

The image distortion in the weak-lensing limit by a radially symmetric lens at an angular diameter distance D_d from the observer can be measured as an average tangential ellipticity about the lens centre. The average tangential ellipticity ε_t of source images is a direct measurement of the reduced shear $\langle g \rangle = \langle \varepsilon_t \rangle$. The reduced tangential shear is related to the shear, γ , and convergence or surface mass density, κ as

$$g(\theta, \beta) = \frac{\gamma(\theta, \beta)}{1 - \kappa(\theta, \beta)}, \quad (3.3)$$

where θ is the angular projected radial distance from lens centre and β is a scale factor for the strength of the lensing effect. It is defined as the angular distance ratios such that $\beta = \frac{D_{ds}}{D_s}$, where D_s, D_{ds} are the angular diameter distances between observer and source and between deflector and source. For redshifts lower than or equal to the cluster redshift, β is equal to zero by construction. For higher redshift sources, $\beta(z_s)$ is a strictly monotonously increasing function of the source redshift z_s . As such, it can be used as a distance measurement compared to redshift and to exclude low distance sources not carrying lensing information.

For a reliable weak-lensing analysis, the selection of background lensed sources and reducing the contamination of such by foreground sources is extremely important. Here, a novel approach for getting as a pure background selection as possible with three colour information was used.

The estimated β_i for each galaxy i in the observation is the weighted mean of β_k of all sources k in the COSMOS photo-z catalogue (Ilbert et al., 2009) within a region in colour-colour-magnitude space defined by the size of the photometric errors in colour and magnitude,

$$\beta_i = \frac{\sum_{k=1}^N w_k(\Delta c_1, \Delta c_2) \beta_k}{\sum_{k=1}^N w_k(\Delta c_1, \Delta c_2)}. \quad (3.4)$$

Each reference source (k) is weighted by a two dimensional Gaussian function, $w_k(\Delta c_1, \Delta c_2)$, where $\Delta c_1, \Delta c_2$ are the distance coordinates in colour-colour region of reference galaxy k , from the observed colour of the source galaxy, i . The dispersion of the Gaussian function is given by the actual measurement uncertainty on the observed colour of the galaxy, i . Due to a limited precision of the

Table 3.2: The 27 targets of the eDXL sample. Each cluster position is taken to be the position of the BCG in the optical data. The weak-lensing derived masses ($M_{\text{WL},500}$), and the re-measured ROSAT luminosities (L_x) are reported.

Name	RA	Dec	redshift z	$M_{\text{WL},500}$ [$10^{14} M_\odot$]	$L_x[0.1 - 2.4 \text{ keV}]$ [$10^{44} \text{ erg s}^{-1}$]
eDXL clusters					
A2204	16:32:46.9	+05:34:32.3	0.152	$6.53^{+1.51}_{-1.67}$	14.2 ± 0.6
RXCJ2014.8 – 2430	20:14:51.7	-24:30:22.3	0.160	$5.37^{+2.22}_{-2.60}$	9.8 ± 1.1
A1689	13:11:29.5	-01:20:27.9	0.183	$20.56^{+1.51}_{-1.59}$	12.5 ± 0.9
A2163	16:15:49.0	-06:08:41.5	0.203	$12.78^{+2.72}_{-3.17}$	19.5 ± 1.2
RXJ1504	15:04:07.5	-02:48:16.5	0.215	$5.25^{+1.25}_{-1.38}$	25.0 ± 1.4
RXCJ0532.9 – 3701	05:32:55.7	-37:01:36.0	0.275	$6.76^{+1.33}_{-1.53}$	6.1 ± 0.7
RXCJ0019.0 – 2026	00:19:08.0	-20:26:28.0	0.277	$7.70^{+1.45}_{-1.53}$	6.1 ± 1.1
RXCJ2337.6 + 0016	23:37:39.7	+00:16:17.2	0.278	$7.08^{+1.37}_{-1.32}$	6.4 ± 1.0
RXCJ0232.2 – 4420	02:32:18.6	-44:20:48.0	0.284	$5.13^{+1.69}_{-1.94}$	11.1 ± 1.2
RXCJ0437.1 + 0043	04:37:09.5	+00:43:52.1	0.284	$8.10^{+2.03}_{-2.15}$	7.6 ± 0.9
RXCJ0528.9 – 3927	05:28:53.0	-39:28:17.8	0.284	$4.38^{+1.26}_{-1.36}$	12.4 ± 1.2
RXCJ2151.0 – 0736	21:51:00.8	-07:36:31.0	0.284	$2.22^{+1.10}_{-1.31}$	7.1 ± 1.4
A2813	00:43:25.1	-20:37:01.2	0.292	$8.30^{+1.54}_{-1.75}$	7.8 ± 1.2
RXCJ0516.6 – 5430	05:16:37.6	-54:30:38.1	0.295	$8.42^{+1.99}_{-2.06}$	10.7 ± 1.4
Bullet	06:58:36.4	-55:57:19.2	0.297	$6.30^{+1.71}_{-2.09}$	21.0 ± 1.7
A2537	23:08:22.2	-02:11:31.6	0.297	$10.46^{+1.80}_{-1.87}$	9.8 ± 1.6
RXCJ0245.4 – 5302	02:45:31.3	-53:02:07.8	0.302	$3.47^{+1.24}_{-1.45}$	6.1 ± 0.8
RXCJ1135.6 – 2019	11:35:21.4	-20:19:56.6	0.305	$4.47^{+1.22}_{-1.29}$	6.9 ± 1.4
A2744	00:14:18.5	-30:22:51.2	0.307	$11.55^{+1.79}_{-2.00}$	12.2 ± 1.6
A1300	11:31:54.2	-19:55:39.8	0.308	$5.82^{+1.62}_{-1.76}$	13.2 ± 1.6
MACSJ1115.8 + 0129	11:15:52.0	+01:29:55.0	0.348	$5.43^{+1.65}_{-1.95}$	14.6 ± 1.7
RXCJ2248.7 – 4431	22:48:44.0	-44:31:51.0	0.348	$9.84^{+2.01}_{-2.16}$	28.6 ± 3.1
RXCJ1206.2 – 0848	12:06:12.1	-08:48:03.4	0.441	$8.67^{+2.68}_{-2.87}$	15.6 ± 2.6
RXCJ2243.3 – 0935	22:43:22.8	-09:35:22.0	0.447	$10.29^{+2.52}_{-2.84}$	14.0 ± 3.4
RXJ1347 – 1145	13:47:30.6	-11:45:09.5	0.451	$14.00^{+2.96}_{-3.02}$	36.9 ± 3.8
RXCJ2214.9 – 1359	22:14:57.2	-14:00:12.3	0.483	$8.39^{+2.45}_{-2.88}$	14.7 ± 3.3
MS0451.6 – 0305	04:54:10.8	-03:00:51.4	0.539	$6.08^{+1.90}_{-2.12}$	16.7 ± 3.4

Table 3.3: Same as table 3.2, but for other clusters in the full APEX-SZ sample that are not part of the X-ray selected complete sample.

Name	RA	Dec	redshift z	$M_{WL,500}$ $[10^{14}M_\odot]$
Other clusters				
A907	09:58:22.0	-11:03:50.2	0.153	$3.38^{+0.85}_{-1.02}$
A3404	06:45:29.5	-54:13:37.1	0.167	$8.40^{+1.91}_{-2.26}$
A383	02:48:03.4	-03:31:45.1	0.187	$5.86^{+1.17}_{-1.13}$
A520	04:54:13.7	+02:56:10.2	0.199	$5.45^{+0.96}_{-0.83}$
A209	01:31:52.5	-13:36:40.7	0.206	$9.08^{+1.12}_{-1.22}$
A2390	21:53:36.8	+17:41:43.7	0.228	$9.40^{+1.28}_{-1.41}$
A1835	14:01:02.1	+02:52:42.6	0.253	$13.74^{+2.45}_{-2.58}$
RXCJ1023.6 + 0411	10:23:39.2	+04:10:58.0	0.280	$8.28^{+1.20}_{-1.19}$
XLSSC-006	02:21:45.2	-03:46:02.7	0.429	$4.97^{+1.28}_{-1.33}$
MACSJ1359.2 – 1929	13:59:10.3	-19:29:24.7	0.447	$4.19^{+1.84}_{-2.16}$
MACSJ1311.0 – 0311	13:11:01.8	-03:10:39.7	0.494	$6.96^{+1.86}_{-1.97}$
MS1054.4 – 0321	10:56:60.0	-03:37:36.2	0.831	$14.20^{+10.06}_{-14.57}$

estimated β_i , a cut $\beta_{\text{cut}} > 0$ was applied to exclude cluster members and foreground galaxies. The first step is finding the $\beta_{\text{cut,max}}$ that maximizes the signal to noise of the lensing signal. Klein et al. (in preparation) show that this cut results in a bias of $\sim 1\%$ in R_{200} due to noise fluctuations. This bias is avoided in the final mass analysis by increasing the applied cut, $\beta_{\text{cut,fin}} = \beta(z_{\text{cut,max}} + 0.05)$, the value that is obtained for a redshift 0.05 higher than that of $\beta_{\text{cut,max}}$.

The mass estimate for each cluster was obtained by fitting a reduced tangential shear profile predicted by a projected Navarro-Frenk-White (NFW) profile (e.g., Bartelmann, 1996) to the observed ellipticities. They derive the best fitting profile parameters R_{200} and c_{200} by minimizing the merit function

$$\chi^2 = \sum_{i=1}^N \frac{|g_i(\theta_i, \beta_i; R_{200}, c_{\text{NFW}}) - \tilde{\varepsilon}_{t,i}(\theta_i)|^2}{\tilde{\sigma}_i^2 (1 - |g_i(\theta_i, \beta_i; R_{200}, c_{\text{NFW}})|^2)^2} . \quad (3.5)$$

Here $g_i(\theta_i, \Sigma_{\text{crit},i}; R_{200}, c_{\text{NFW}})$ is the model prediction for galaxy i and $\tilde{\varepsilon}_{t,i}$ the observed ellipticity times 1.08 for the same galaxy. The factor 1.08 is the multiplicative shear calibration bias of the used KSB+ pipeline (Kaiser, Squires and Broadhurst (1995) and Erben et al. (2001)) to convert from measured to true ellipticity. This calibration bias has an uncertainty of $\sim 5\%$. This uncertainty is a dominant source of systematic uncertainty in the mass measurements. Each shear profile was centred on the BCG, using distances in the range of 0.2 to 4.2 Mpc for the fitting procedure. They minimised the χ^2 on a grid of R_{200} and c_{200} . Finally, the mass-concentration relation described by Bhattacharya et al. (2013) was used to put priors on the concentration parameter to break the degeneracies in the profile models. The initial mass estimates from Equation (3.5) are biased. In evaluating the NFW shear profile the ratio in Equation (3.3) was used when averaging the value of β over the reference

catalogue sources. However, $\frac{\gamma(\langle\beta\rangle)}{1-\kappa(\langle\beta\rangle)} \neq \left\langle \frac{\gamma(\beta)}{1-\kappa(\beta)} \right\rangle$. Given the finite width of the β distribution that are averaged over when calculating β_i from a reference catalogue (Equation 3.4), the point estimate for β_i would be a biased estimator. Especially in the inner regions of the cluster, this would model the shear profile incorrectly. The estimate of the final masses was obtained by correcting for the averaging over β in two subsequent iterations. The best-fit mass estimate from the zeroth iteration was used to predict the reduced shear, g , at the projected distance θ from the cluster centre, and β_k . Then β'_i was introduced, which satisfied the equation:

$$g(\beta'_i) = \frac{\sum_{k=1}^N w_k g(\theta_i, \beta_k)}{\sum_{k=1}^N w_k} \frac{1}{v_b(c_1, c_2)}. \quad (3.6)$$

Here, the first term is the weighted average of the reduced shear given the projected distance θ_i of galaxy i to the cluster centre and angular diameter distance ratios β_k of references sources. The weights w_k are identical to those used to derive β_i and solely depend on the distance between reference and observed source in colour-colour space.

The second term in equation (3.6) contains the map $v_b(c_1, c_2)$, an estimator of the overdensity of galaxies in colour-colour space with respect to a background estimate. This term addresses the different redshift distributions in the reference and cluster fields, assuming that they are caused by the addition of cluster galaxies. We divide the cluster field into annuli. The background annulus is chosen to be beyond R_{200} (using the R_{200} estimate from the first iteration, equation 3.5). The region inward of R_{200} is split into three overlapping annuli of width $0.3 - 0.5R_{200}$. For each galaxy i , we compute $v_b(c_1, c_2)$ for one specific bin b (depending on its angular separation θ_i from the cluster centre), with respect to the background annulus.

Under the assumption that the redshift distribution in the outskirts of the observed fields is close to the reference distribution, the density ratio maps v_b in colour-colour space reflect the difference of the two distributions at a given position in colour-colour space. Ignoring the impact of lensing magnification, the cluster always causes an excess of galaxies compared to the average distribution. To avoid correcting to insignificant noise fluctuations, $v_b(c_1, c_2)$ is set to 1 for all colour-colour regions with an excess smaller than two sigma above the mean value. Visual inspection of the v_b images was performed to ensure that overdensities caused by additional clusters in the observed fields with redshifts higher than those of the targeted clusters are not considered in the correction.

Equation (3.6) describes the expected reduced shear at position θ_i given the expected redshift distribution of reference sources and the expected contamination caused by cluster galaxies given the colours of the observed galaxy i . As such β'_i is a less biased estimator than β_i . The χ^2 was re-computed in a grid of c_{200} and R_{200} with the updated β' . They re-iterated this once more and computed final mass estimates by minimising Equation (3.5) and applying a procedure identical to the first iterations of mass estimates. In Tables 3.2 and 3.3, the final spherical masses within R_{500} defined in Equation (2.1.2) and the cluster profile centres adopted for the lensing measurements are given.

We compare the R_{500} obtained from the lensing analysis described above to the R_{500} determined from ROSAT data in Section 3.2.3 for the eDXL sample. The luminosities measured within this X-ray R_{500} is shown in Table 3.2. On average, the ratio of lensing aperture to the X-ray derived aperture is 0.89 ± 0.03 , where the uncertainty is the standard error for the sample of 27 clusters. This indicates that the X-ray aperture is significantly larger than the lensing estimate on average for this sample. This is a bit surprising, however, the lack of consistency in these two measurements are not important

for our final scaling relation work in Chapter 6. The reason being that the X-ray luminosities are essential for only determining the sample selection biases. The luminosities measured in the manner described in Section 3.2.3 were performed in a manner similar to how the parent populations were selected. This makes sure that we are able to get an accurate modelling of the selection function for this sub-sample of 27 clusters.

3.4 APEX-SZ data analysis performed with bolometer analyser (BoA)

Note: This work was led by Dr. Martin Sommer and I have minor contributions to this pipeline. Most significant contribution made to this pipeline by me was the implementation of a the polynomial+secant function fitting to time stream data (encountered in Section 3.4.2) in python for the BoA pipeline. Some other minor contributions include aiding in the determination of the optical depth of the atmosphere from Skydip data.

APEX-SZ had a resolution of one arcmin and was used to observe 47 known massive galaxy clusters, with a total observation time of over 800 hours.

The APEX-SZ data were flagged and filtered using the bolometer array data analysis software BoA¹. A series of linear filtering steps was carried out on the time-stream data of each target, using universal settings to ensure a uniform analysis. We begin this Section with summaries of the calibration and time-stream filtering steps, and proceed to discuss our analysis in terms of the *point source transfer function* (described in section 3.4.3), constructed to take into account both the filtering steps and the instrument beam when modelling the sky signal.

3.4.1 Calibration

The beam position and shape of each bolometer flux in the focal plane were measured from daily scans of a calibrator (Mars, Uranus, Saturn or Neptune). Side lobes were characterized by combining the individual detector beams into a composite beam. Absolute flux calibration was performed based on the response of each detector using scans of Mars and Uranus. Depending on visibility of the primary calibrators, bootstrapped observations of secondary calibrators were also used. To account for differences in atmospheric opacity between the data and calibration scans, a correction was applied based on radiometer readings. A further correction was applied to account for gain fluctuations due to bolometers being biased near the upper edge of the superconducting transition. The total calibration uncertainty for APEX-SZ is $\pm 10\%$. The details of all these steps were discussed by Bender et al. (2016), and are thus only summarized here.

3.4.2 Time stream processing

The time stream processing of the APEX-SZ data is similar but not identical to that performed by Bender et al. (2016). Thus, we give a relatively detailed account of this process here. The observations with APEX-SZ were carried out using circular drift scan patterns centred on a constant horizontal

¹ <http://www.apex-telescope.org/bolometer/laboca/boa/>

coordinate, allowing the target to drift through the pattern and the FoV. Circle radii were chosen such as to maximize the integrated signal-to-noise ratio of each target, based upon considerations of filtering effects (see Section 3.4.3). The details of the APEX-SZ drift scan pattern were discussed by Bender et al. (2016). As a first step, the data were parsed into separate, full circles on the sky, and re-grouped based on a common centre in horizontal coordinates, resulting in what we shall call subscans. Data not belonging to circle sets were discarded. Optically unresponsive channels (bolometers) were rejected. Spikes were cut using sigma clipping, and jumps (in DC level) were identified and corrected for using a wavelet-based algorithm. An additional data cut was performed by analysing the correlation between channels; channels found to correlate poorly with their neighbours were rejected along with channels exhibiting levels of noise significantly higher than the median noise level. Typically, 140-170 live channels were used for further analysis. After these initial steps, an optical time constant (time delay in bolometer response) was de-convolved from each channel, using the approach of Bender et al. (2016).

The polysecant (a polynomial + secant model) fitting employed by Bender et al. (2016) was also used here. To the time stream of each channel and subscan, we fit a 6th order polynomial plus a normalization of the expected variation of signal along a circle due to air mass load, and subtracted this baseline from the data. Following this step, we removed a signal correlated across all channels, constructed by taking the mean signal adjusted for individual channel normalisations. Finally, a polynomial of order 3 was fit to each set of two circles on the sky, before the data were again de-spiked using sigma clipping.

3.4.3 Point source transfer function

APEX-SZ observations were generally carried out at relatively high (for the site) levels of precipitable water vapour due to significant amount of observation run concurring with the *Bolivian Winter*. For this reason, the APEX-SZ data suffer from excess low-frequency noise correlated on scales much smaller than the FoV, requiring high-pass filtering of individual bolometer time streams to be applied after removing the correlated atmospheric signal. While this step enhances the signal-to-noise ratio of detections, it also significantly attenuates astrophysical signals. To account for this, a point source transfer function (as described by Halverson et al., 2009; M. Nord et al., 2009) is used to model the systematic signal loss. The point source transfer function is unique for each target. It is constructed from a noiseless simulation of a perfect point source at the position of the target, convolved with the instrument point-spread function, de-gridded to the bolometer time streams and processed in parallel with the data, applying identical filtering to both the data and the simulation. After filtering, the point source transfer function represents the impulse response of the filter, and can be used, under the assumptions of directional independence and linearity of the filter, to model the response of any model that one may wish to compare to the data.

Images of the data and the transfer functions were made using the methods outlined by Bender et al. (2016). For each target 100 noise images were also created by randomly inverting half the data (randomly chosen subscans), to characterize instrument noise properties.

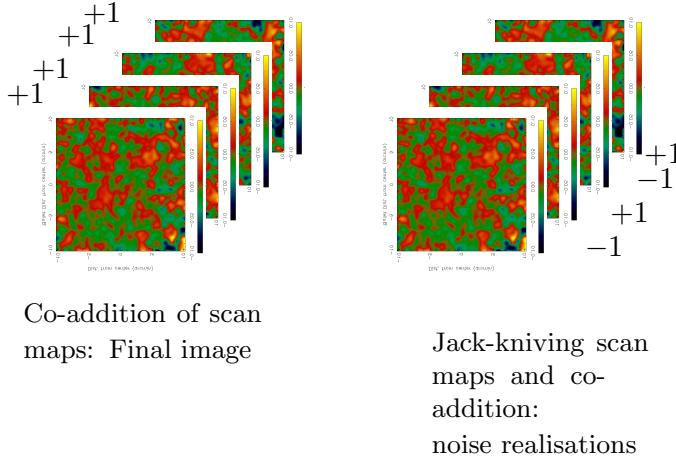


Figure 3.3: The final co-added map is produced by average summing all of the un-flagged individual scan maps. The final co-addition should down-weight the noise and boost the overall signal to noise of the astrophysical signal. *Right:* The same set of scan maps are randomly split into two halves. One half of scan maps are multiplied by +1 and other half by -1. These scans are then added together with the assigned sign to effectively get rid of astrophysical sources and produce maps that contains the noise information. Typically, 100 realisations of noise images were produced in this manner. This is discussed in Section 3.5.2

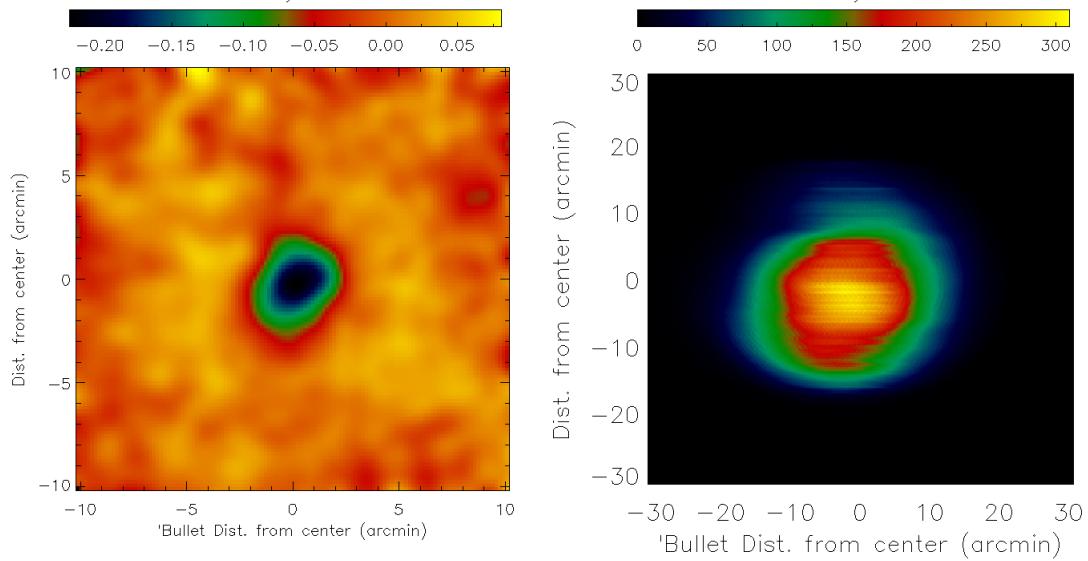


Figure 3.4: *Left:* The final APEX-SZ coadded image of Bullet cluster is shown here. The map has been smoothed with a Gaussian kernel with σ equals half arcmin. This is filtered image, the peak temperature decrement seen here is attenuated due to the filtering of the signals for noise. *Right:* (Pixel weights) The map pixel weights for the Bullet cluster is shown here. The weights are inversely proportional to the square of the noise RMS at the pixel positions. The weights fall off radially from the center of the pointing due to the circular scan strategy.

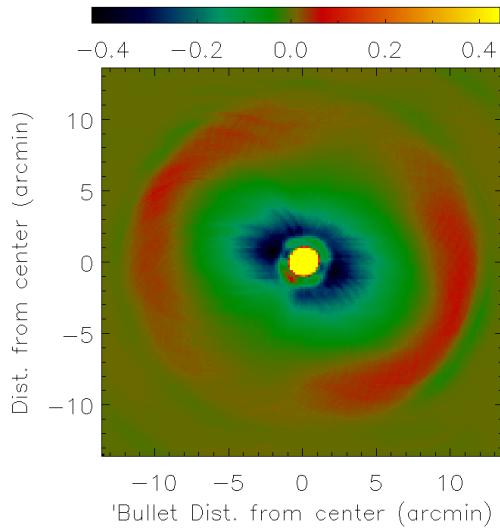


Figure 3.5: (Point source transfer function) The image shown here is the point source transfer function of the Bullet cluster. The negative lobes around the center, where a bright point source was simulated, is due to the low order polynomial baseline fitting.

3.5 Post-processed data products

The map making process developed for BoA is described by M. Nord (2009). The general scheme going from the time streams to projected map domain incurs loss of information due to the averaging over the time stream information. For all the data products obtained with the process described in the previous section, the final data products that would be desirable to use are the representation of the final products in the two dimensional angular map space. The filtering process, therefore, has been designed to behave in a stationary manner (essentially devoid of low frequency elements) so that the averaging over time samples can reliably represent the final product. Following sections introduce these final products that will be used in rest of this work for further analysis and measurements of the integrated Compton parameter of the SZ effect (see Section 2.2.2 of Chapter 2).

3.5.1 Co-added reduced source maps

The time signals of all the bolometers that are not flagged by the data reduction pipeline were coadded using the weights associated with each time signal series. Small units of time streams were reduced independently and each of these units were mapped using the weighted averaging the time streams. The final co-added image was produced by averaging over numerous scans (or units). Such a co-addition is represented in Figure 3.3. A final co-added image of Bullet cluster from APEX-SZ is shown in Figure 3.4. Each pixel of the final map has an estimated weight which is a representation of the effective integration time. An example of weights obtained for Bullet data is also shown in Figure 3.4.

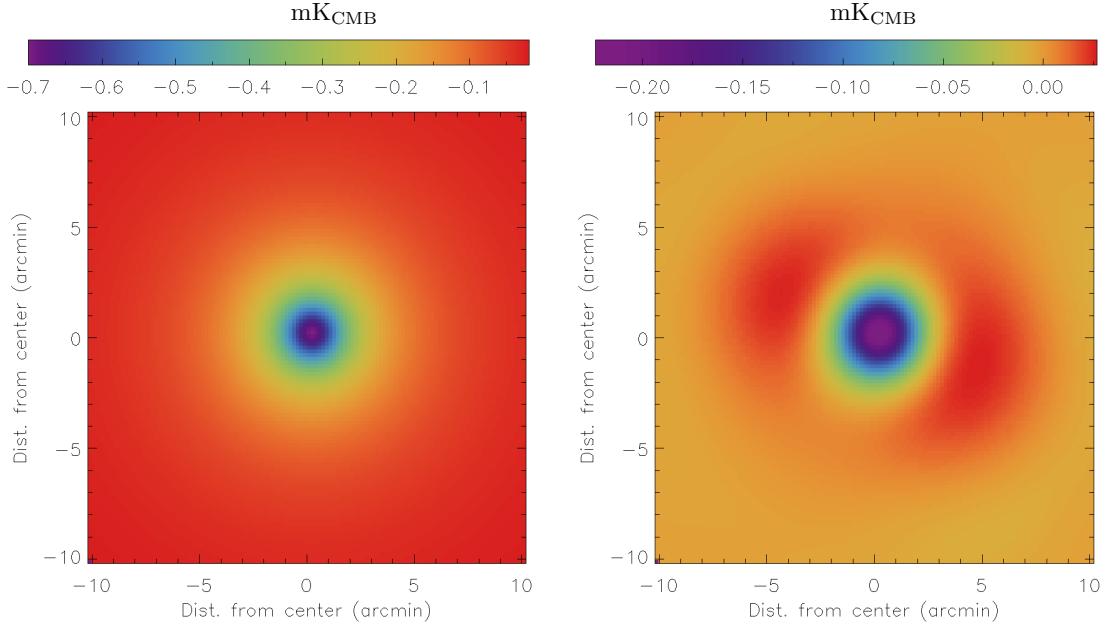


Figure 3.6: *Left:* Image of β model with core radius of $1.5'$ and ΔT_0 of $-0.7\text{mK}_{\text{CMB}}$. *Right:* The filtered β model produced by convolving the image on the left with the point source transfer function (PST) obtained for Bullet data. The attenuation of the SZ signal is illustrated in peak signal of the filtered model, which is less than half of the peak signal of the input model. The positive ring around the central region is due to the negative feature found in the PST (see Figure 3.5).

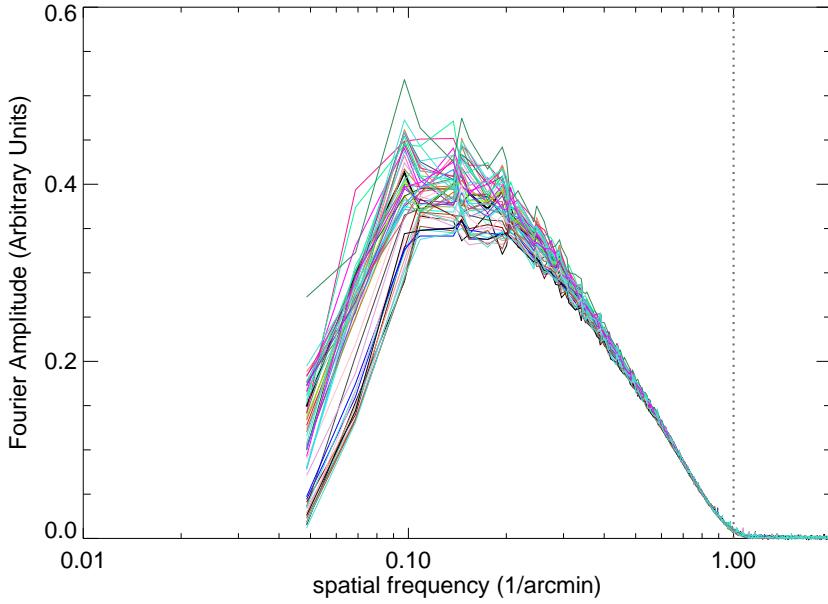


Figure 3.7: The Fourier transform amplitude of the point source transfer function for all APEX-SZ clusters in the sample are plotted here. The transfer function drops after the 10 arcmin scale, hence damping all the features larger than typically 10 arcmin scale in the reduced map. The dotted vertical line shows the scale size of the telescope beam at 150 GHz.

3.5.2 Jack-knived noise maps

The noise in the coadded maps of the APEX-SZ can be estimated by using the Jack-knived approach. The coadded maps were produced by averaging the signal at each pixel over several scans as described in previous subsection. The noise maps can be constructed by randomly flipping the sign of half of the scans used in the co-addition. This technique is referred to as jack-kniving, where the sample of scans or units are partitioned into approximately equal numbers and are subtracted from each other. In doing so, any astrophysical source that is likely to be present in all the scans would essentially be canceled out, for example, the primary CMB anisotropies. The resulting maps represent the noise arising from non-astrophysical origins. This procedure was used to produce 100 independent noise realisations from the scan maps. The jack-kniving is schematically shown in Figure 3.3.

3.5.3 Point source transfer function (PST)

The approximation of a linear PST allows us to construct a filtered mapping of any extended signal by convolving the true signal with the PST. This is used as a transfer function that modifies the original astrophysical signal on the sky into the filtered source. The total sum of the transfer function is expected to be zero, however, due to uneven weights in different part of the sky, when the polynomial baselines are applied to the time streams the sum is not strictly zero.

The data reduction method is tuned to filter out any low frequency noise. This limits the recovering of the full extent of the astrophysical source, which itself is thereby heavily filtered. Therefore, there is a partial loss of the source flux. This is demonstrated in Figure 3.6 using a simulation of an extended astrophysical signal. The scales upto which the signal can be recovered optimally depends on various factors including the size of the array, separation between the bolometer channels, and the radius of scan pattern. If the extent of the source signal is larger than the array size then the large scale source signal would be partially or fully lost. It is also important to note that the only case where the full extent of the signal can be fully recovered rests only for the sources which are smaller than the minimum separation between the channels. That is, the source is small enough that it does not produce any correlation between neighbouring channels. It must be noted that the time scale filtering that occurs in data reduction process does not relate to spatial filters on one-to-one level. But a Fourier transform of the PST helps us see that the filters in spatial scales behave effectively as a spatial scale band pass filter. Figure 3.7 show the Fourier transform amplitude of the PST for all APEX-SZ clusters.

3.6 Summary

The APEX-SZ experiment, data and analysis are introduced in this chapter. For a sample of nearly 40 galaxy clusters, the APEX-SZ collaboration has observed these targets in SZ, X-ray and lensing. A sub-sample of clusters forming a complete sample with a well-defined selection criteria is constructed. This sub-sample is an X-ray selected sample. The lensing mass measurements of these clusters are provided by Klein et al. (in preparation). The APEX-SZ observations of these targets have been reduced and final data products can be used further to extract the integrated Comptonization. This makes it possible to measure cross-correlation between lensing mass and the integrated Compton parameter, where the Comptonization can be obtained from APEX-SZ.

For this, an accurate measurement of the integrated Compton parameter from filtered the APEX-SZ data products is required. A measurement of this quantity has been done for APEX-SZ data processed through the MATLAB version of pipeline and is given by Bender et al. (2016). However, that work dealt with scaling between X-ray observables and SZ observable using various literature estimated X-ray observables. In this work, the global observable measured from the BoA pipeline is used for finally measuring the mass-observable scaling relations. The purpose is multi-fold as this provides independent measurements of the Y parameter, which is useful as a consistency check between both the pipelines, and a homogeneous re-estimation of Y using some knowledge from the lensing analysis (that is analysed in a homogeneous manner) is useful for controlling for any systematic effect and inferring the underlying relations more accurately. The actual method used for extracting the cluster observables from BoA data products and the results from such is described in detail in Chapter 4.

CHAPTER 4

Methods: measuring integrated Comptonization from APEX-SZ

Overview

The determination of an optimal method, in the sense that the extended Comptonization is both accurate and as precise as possible, to model the integrated SZ signal (Y_{SZ}) in the filtered APEX-SZ maps is the key focus of this chapter. The primary goal is to have an unbiased modelling of the SZ signal of galaxy clusters from APEX-SZ bolometer measurements in the presence of the strong filtering processes described in Chapter 3. By making use of a number of mock reduced APEX-SZ cluster maps, two methods of fitting the maps are discussed, namely: a) a Fourier domain fitting, where the fit of a model to data is performed in the 2-D spatial frequency space; b) Radial binning method, wherein the model and data are averaged into radial bins from the cluster centre before fitting. The quantitative analysis of the recovered parameters from mock simulations is used to find the optimal fitting routine to apply to all cluster dataset.

In Section 4.1, the motivation for exploring and testing different methods for extracting information on the SZ signal from APEX-SZ observations is elaborated. Following this in Section 4.2, all the necessary set-ups that would be used in this Chapter to fit models to data are established. The basic tools are introduced and parametric models, describing the ICM distribution and established in previous literature, are discussed. A quantitative and qualitative analysis of the two fitting methods are discussed in detail. The results from different modelling choices are discussed in Section 4.3. The conclusions and discussion based on the results are presented in the final section 4.4.

4.1 Motivation

Towards the aim of studying scaling properties of the intra-cluster gas medium pressure with total galaxy cluster mass, this Chapter will deal with the methodologies required for measuring the total SZ signal from APEX-SZ filtered images for the full sample of targeted clusters. Due to the strong filtering process of the data analysis described in Chapter 3, the filtered maps have attenuated astrophysical signal. In addition to the attenuation, the point source transfer function of the telescope

optics and data reduction introduce correlations in the image noise.

It is imperative to employ a fitting method that is capable of extracting an unbiased estimate of the integrated Compton signal (often denoted by Compton-Y or integrated Comptonization) from the filtered APEX-SZ co-added images. To set up the appropriate apparatus for measuring the integrated Comptonization, parametric models for the ICM pressure are used to simulate the APEX-SZ filtered images and each apparatus is rigorously tested on them. Two methods of fitting a parametric pressure profile to the APEX-SZ data are considered.

4.2 Method

4.2.1 Parametric models

A parametric modelling assumes a known template profile of the ICM pressure which are parameterised by a power law slope of cluster radius, a normalization, and a characteristic radius related to the cluster size. The models can be fit for their parameters using the data and employing a fitting scheme for a χ^2 approach for the likelihood of cluster model. The ICM can be characterized in different ways, bearing the thermodynamic state in mind. Below, I present two parametric models for the ICM which are widely used.

Isothermal β model

Assuming that the equation of state of cluster ICM to be in thermal equilibrium with the gravitational potential of galaxy cluster, the gas distribution relates to the galaxy count distribution (Cavaliere and Fusco-Femiano, 1976). Using this formalism, the gas distribution for an isothermal gas can be modelled with the King profile

$$\rho_{\text{gas}}(r) = \rho_{\text{gas}}(0) \left(1 + \frac{r^2}{r_c^2}\right)^{-\frac{3\beta}{2}}, \quad (4.1)$$

where r_c is a core radius. The β can be interpreted as the ratio of the specific kinetic energies of galaxies and gas medium (Cavaliere and Fusco-Femiano, 1976). Specifically,

$$\beta = \frac{\mu m_H \sigma_v^2}{k_B T_{\text{gas}}}, \quad (4.2)$$

where σ_v is the RMS line of sight velocity dispersion.

The projection of the β profile along the line of sight can be done analytically, yielding

$$y = \frac{\sigma_T k_B}{m_e c^2} n_{e0} T_e \left(1 + \frac{r^2}{r_c^2}\right)^{-\frac{3\beta}{2} + \frac{1}{2}}. \quad (4.3)$$

For all purposes in this work, β in equation (4.2.1) is set to 0.86 following the estimate from stacked SZ profiles of a few massive clusters (Plagge et al., 2010).

The Y parameter estimated using a β model is known to be sensitive to the assumed cluster size (e.g., Arnaud et al., 2010; Hallman et al., 2007). The SZ signal scales linearly with temperature and density of electrons in the ICM, when the temperature falls near cluster outer radii this is not sufficiently captured by the isothermal assumption. Hallman et al. (2007) comprehensively discusses such issues with fitting SZE using this model. Measuring Y_{500} using this model to fit the SZ signal would lead to a biased estimation. However, keeping this caveat in mind, the β model is still powerful for testing the apparatus due to its simple parameterisation compared to other parametric models. With an appropriate choice of the aperture, it can be used for measuring an integrated Comptonization and calibration to cluster masses (Saliwanchik et al., 2015). Besides this, it finds its use in being a spatial template for discovering galaxy clusters in SZ (e.g., Vanderlinde et al., 2010; Reichardt, Stalder et al., 2013; Bleem et al., 2015) and X-rays (e.g., Pacaud, Clerc et al., 2016).

Generalised Navarro-Frenk-White profile

A more realistic parameterisation of the electron pressure in the ICM was given by Nagai, A. V. Kravtsov and Vikhlinin (2007) based on the internal structure of dark matter halos found in simulations (Navarro, Frenk and White, 1995; Navarro, Frenk and White, 1996). The pressure is given by

$$P(r) = \frac{p_0}{\left(\frac{r}{r_s}\right)^\gamma \left(1 + \left(\frac{r}{r_s}\right)^\alpha\right)^{\left(\frac{\beta-\gamma}{\alpha}\right)}}, \quad (4.4)$$

where p_0 is a normalization, the logarithmic slope parameters α , β and γ describe the intermediate, outer and inner part of the pressure profile, respectively, and the scale radius (r_s) is related to R_{500} by the concentration parameter c_{500} as

$$r_s = \frac{R_{500}}{c_{500}}. \quad (4.5)$$

4.2.2 Generating mock APEX-SZ Compton-y images

For each cluster observed with APEX-SZ, there is a corresponding filtered co-added image of the cluster which consists of the filtered cluster SZ signal, noise and CMB anisotropies; a corresponding point source transfer (PST) function unique to each co-added cluster image; a set of 100 jack-knived noise maps which is devoid of any astrophysical source due to the jack-kniving process; and an image of the pixel weights for each cluster co-added image data. The pixel weights are an approximate representation of the integration time on each pixel of the map. Due to differences in sensitivity between bolometer channels, these weights are not one-to-one representations of integration time. The typical size of the pixel has dimensions of $10''$ by $10''$.

Filtered model

In order to compare a parametric model for the APEX-SZ reduced maps, it is required to predict how the model is affected by the filtering. The most accurate way of accomplishing this would be to model this effect on the true SZ signal, the source model would have to be sent through the data reduction pipeline to accurately represent the filtered SZ model. However, running this through the

Table 4.1: Summary of input parameters of the β model, type of filter function and type of noise used for mock data.

Mock data	core radius [arcmin]	ΔT_0 [mK _{CMB}]	Convolved with	Noise
Model 1	1.5	-0.8	PSF	White ($78 \mu K_{CMB}$)
Model 2	1.0	-0.3	PSF	White ($78 \mu K_{CMB}$)
Model 3	1.5	-1.0	PST	White ($78 \mu K_{CMB}$)
Model 4	2.17	-2.0	PST	White ($78 \mu K_{CMB}$)
Model A	1.0	-0.7	PST	Bullet
Model B	1.5	-0.7	PST	Bullet
Model C	2.17	-0.7	PST	Bullet

pipeline for each model at every step of a model fitting routine would be time consuming, given that the pipeline runs for few hours for reducing data of each target, making such a process prohibitively computationally expensive. The filtering steps have been carefully designed to make the transfer function approximately linear (M. Nord, 2009). All the subsequent analysis is carried out under the assumption that the point source transfer function (PST) is approximately linear. The filtered model in the image plane can be obtained by convolving the model with the point source transfer function (PST). The typical process for generating a mock filtered APEX-SZ image first begins with assuming a radial profile using either the β model profile or gNFW and create a two dimensional azimuthally symmetric image of the profile. The map size was typically 90' by 90'. This size is the same as that for the image of the PST. The filtered SZ model was produced by convolving the model image with the point source transfer (PST) function. In practice, it was produced using the IDL function `fft_convolution`.

Noise realizations

Since the aim is to test different methods and determine the robustness of these methods on obtaining estimates on global SZ properties, the mock images generated included two different types of noise. A set of mock images used the *white noise* with no pixel-to-pixel correlation. These mock data were used as a control for testing the methods. Another more realistic set of mock data were used for the main analysis for drawing inference on the optimal method. These were generated using the jack-knived noise realizations from APEX-SZ observations. Primarily, the Bullet cluster noise field was chosen. The Bullet cluster was detected with highest peak signal to noise of ~ 16 in the APEX-SZ beam-smoothed maps. A detailed study of the cluster profile using this data was done by Halverson et al. (2009).

Table 4.1 shows a list of model parameters, kind of transfer function used to model a filtered signal and the type of noise added to various mock data. The first two models in the table were generated by convolving the true signal with the point spread function of the APEX-SZ beam. The rest of the models used the point source transfer function (PST). The white noise maps were generated with an rms of $78 \mu K_{CMB}$.

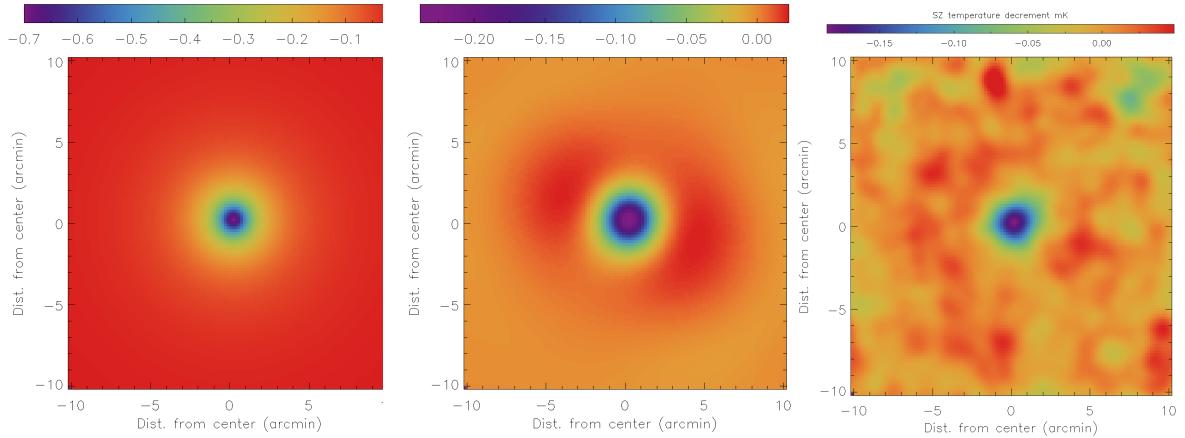


Figure 4.1: *Left:* An input model used in the simulations of mock SZ images. Here, a β profile for the SZE profile was assumed with a core radius equal to the beam size i.e. 1' and a temperature decrement of $-0.7 \text{ mK}_{\text{CMB}}$. *Middle:* The input model was convolved with the point source function which is a linear representation of the filtering process of the reduction pipeline. This gives the filtered SZ image shown here. *Right:* One of the jack-knived noise realisation of the APEX-SZ observations of the Bullet cluster field was added on top of the filtered SZ image as an example. This is a typical realistic mock image of APEX-SZ observations of a galaxy cluster's SZ intensity.

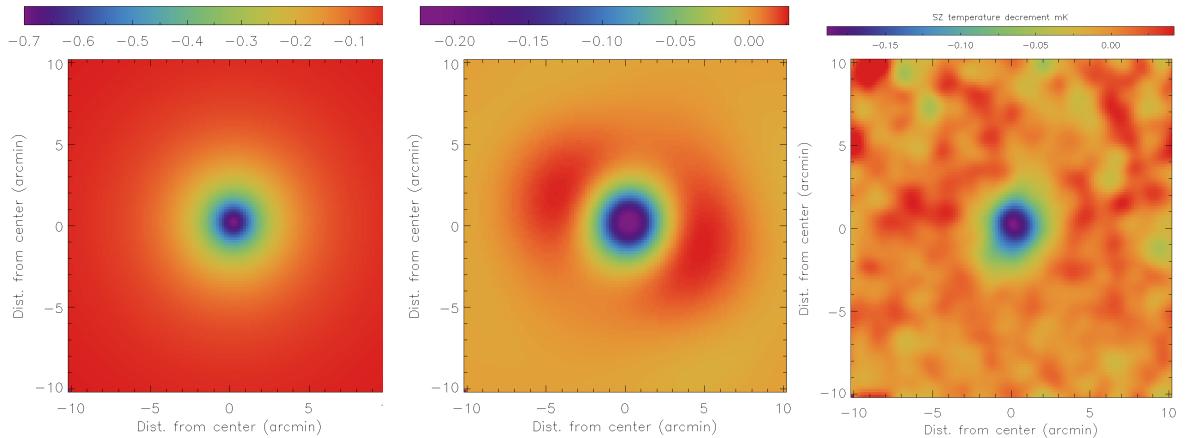


Figure 4.2: Same as Figure 4.1 for an input model with core radius of 1.5 arcmin.

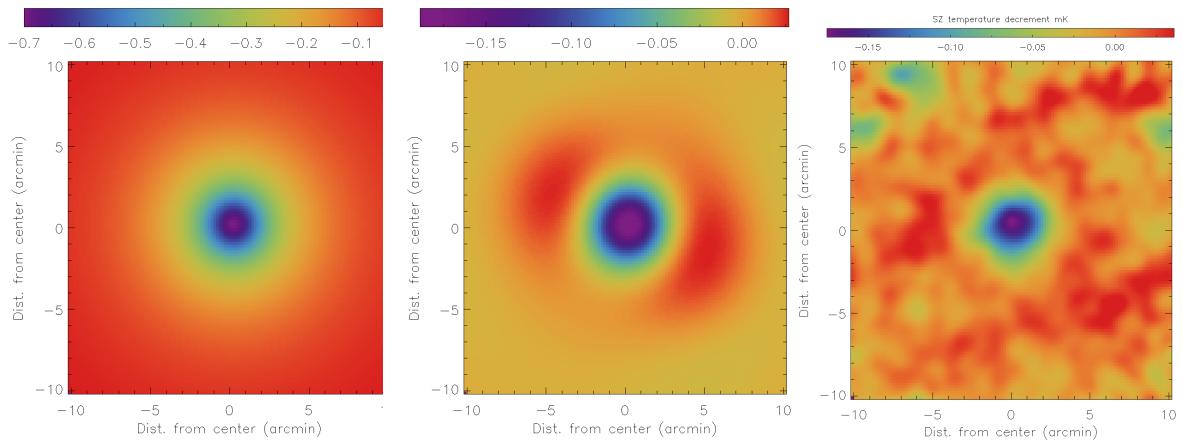


Figure 4.3: Same as Figure 4.1 for an input model with core radius of 2.2 arcmin.

4.2.3 Bayesian fitting approach

A general and more formal description of the Bayesian fitting processes can be found in Gelman et al. (2014). Here, a short description of the formalism is presented. In the Bayesian approach, the underlying model parameters, θ , are inferred by sampling the distribution of θ for a given set of observations, \mathbf{d} . This as per the Bayes theorem gives the posterior density:

$$P(\theta|\mathbf{d}) = \frac{P(\mathbf{d}|\theta)}{P(\mathbf{d})} P(\theta). \quad (4.6)$$

$P(\theta|\mathbf{d})$ is the conditional probability of θ given the observation \mathbf{d} . $P(\mathbf{d}|\theta)$ is the conditional probability that \mathbf{d} occurs given θ . $P(\mathbf{d})$ amounts to $\int P(\mathbf{d}|\theta)P(\theta)d\theta$, where the θ is set of continuous variables. $P(\theta)$ is the prior probability on the model parameters. $P(\mathbf{d})$ generally is a constant for a set of observations. Hence, the denominator of the equation can be ignored while fitting a given set of data.

The likelihood, $\mathcal{L}(\theta|\mathbf{d})$, of a model given the observation is related to the χ^2 and is given by $\mathcal{L} \propto \exp(-\chi^2/2)$. The χ^2 is defined as follows:

$$\chi^2 = (\mathbf{d} - \mathbf{m}(\theta))^T \mathbf{C}^{-1} (\mathbf{d} - \mathbf{m}(\theta)). \quad (4.7)$$

This non-linear model with multiple parameters can be efficiently fit using the Markov chain Monte Carlo (MCMC) method. The Metropolis-Hastings algorithm is used to give an approximate posterior distribution of the parameters.

4.2.4 Fitting methods

4.2.4.1 Fourier Domain (FD) Fitting method

The analyses and reduced maps from the APEX-SZ data typically have $10'' \times 10''$ pixel size and the region of sky mapped is of the order of $30'$. Fitting the model in the pixel space constitutes using a large and cumbersome array in pixel vector space . The noise covariance matrix would require an

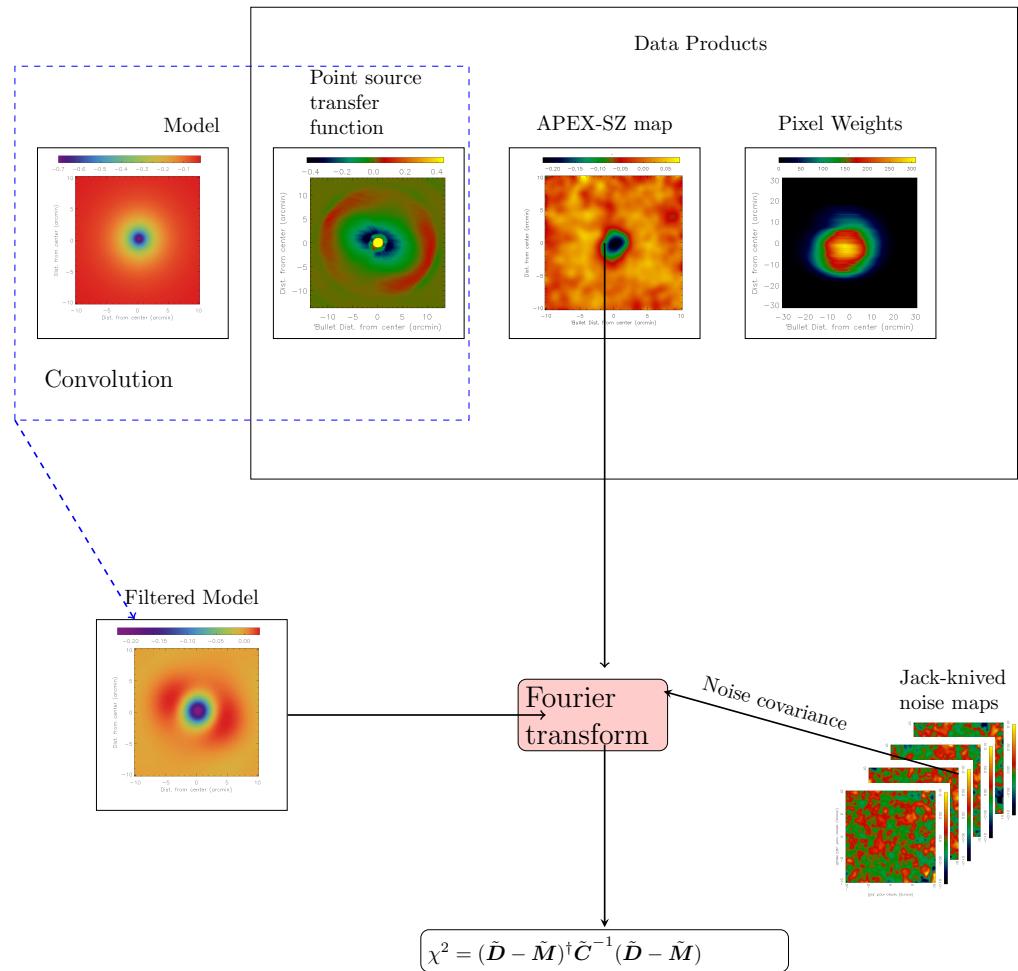


Figure 4.4: Schematic description of the Fourier domain fitting.

equal number of no. of noise realizations to account for the pixel to pixel correlations. Since the χ^2 computation requires inverting this covariance matrix, for it to be invertible, the pixel to pixel covariance should have non-zero determinant. Due to the limited integration time per target, there is not a sufficient number of independent noise realizations to determine all the eigenvalues of the covariance matrix. It is possible to simplify this problem, if noise stationarity across the map is assumed (J. Sayers, Golwala et al., 2011). Under this assumption, a Fourier transform of the noise covariance matrix from real space to spatial scales provides a diagonal covariance matrix, whose elements represent the noise power spectrum. This has the advantage of using the information from the data to the largest degrees of freedom and at the same time is useful for fitting spherical models (e.g., Halverson et al., 2009).

The conditions for a stationary noise in the 2-D map space are characterized by the following:

1. the noise in the time streams are stationary. By definition, the filtering steps have been designed so that the time streams have flat spectra and non-stationary components would be filtered out. Hence, this is fair assumption in terms of APEX-SZ bolometer time-streams.

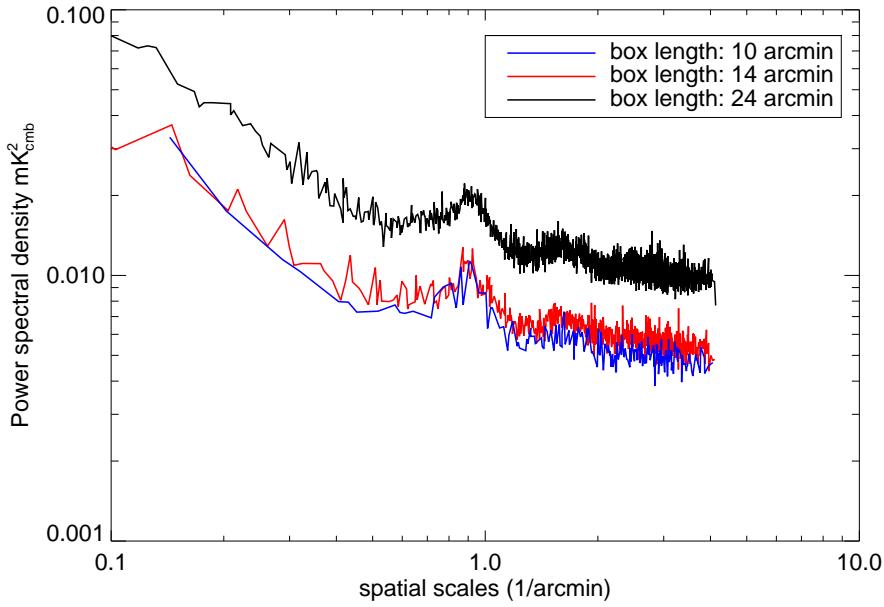


Figure 4.5: The noise power spectral density (PSD) in the Bullet cluster field inside three different central areas. The PSD is a measure of average noise power in the region and this is essentially the noise covariance matrix for the χ^2 . A higher PSD implies the weights for the spatial scales are lower in the χ^2 .

2. the coverage across the map is uniform. This assumption for the APEX-SZ observations is valid only in a small area of the map, typically within $\sim 5'$ radius about the center of the map. The dominant source of non-uniformity of the weight is due to the scan strategy, which is a circular pattern that optimises the time spent on target. This gives rise to a slow decrease in weights in radial direction away from the center of the target.

Fitting procedure

1. The reduced APEX-SZ map of a cluster is cut to a desired box size (typical choice of $10' \times 10'$) about the center of the cluster.
2. The cut out image is Fourier transformed.
3. The noise realizations are cut in the same way. The Fourier amplitude is average summed to obtain the power spectral density of the noise inside the $10' \times 10'$ region. This is essentially the noise covariance matrix.
4. The SZ model image is generated using a set of β model parameters sampled by the MCMC. The size of the image is typically $90' \times 90'$. This size corresponds to the image size of the APEX-SZ PST. In order to compare with the filtered image data, the SZ model is convolved with the PST image.
5. The filtered model map is cut about the same center and the same box length as in step one and Fourier transformed.

6. The χ^2 is computed as per the equation:

$$\chi^2 = (\tilde{\mathbf{D}} - \tilde{\mathbf{M}})^\dagger \tilde{\mathbf{C}}^{-1} (\tilde{\mathbf{D}} - \tilde{\mathbf{M}}), \quad (4.8)$$

where $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{D}}$ are the Fourier transform of the images of the filtered model and data respectively. $\tilde{\mathbf{C}}$ is the covariance matrix defined in the spatial scales domain. The χ^2 is used to compute the likelihood for the set of input β model parameters.

7. The process is repeated for steps 4,5,6 at each step of the MCMC while exploring the posterior of the model parameters.

A schematic representation of the Fourier domain fitting method for APEX-SZ is given in Figure 4.4.

The noise covariance is diagonal in the spatial frequency domain due to the stationary noise assumption. In a weakly stationary process, the covariance between two events occurring at different times is dependent only on the time lag between events and not on the actual times when the events occur. Such a covariance for stationary noise is an auto-covariance which is essentially the power spectral density.

For each noise realisation map, its Fourier transformation yields a 2-D complex vector \tilde{N}_m . The power spectral density is the average Fourier amplitude of the M number of noise realizations $(1/M)\sum_{m=0}^{M-1} |\tilde{N}_m|^2$. In the case of white noise, the power spectral density is a flat spectrum. In the APEX-SZ filtered noise, the power spectral density is shown in Figure 4.5 for different area around the cluster center in the Bullet cluster field. A reminder is that this is essentially the noise covariance matrix (diagonal) for the χ^2 . Note that for larger area, the noise power is actually estimated to be larger than in smaller regions. The PSD itself does not encapsulate any directional information or position information of noise properties, it estimates only the noise amplitude under the stationary process assumption. The increase in noise power is due to the variation in the pixel noise of the APEX-SZ maps away from the center of the cluster where the RMS actually increase due to lower integration time. This limits the use of this method to sufficiently smaller regions (nearly 1/9 th area) of the cluster map.

4.2.4.2 Radial binning (RB) method

As we are concerned with fitting an azimuthally symmetric model about a center of the cluster, a radial binning method could be adopted where the image is binned into radial bins. In each such bin the SZ intensity is weighted averaged using the pixel weights. This method takes into account the varying noise rms in the maps. Thus, in the process the pixel weights are taken into account while performing the fitting, unlike the Fourier domain. The two dimensional information is reduced due to the assumption of azimuthal symmetry.

Even though there is effectively some loss of the morphological information due to the symmetry assumption, this method can be employed to recover the radial information up to the extent of the maps in the APEX-SZ (typically radial extent of nearly 15').

The choice of the center for the radial profile might be the cluster center that best describe the gas distribution of the ICM. By construction, the radial binning method requires some prior knowledge of the cluster centre and is not meant for changing centres of the data and profile for the χ^2 comparison.

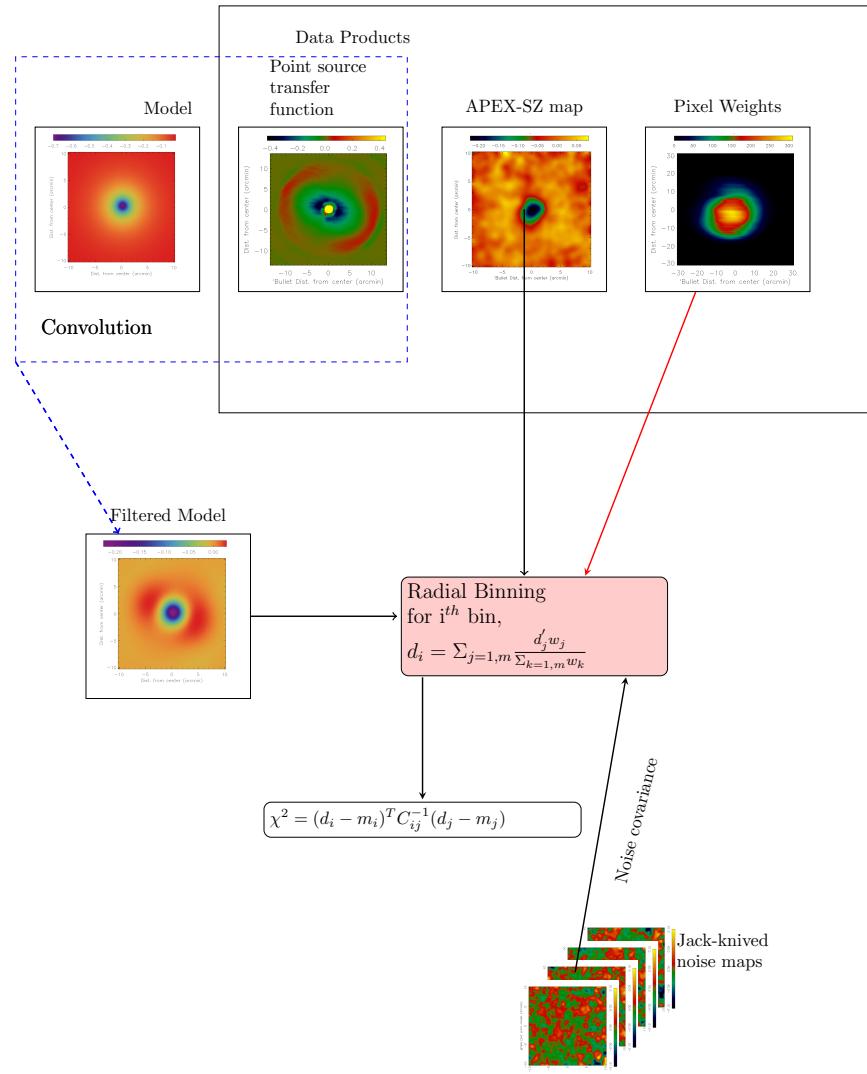


Figure 4.6: Schematic representation of the radial binning method. This method takes into account the variations in the pixel weights across the APEX-SZ maps.

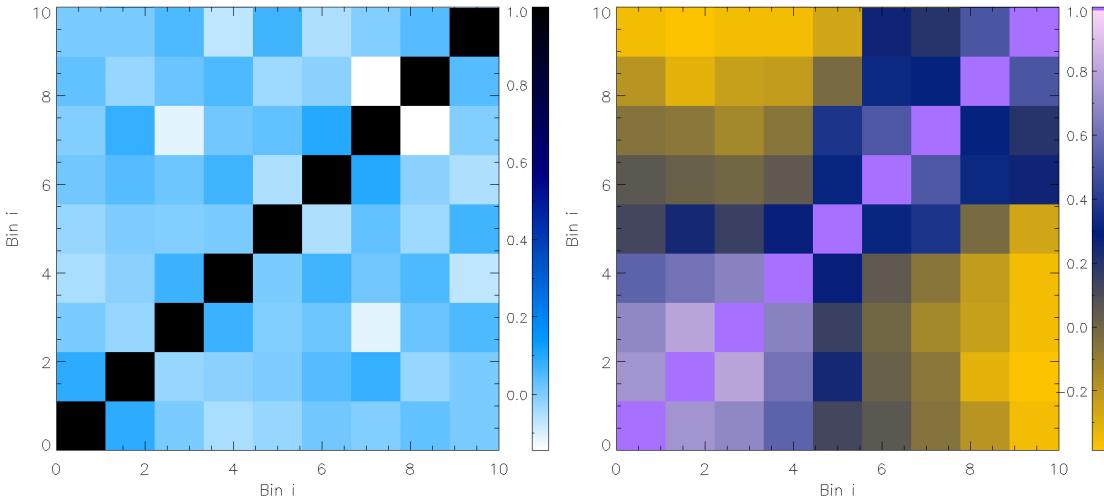


Figure 4.7: Bin-to-Bin correlation matrix. *Left*: White noise. The correlation matrix is essentially a diagonal matrix. *Right*: Bullet cluster field noise. The noise correlation matrix for the real noise show correlations between neighbouring bins due to the telescope beam. The anti-correlation between the noise in inner and outer bins due to the low-order polynomial baseline applied to the APEX-SZ observations.

Therefore, for this method the centres for the fit is always kept fixed in the analysis. The χ^2 in Equation (4.2.3) is computed with \mathbf{d} , \mathbf{m} vectors of the radially binned data and the filtered model respectively, and \mathbf{C} is the bin-to-bin covariance.

The procedure is described in following steps:

1. An appropriate centroid for the radial profile of the pressure model is first chosen. This centroid, once chosen, is kept fixed throughout rest of the process. A maximum radius of the outer most annulus is set by the region of data chosen for the fitting.
2. From the center of the radial profile, annuli are chosen to have a width corresponding to t in units of arcmin or pixel are identified. The central most bin is a circular area of πt^2 . The subsequent annuli have the area $\pi(2i - 1)t^2$. There are less number of pixels in the inner annuli bins than in the outer bins. Hence, the noise rms in outer bins are expected to be smaller than the inner ones (until the noise increase in the outer bins takes over).
3. In each radial bin i the average signal is obtained by using the relation

$$d_i = \sum_{j=1,m} \frac{d'_j w_j}{\sum_{k=1,m} w_k}, \quad (4.9)$$

where indices j and k ranging from 1, ..., m refer to the pixel indices within the annulus. The weights w represents the pixel weights. Whenever white noise is considered, the pixel weights are constant across the map. In case of APEX-SZ maps, the pixel weights for the corresponding cluster data is used.

4. Each noise realisation is radially binned from the chosen center. This ensemble of noise realizations in each bin is used to compute the bin-to-bin covariance.

5. A filtered SZ model is generated by convolving the true SZ model signal in 2-D image with the PST image is generated. This is radially binned in the same way as described in steps above.
6. Equation 4.2.3 is used to compute the χ^2 .
7. The steps 5 and 6 are repeated at each step of the MCMC to determine the posterior of the model parameters.

To be able to fit model to data in radial bins using the χ^2 , a bin-to-bin noise covariance matrix needs to be computed from an ensemble of noise realizations.

The bin-to-bin noise covariance matrix for white noise and Bullet cluster field noise are shown in Figure 4.7. It can be seen that the white noise covariance is essentially a diagonal matrix. The Bullet cluster noise field covariance matrix shows correlations between adjacent bins and anti-correlation between inner and intermediate bins. The correlations are essentially due to the telescope optics. The anti-correlation is due to the low-order polynomial baseline applied to the time streams.

4.3 Results

After generating the mock data for different stages of testing and introducing the two candidates for fitting models to data, the results from different stages of testing are presented in this section. First, the Fourier method is employed for the analysis and is tested on white noise and real noise mock data. These are presented together in Section 4.3.1. Later, the radial binning method is employed for similar analysis with mock sets and these results are presented in Section 4.3.2. Following this, a comparison of the two methods is performed to have a qualitative understanding of the differences and relative advantages and disadvantages in these methods.

4.3.1 Testing the Fourier Domain (FD) method

White noise generated mock data

To test the method, two models (Model 3 and 4) from Table 4.1 were used. Each model was produced with different noise realizations. Over 80 mock realizations of each of these models were fit using the Fourier transform method described in Section 4.2.4.1. Four parameters of the β model were fit for each realizations: the position of the centroid, ΔT_0 and r_c . The noise covariance for this case is the power spectral density, which is a flat spectrum. The χ^2 were computed within a $10' \times 10'$ box about the centre of the image, where the SZ signal was originally simulated. The size and the region of the data to fit were always kept constant throughout. As described before, the multivariate posterior distribution for each of the model parameters were obtained through an MCMC fit.

The posterior distribution obtained for each mock data was used for further analysis. The marginalised posterior distribution of each of the four parameters were considered separately. For each parameter, the mode of every recovered posterior is sampled (in total there were over 80 such values each one corresponding to one mock realisation). The drawn modes together form a sample distribution of the modes of the recovered posteriors. This sample distribution of modes for each parameter are shown in Figure 4.8 and 4.9 for Model 3 and Model 4 respectively.

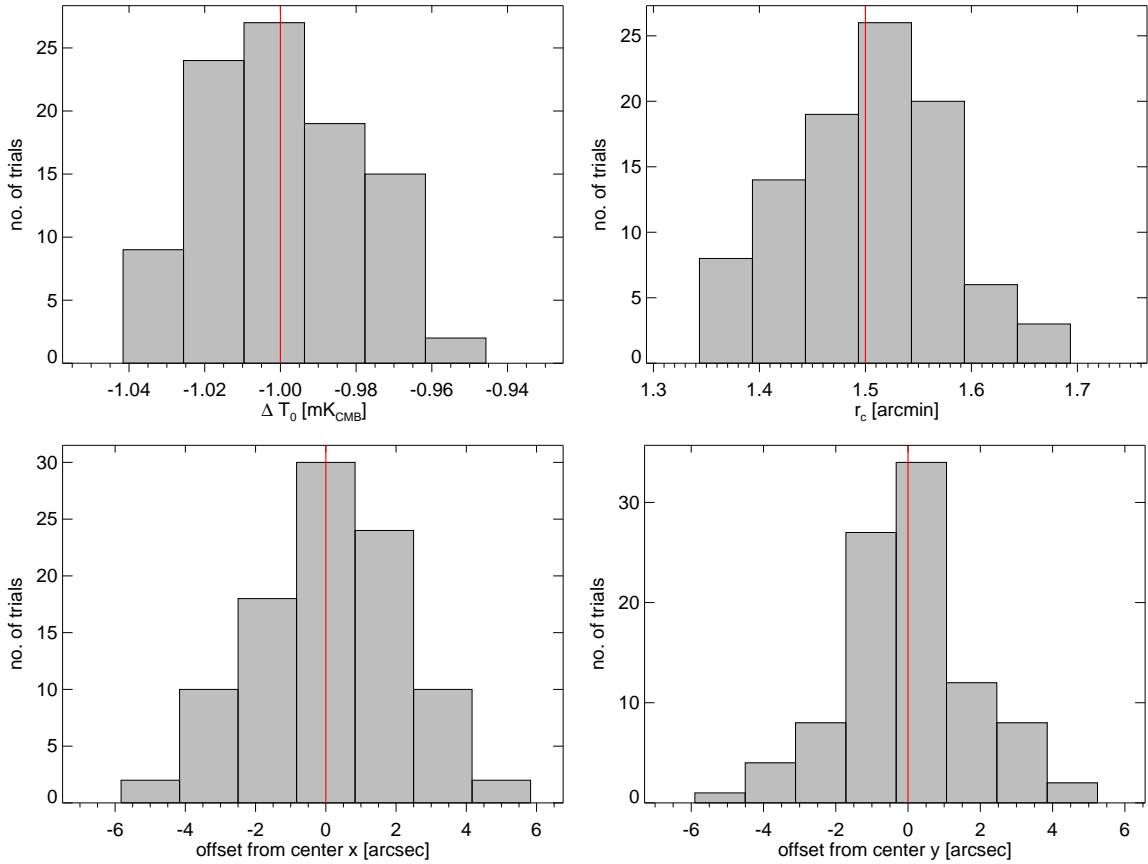


Figure 4.8: Model 3: Distribution of modes for each parameter from multiple mock data. The red lines indicate the input parameter values used for generating mock data with white noise.

The means of the sample distributions were chosen as a statistical indicator for estimating the bias in the recovered parameters. The estimated mean for each parameter from the sample distribution was used to quantify a bias in recovering the corresponding parameter by the method. For this, the 68% percent confidence level on the mean was computed using a bootstrapped procedure. Random realizations of the sample were obtained by randomly drawing elements from the sample and allowing for multiple copies of an element. Mean of each such sample was computed and the distribution of this mean was used to obtain 68% confidence level of the mean. The bootstrapped distribution of the mean for ΔT_0 and r_c is shown in Figure 4.10. This confidence level is expected to be much smaller than the square root of the sample variance due to numerous mock fits. If the sample distribution were perfectly Gaussian, then the confidence level on the mean corresponds to the standard error of the mean. This relates to the standard deviation σ as $\sigma_{\text{standard error}} = \sigma / \sqrt{N}$, where N is the number of mock fits used to obtain the sample distribution of modes. The significance with which the bias in the recovered parameters can be quantified depends on the number of mock data used. Having more mock data to fit improves the standard error or the confidence level on the mean of the sample distribution.

The input parameters of the models used in the mock images are recovered by the fitting procedure with no evidence of any bias in the recovered parameter confidence regions.

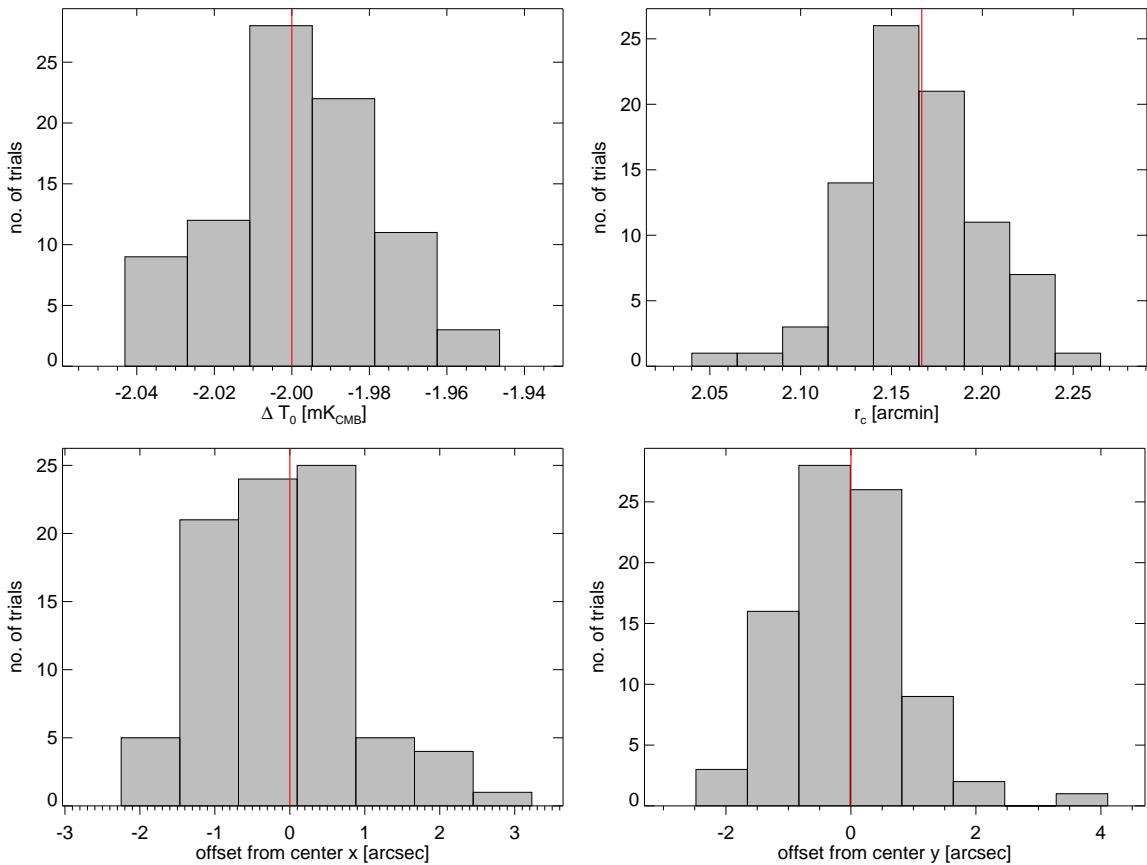


Figure 4.9: Model 4: Distribution of modes for each parameter from multiple mock data. The red lines indicate the input parameter values used for generating mock data with white noise.

Real noise generated mock data

Here the models Model A, B and C were considered for testing the method on realistic APEX-SZ mock images. These models differ only in their values of core radius. The purpose is to analyse the performance of the Fourier domain method in its accuracy of recovering the simulated model parameters. In other words, by using the results from running the fit on multiple mock images for each model, an estimation of a systematic bias in recovering parameter values is made possible.

For each model, over 75 mock images were fit. All images were fit using the data in a region within 10' by 10' box about the centre of the simulated cluster profile. Following the procedure in the previous section, the mode of a posterior distribution of each parameter from each of the fit is sampled. Together for each parameter, a sample distribution of recovered values is obtained. The resulting sample distribution of recovered values from over 75 mock realizations are shown for the Model A, B and C in figures 4.11, 4.12 and 4.13 respectively. Similar to the procedure followed in previous section, the bootstrap of the mean of the sample distribution was used to obtain 68% confidence level on the sample means. The bootstrapped distributions showing the confidence regions and the input values of each parameter are plotted in Figure 4.14. The focus is placed on just the ΔT_0 and r_c as they are of primary interest. From the figures, the input values can be seen to be within 68% confidence region of the sample mean of recovered parameter space for both ΔT_0 and r_c . Although not shown

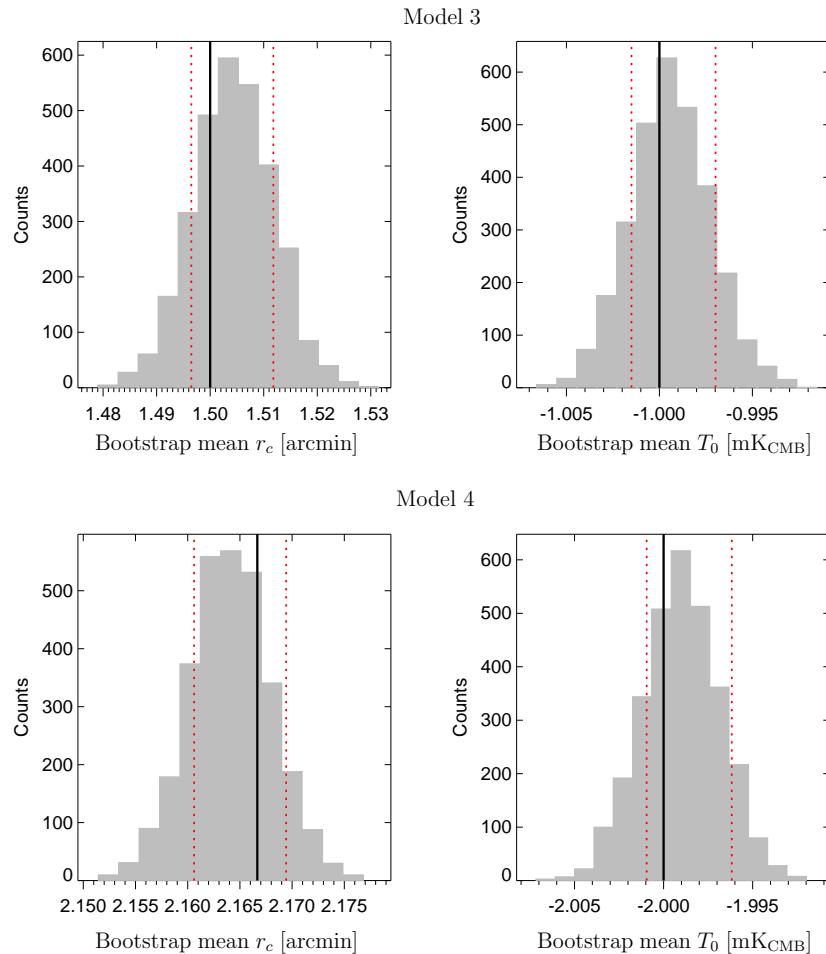


Figure 4.10: Model 3 & 4: Bootstrapped confidence region of the mean value recovered from mock fits for ΔT_0 and r_c . The black solid line indicates the input value of the parameter used for simulating the mock Compton-y images. The red dotted lines represent the 68% confidence region of the distribution. The input values are well-within this confidence region on the average estimated value.

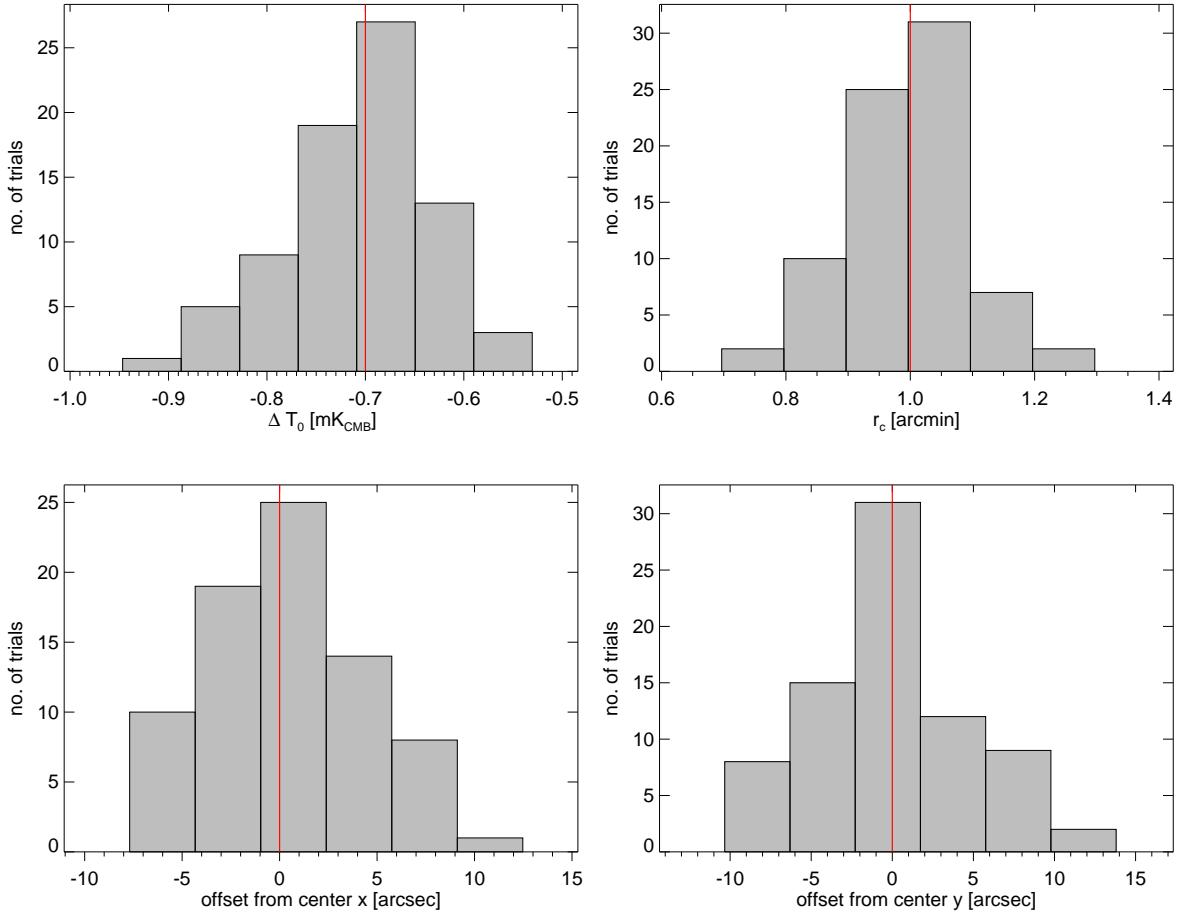


Figure 4.11: Results for Model A. The distribution of the recovered modes from fitting each of the mock images generated using the Bullet noise fields.

here, the position parameters recovered do not show any hint of bias in them.

The tests within $10' \times 10'$ box size show that the recovered parameters remain unbiased with respect to the input model parameters used for simulated mock images. This can be said for all the four parameters. Consequently, the integrated Compton-y parameter is recovered in an unbiased manner for all the three models used for testing. No special trend could be recognized in the recovered parameters with respect to the cluster core radius. In addition to the bullet noise field, XMM-LSS field observation was also used for generating mock data. Over 90 realizations of the mock images for two sets of β models were used for the test. The resulting tests with this field provide confirmation for concluding that there is no evidence for any bias in the recovered parameter space.

A summary of results from all of the above tests is given in Table 4.2. The values quoted in the Table correspond to the average deviations of recovered value of parameters from the input simulated model values. For an unbiased estimate this mean deviation should be consistent with zero within its 68% confidence level. From the tabulated results, we indeed find the recovered best-fit values from numerous mock fits to be consistent with the simulated input model parameters. Therefore, we conclude that there is no significant bias in recovering the best-fit values of the model parameters.

Table 4.2: Summary of recovered input values from different mock realizations. The fourth column gives the mean deviation from the input parameter value (p_{input}) and 68% confidence level about the mean deviation of the sample distribution. The final column gives the mean deviation and its confidence level determined by bootstrapping the sample mean.

Mock	No. of	Model input values	Mean deviation	Mean deviation and
data	mock data		and 68% confidence	the 68% confidence level
model	fit		levels of the sample	on the mean
	N	p_{input}	$\langle p - p_{\text{input}} \rangle_N [10^{-3}]$	$[10^{-3}]$
White noise mock data				
Model 4	85	$\Delta T_0[\text{mK}_{\text{CMB}}] = -2.0$	$-1.0^{+23.0}_{-16.0}$	-1.0 ± 2.0
		$r_c[\text{arcmin}] = 2.167$	$-3.0^{+50.0}_{-30.0}$	-3.0 ± 6.0
Model 3	95	$\Delta T_0[\text{mK}_{\text{CMB}}] = -1.0$	$-1.0^{+47.0}_{-18.0}$	-1.0 ± 3.0
		$r_c[\text{arcmin}] = 1.50$	$+4.0^{+67.0}_{-82.0}$	$+4.0^{+9.0}_{-7.0}$
XMMLSS noise mock data				
XMMLSS-1	85	$\Delta T_0[\text{mK}_{\text{CMB}}] = -1.0$	$+3.0^{+70.0}_{-20.0}$	$+3.0 \pm 3.0$
		$r_c[\text{arcmin}] = 1.50$	$+10.0^{+90.0}_{-50.0}$	$+1.0 \pm 10.0$
XMMLSS-2	97	$\Delta T_0[\text{mK}_{\text{CMB}}] = -1.0$	$+4.0^{+43.0}_{-39.0}$	$+4.0^{+5.0}_{-4.0}$
		$r_c[\text{arcmin}] = 2.167$	$+13.0^{+130.0}_{-100.0}$	$+13.0^{+10.0}_{-20.0}$
Bullet field noise mock data				
Model A	77	$\Delta T_0[\text{mK}_{\text{CMB}}] = -0.7$	$+5.0^{+57.0}_{-63.0}$	$+5.0^{+10.0}_{-7.0}$
		$r_c[\text{arcmin}] = 1.0$	$-15.0^{+62.0}_{-102.0}$	$-15.0^{+15.0}_{-10.0}$
Model B	79	$\Delta T_0[\text{mK}_{\text{CMB}}] = -0.7$	$-1.0^{+60.0}_{-80.0}$	$-1.0^{+11.0}_{-8.0}$
		$r_c[\text{arcmin}] = 1.50$	$-10.0^{+280.0}_{-120.0}$	$-10.0^{+30.0}_{-10.0}$
Model C	98	$\Delta T_0[\text{mK}_{\text{CMB}}] = -0.7$	$-2.0^{+67.5}_{-67.3}$	$-2.0^{+13.0}_{-9.0}$
		$r_c[\text{arcmin}] = 2.167$	$+7.0^{+320.0}_{-190.0}$	$+7.0 \pm 20.0$

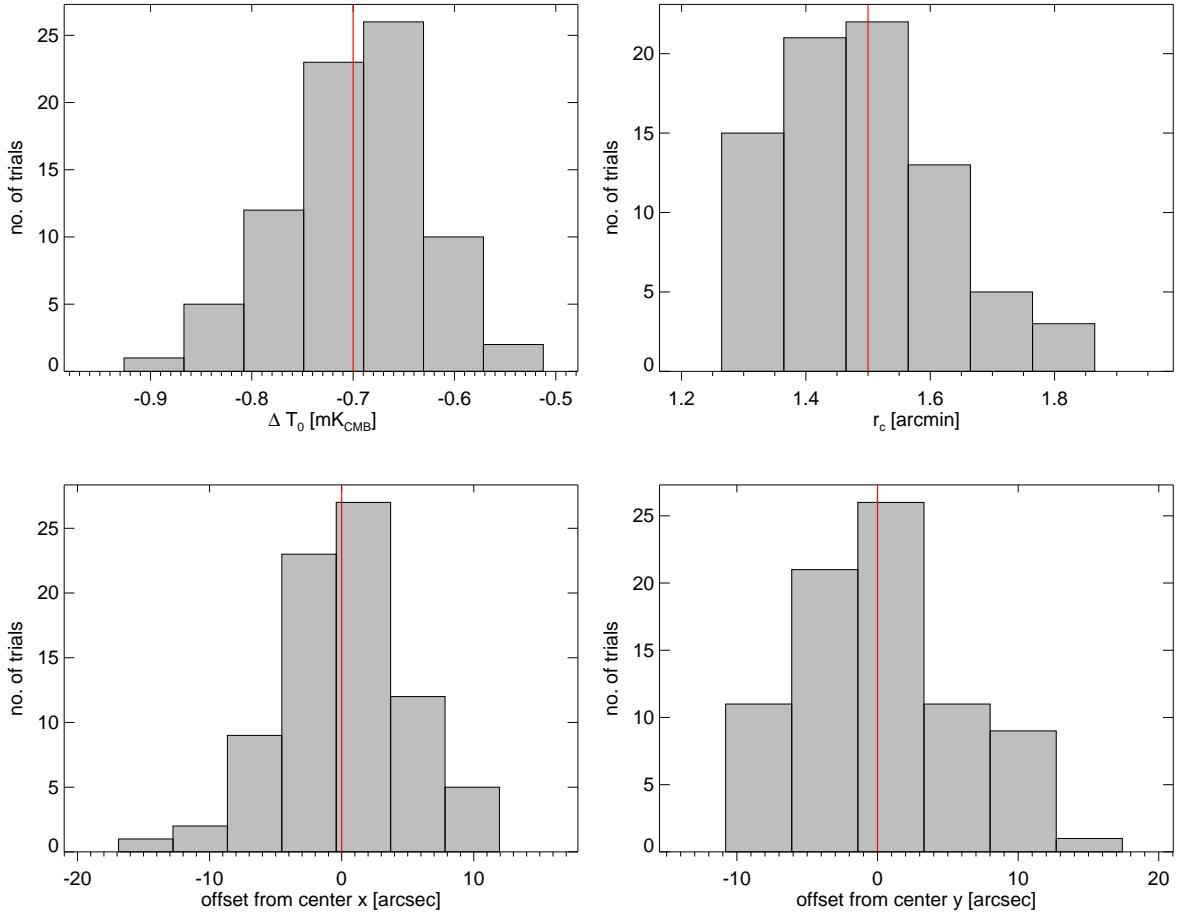


Figure 4.12: Same as Figure 4.11 but for Model B with core radius of $1.5'$. Although it appears as though the distribution of the recovered core radius is highly skewed and shows some lack of preference for smaller core radius, it is not to be concluded that this is a hint of a biased distribution. Rather this is an indication of incompleteness of the noise realizations limited by the integration time. A similar model was used to test with another noise field (XMM-LSS) and that did not show such level of skewness.

4.3.2 Testing the Radial Binning (RB) method

The procedure for testing the radial binning method is same as done for the Fourier method.

White noise generated mock data

The Model 3 reported in Table 4.1 was considered for testing the method. The input values of the β model parameters were $\Delta T_0 = -1.0 \text{ mK}_{\text{CMB}}$ and $r_c = 1.5'$. The mock data were generated with white noise without any pixel to pixel correlation for the purpose of testing the method. The model was convolved with the PST in the image plane to obtain the filtered SZ model. White noise was used in the mock images with an RMS of $78 \mu\text{K}_{\text{CMB}}$. The method described in Section 4.2.4.2 was followed. The schematic description of the method is shown in Figure 4.6. The fit was done within a circular region of radius $9.5'$. This choice of radius is arbitrary as the radial binning

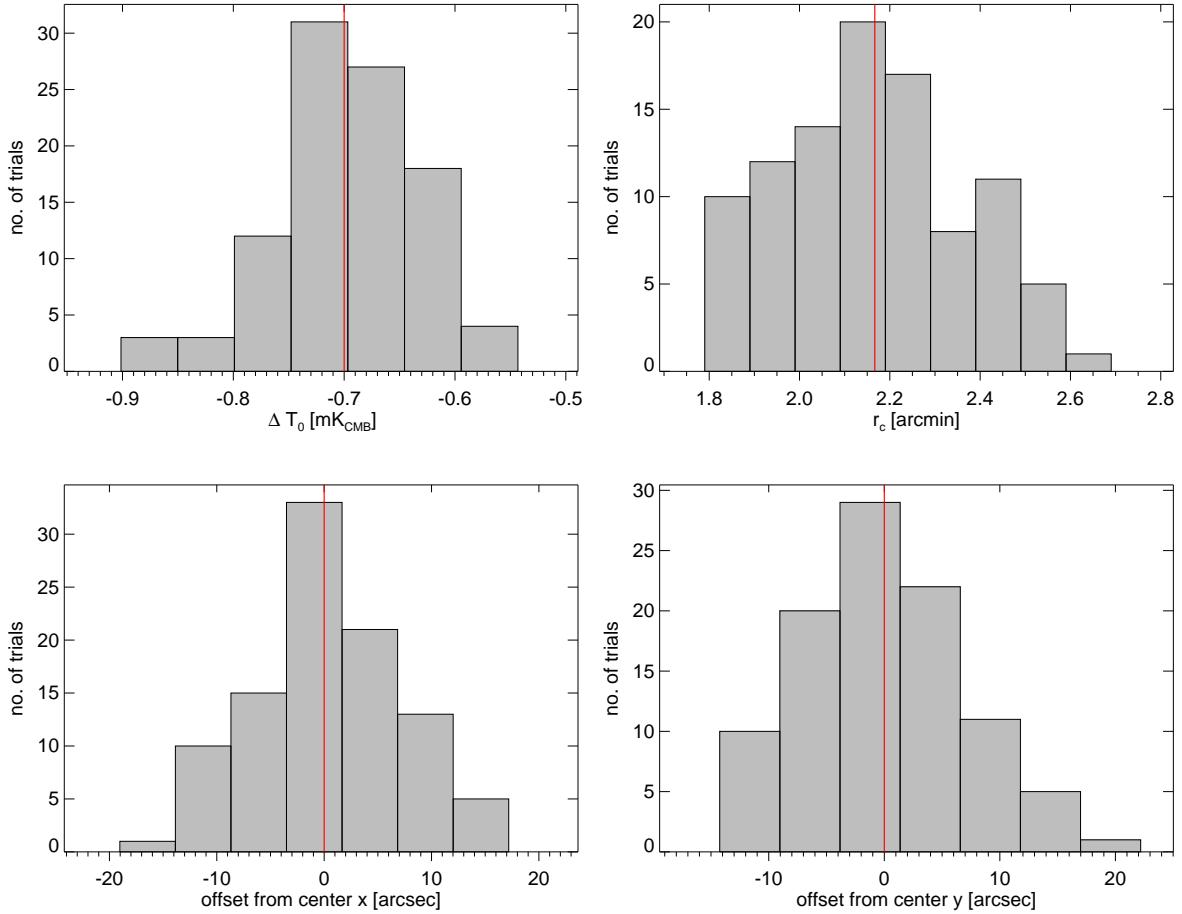


Figure 4.13: Same as Figure 4.11 but for Model C with core radius of 2.17'.

method does not place any theoretical restrictions on the size of the region to be used for fitting. The pixel weights were considered to be uniform. A total of 20 realizations of the mock images were fit using this method. Two parameters of the β model were fit, the ΔT_0 and r_c . The sample distribution of modes from 20 realizations gave the sample mean and the 68% confidence level for $\Delta T_0 = -1.004^{+0.020}_{-0.025}$ mK_{CMB} and $r_c = 1.502^{+0.057}_{-0.069}$. The mean of the sample distribution of recovered modes and the bootstrapped confidence level on the mean (equivalent to standard error of the mean) from the fit are $\Delta T_0 = -1.004^{+0.005}_{-0.005}$ mK_{CMB}, $r_c = 1.502^{+0.014}_{-0.014}$. The input values for both parameters of the simulated β models used in the mock images were recovered within the 68% confidence level of the mean of the recovered parameter sample distribution. The method showed no sign of a bias in the recovered parameters.

Real noise generated mock data

In this section, models A, B, and C reported in Table 4.1 were considered. The bin-to-bin correlation for such a noise is shown in Figure 4.7 and was discussed in Section 4.2.2.

The weights of the Bullet data were used in the radial binning as per Equation 4.9. Initially, a circular area of 100 sq arcmin about the simulated cluster center was considered for fitting with the radial

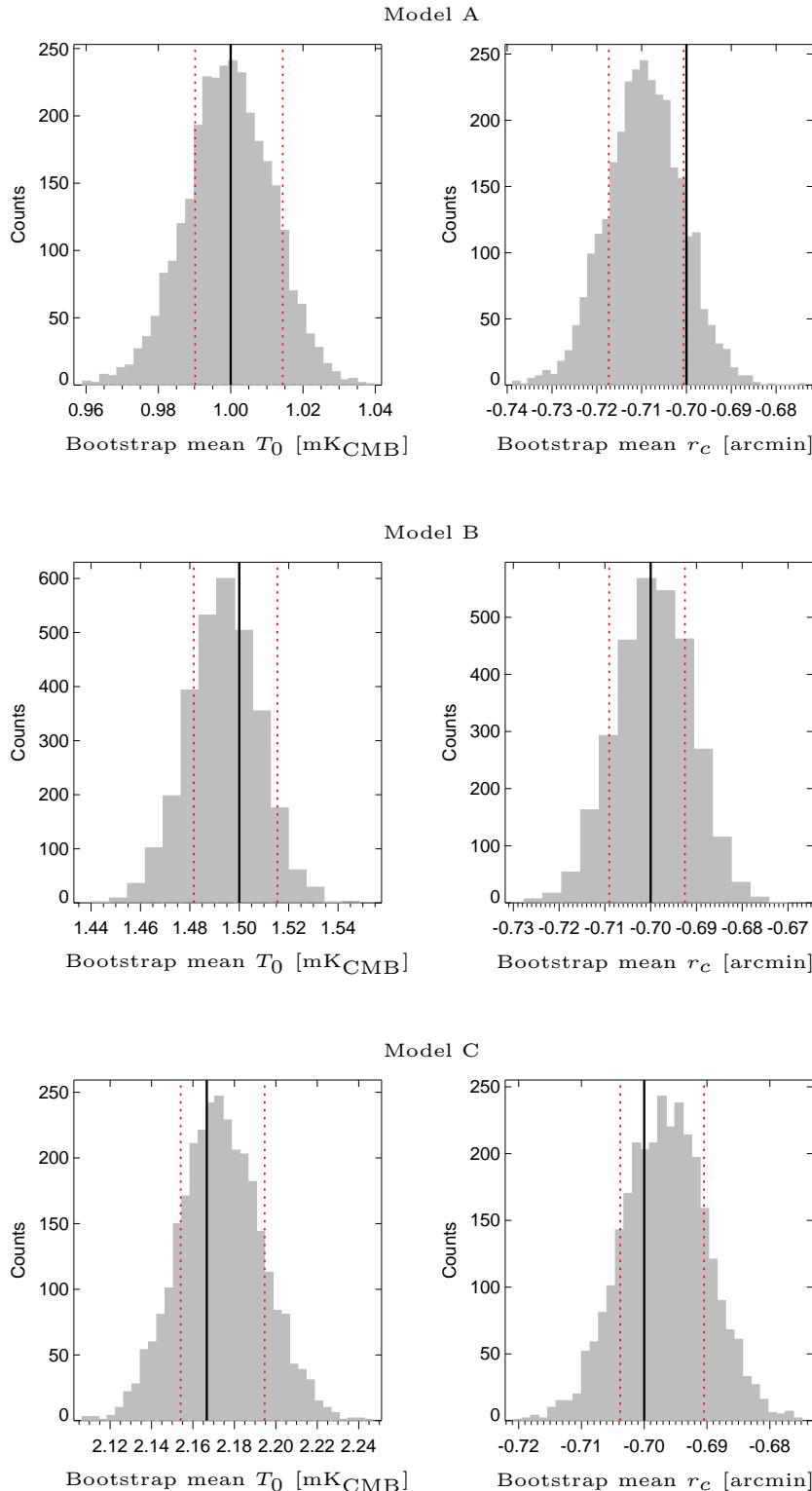


Figure 4.14: The bootstrapped distribution of the mean of the sample distributions (shown for each model in figures 4.13, 4.12, 4.11) of core radius and peak temperature decrement of the beta model obtained from the Bullet noise simulated mock data. The black solid line indicates the input values of the parameters, which is within the 68% confidence region (red dotted lines) on the average estimated value.

Table 4.3: Results on Y estimations from fitting multiple realizations of mock APEX-SZ maps.

Mock data	Area of map used	No. of realizations	input $Y \times 10^{-11}$ sr	average best-fit $Y \times 10^{-11}$ sr	68% confidence interval of the mean $\times 10^{-11}$ sr
Model A	100 arcmin sq.	25	43.5	42.8	[42.3, 48.0]
Model B	S	30	68.5	67.3	[66.2, 69.0]
Model C		38	98.3	97.6	[95.8, 98.8]
Model A	552 arcmin sq.	14	43.5	43.2	[41.7, 44.0]
Model B	L	30	68.5	67.6	[67.5, 71.2]
Model C		30	98.3	98.5	[97.5, 99.9]

Table 4.4: ΔT_0 parameter estimated from numerous mock realizations of APEX-SZ mock maps. The input value of ΔT_0 was $-0.7\text{mK}_{\text{CMB}}$.

Mock data	Area of map used	No. of realizations	average best-fit $\Delta T_0 \text{ mK}_{\text{CMB}}$	68% confidence interval of the mean
Model A	100 arcmin sq.	25	-0.71	[-0.70, -0.73]
Model B	S	30	-0.70	[-0.71, -0.68]
Model C		38	-0.71	[-0.70, -0.73]
Model A	552 arcmin sq.	14	-0.714	[-0.720, -0.695]
Model B	L	30	-0.706	[-0.714, -0.698]
Model C		30	-0.707	[-0.714, -0.699]

Table 4.5: r_c parameter estimated from numerous mock realizations of APEX-SZ mock maps.

Mock data	Area of map used	No. of realizations	input r_c [arcmin]	average best-fit r_c [arcmin]	68% confidence interval of the mean [arcmin]
Model A	100 arcmin sq.	25	1.0	0.98	[0.96, 0.99]
Model B	S	30	1.5	1.47	[1.45, 1.49]
Model C		38	2.17	2.13	[2.10, 2.17]
Model A	552 arcmin sq.	14	1.0	0.97	[0.96, 1.0]
Model B	L	30	1.5	1.48	[1.47, 1.50]
Model C		30	2.17	2.14	[2.11, 2.16]

binning method. This is the same area considered for testing the Fourier domain method in the sense as shown in Figure 4.15. One of the main purposes of using this method is indeed to use as much the information as possible to fit models to data. Hence, the procedure was tested also within a larger region with a circle radius of approximately $13'$ (circular region of area 552 sq. arcmin).

The number of mock images that were fit for each model is presented in Table 4.3. From a sample distribution of recovered modes for ΔT_0 , r_c and Y the sample means are quoted in the tables. In the final column, the confidence interval range of the mean determined by the bootstrap procedure is given.

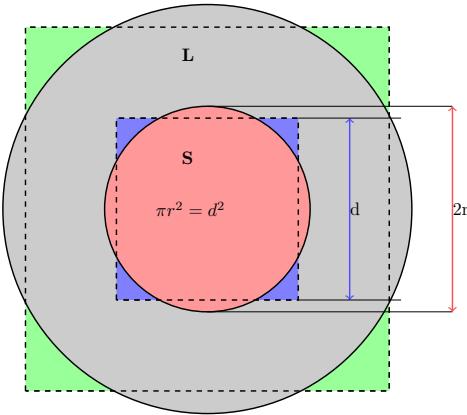


Figure 4.15: Schematic representation of equal areas of circular and square regions. The smaller radius circle and the smaller length box represent a smaller region typically used in the fitting with the FD method. The larger circle and square represent the region of data that is still available for fitting. The smaller area is referred to by the symbol **S**, and the larger area is represented with the symbol **L** for simplicity.

4.3.3 Comparison of the two methods

In this Section, we compare the performance of the two methods considered. First, the comparison is done with white noise mock data. This acts as a control for determining the performance differences in fitting for mock data with realistic noise. In the white noise case, both methods are expected to give unbiased and near similar results. This is due to the stationary noise assumption for the FD method being upheld in this case and for the radial binning method this is simply a white noise with a constant rms, which should boost the signal to noise in annular bins as expected. In Section 4.3.3.1, this discussion is provided with the help of few examples. In Section 4.3.3.2, a similar comparison is done with few examples with real noise but the image size used for fitting is kept limited to 100 arcmin square in this case to approximately uphold the noise stationary condition. Later in Section 4.3.3.2, the choice of the area of the image used for fitting is relaxed to have 552 arcmin square. The comparison of both methods on fitting within such larger region is given and this forms the basis for the final conclusion for choosing an optimal method for the fitting procedure to be applied to the entire cluster sample in Chapter 5.

4.3.3.1 White noise mock maps

Models 1 & 2 listed in Table 4.1 have been convolved with the point spread function of the telescope that is modelled as a symmetric Gaussian beam with σ of 25.5'' (i.e. FWHM of 60''). Using these mock images to fit the models offers a control for the comparison of fitting procedures. The third model represented in the table has been convolved with the PST of Bullet data. This transfer function is not in practice strictly symmetric. The noise added to these models were white noise with no pixel-to-pixel correlation.

For a qualitative comparison, one mock image corresponding to each of the two models were fit using Fourier domain method within a box of size 16.8' by 16.8' and radial binning method applied to the same area of circular radius 9.5'. This was designed to keep the area of the data used in both

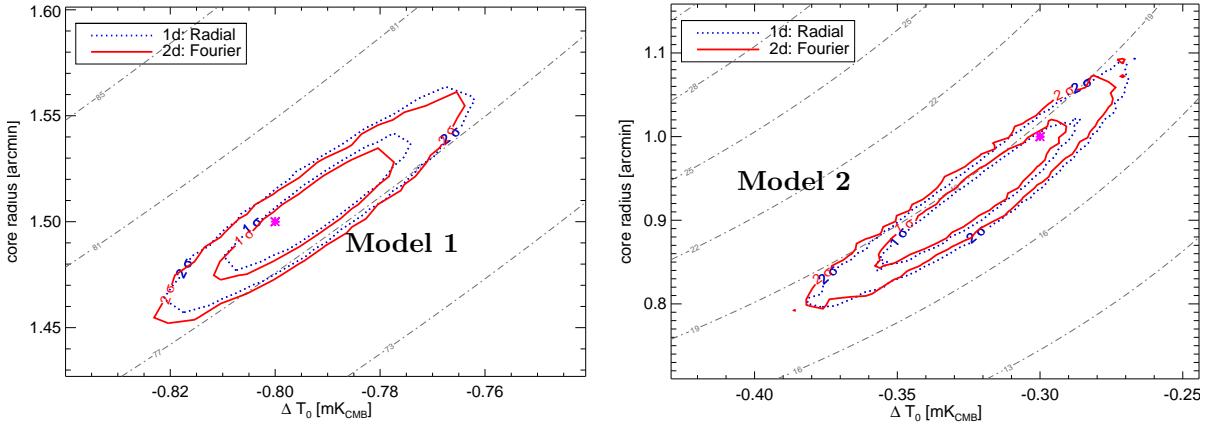


Figure 4.16: Comparison of the FD and the RB methods using Model 1, and 2 mock data. Results from one mock fit from each model are shown. The contours show the 68% and 95% confidence regions of the parameters from one mock fit. The dashed curves represent the iso-Compton- Y regions (the parameter space which give the same value for Y). The magenta point indicates the input value for the simulated model.

methods same. In both cases, ΔT_0 and r_c were fit freely by keeping the centroid of the models fixed. The resulting constraints on the parameters are shown in Figure 4.16. Both methods recover almost identical parameter space. The alignment of the contours trace the iso-integrated Compton parameter curve.

Now, relaxing the symmetry of the simulated filtered models, one mock image for Model 3 from Table 4.1 is fit by leaving ΔT_0 and r_c free. The result is shown in Figure 4.17. Both methods recover similar parameter space and the alignment of the constraints are along the iso-integrated Y curves. The uncertainties have increased due to the strong attenuation of the source by applying the PST.

In conclusion, when the images have uncorrelated noise, both methods perform in very similar fashion.

4.3.3.2 Realistic mock maps

It is instructive to compare the performance of the radial binning method to the Fourier method with real noise. It is done by first taking a small region of the maps to fit for models, and second, the fitting region is increased to observe the performance of each method.

Fitting inside small region of the maps

First, we consider both methods applied on same set of mock data within the smaller region of the maps such that both methods use the same amount of information. The information used by both methods only differ slightly at outer regions, due the choices of the shape of the fitting region (See Figure 4.15). Following the procedure described in the early part of Section 4.2.4.2, numerous mock images were fit with β models with centres fixed. For comparison, the β models were also fit to the data using the method prescribed in 4.2.4.1 with fixed centres and same area of the map. Three examples of recovered posterior from fitting mock data, one of each β model with different core

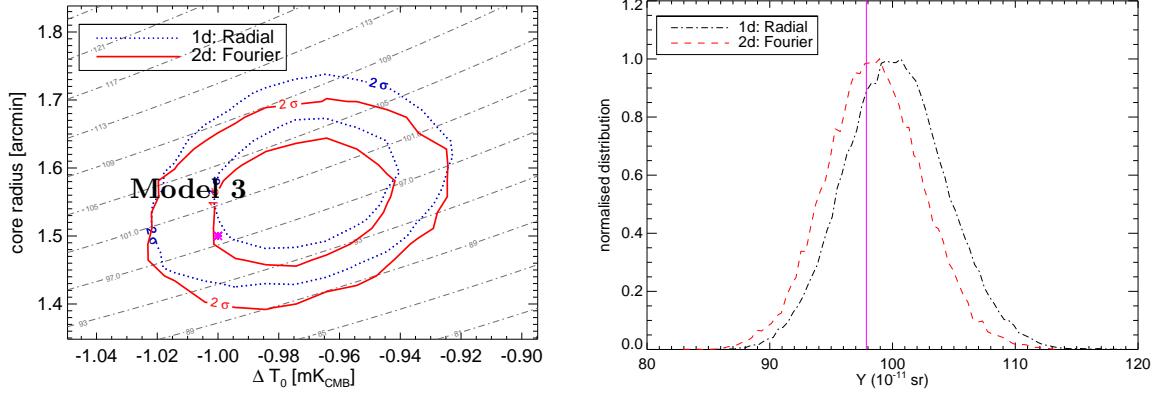


Figure 4.17: *Left:* The parameter covariance for one example Model 3 mock image. The parameter constraints from the radial binning method (blue), and the Fourier domain method (red) are shown. The parameters recovered by both methods are consistent. The input value used for the simulation is shown by the magenta point. The iso-Compton-Y regions in the parameters space is shown as the dotted-dashed lines. *Right:* The marginalised posterior distribution of the Compton-Y from both the methods for the parameter covariance in the left figure are shown here. Not surprisingly, the posterior distribution from both the methods are quite similar.

radius used to generate the mock data, are shown in Figure 4.18. The solid contours are the recovered confidence regions for both methods for the same data. In the figure, the dashed contours indicate the iso-Compton-Y regimes for the β model. The posterior distribution of the integrated Compton parameter is also shown for each example from both methods. In the Compton-Y, little variation from different methods could be observed. However, the confidence regions in the two dimensional parameter space indicate some differences. The cause of such differences can be attributed only to the noise information used by the methods in the form of a power spectral density or bin-to-bin covariance. To demonstrate this, the mock data for Model C was re-fit using the radial binning method. But this time, the off-diagonal elements of the bin-to-bin noise covariance matrix were ignored. The resulting constraints on the parameters are shown in Figure 4.19. It can be seen that the confidence region from the radial binning method in such an instance has shrunk in comparison to using the full noise matrix and they seem to align in a similar fashion as the contours from Fourier domain fitting. This indicates that the noise covariance in the Fourier domain fitting may not fully capture the noise information due to the assumption of a diagonal noise matrix. This is a hint of a slight under-performance of the Fourier domain method for the APEX-SZ cluster data in reliably extracting the parameters of interest.

A higher number of mock data sets ($\sim 10\text{--}20$) for each of the three Models A, B, C were fit in Fourier domain method and radial binning method for a statistical comparison of the recovered parameters. From the posterior of each mock fit, the modes of the parameters were sampled. The sample distribution of recovered modes from the radial binning method and Fourier domain method are compared. The integrated Compton parameter (Y) was also sampled from the posteriors. In figure 4.20, the distribution of the best fit values of Y are plotted for both methods. In Figure 4.21, the scatter plot of the best fit values of ΔT_0 and r_c from both methods are shown.

It can be concluded from looking at Figure 4.20 that both the methods recover similar or at least consistent values for the Y parameter. On similar lines, Figure 4.21, from the scatter plots, there is

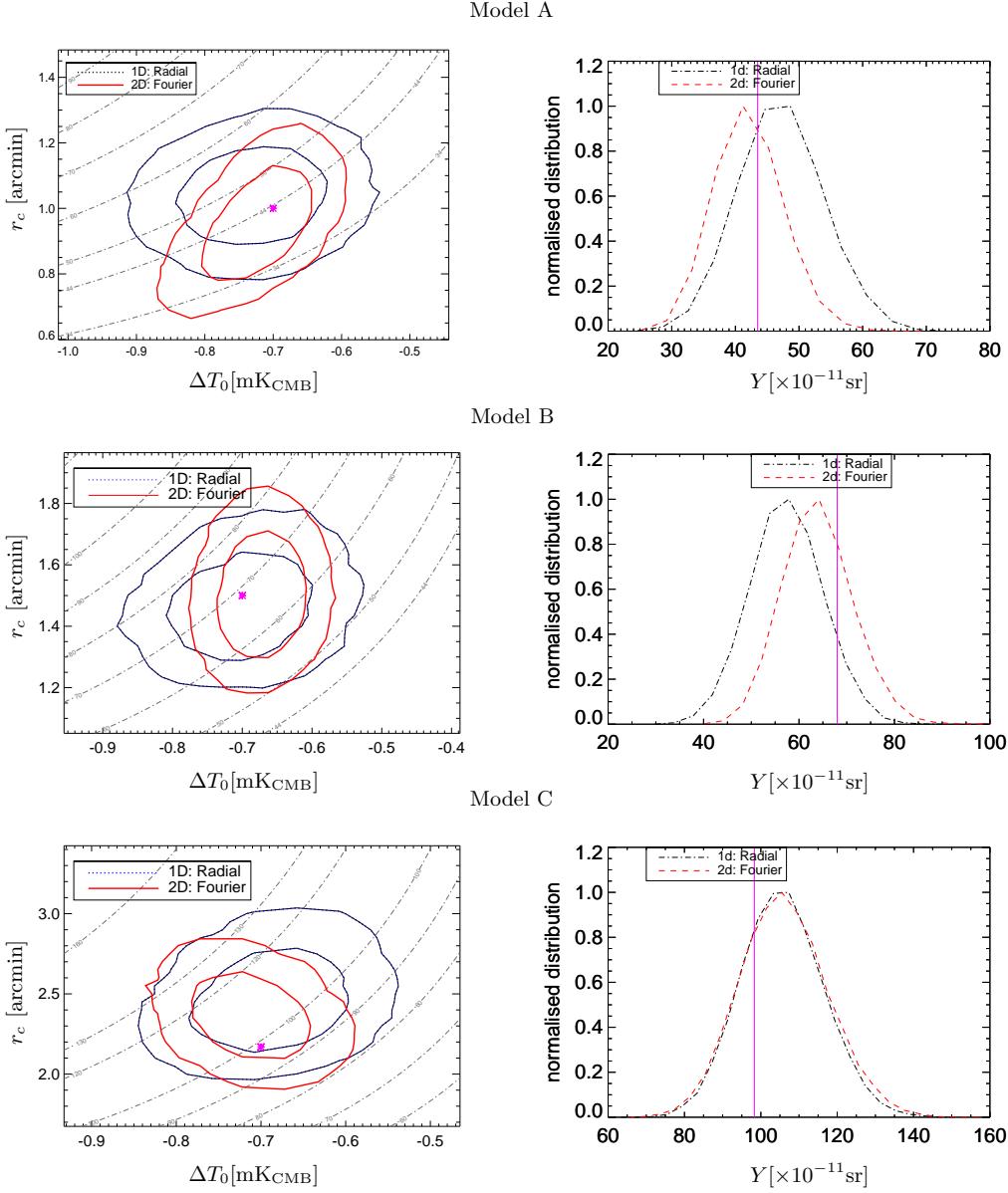


Figure 4.18: *Left:* Example parameter constraints from mock data of each input model A, B and C. The input values of the parameters, used for the mock data, are indicated by the magenta point. The 68% and 95% confidence regions of the constraints obtained by using the Fourier domain method (red), and radial binning method (blue) are shown. *Right:* The obtained constraints on the Compton-Y for each example shown on the left is shown. The black curve shows the marginalised posterior distribution of the Y obtained from radial binning method. The red curve correspond to the same for the Fourier domain fit. The magenta line indicates the Y value corresponding to the input model for the mock data. The parameter constraints recovered by each of the two methods from fitting mock images with real noise prefer slightly different orientations. The iso-Compton-Y in the parameter space are indicated by dashed lines. The area of the map used by two methods are kept the same. The input values used for the simulation are shown by the magenta point.

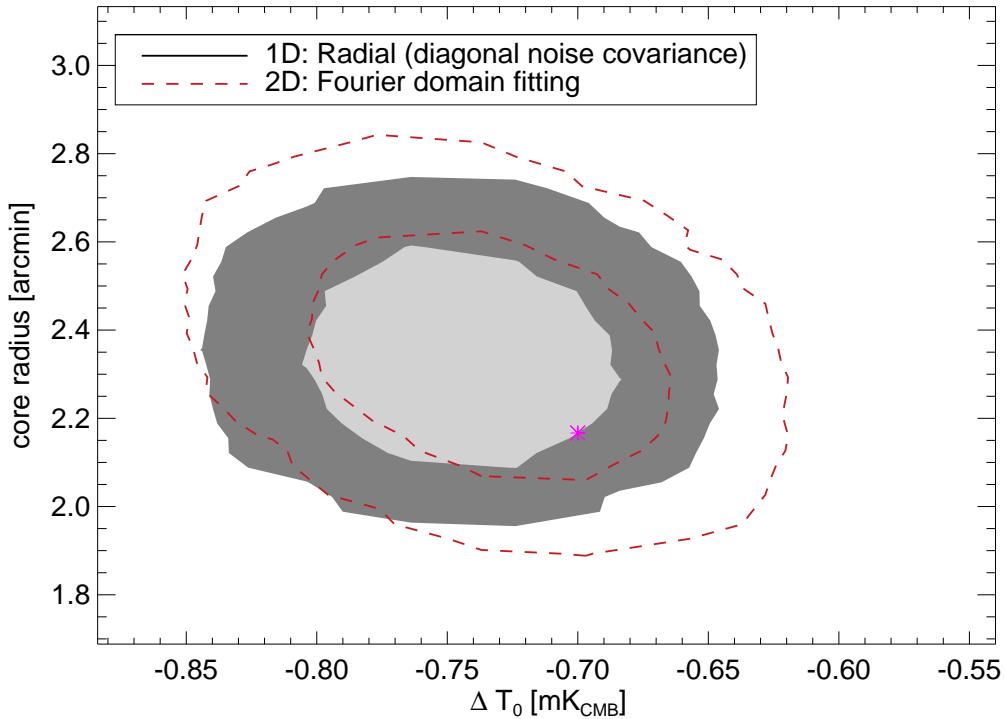


Figure 4.19: The recovered parameter covariance from the two fitting methods. The noise covariance matrix of the radial binning method included only diagonal elements, the off-diagonal elements were purposely ignored.

no evidence of any quantitative and qualitative difference in the preferred parameter space traced by the best-fit from both methods. So, within the $10' \times 10'$ region of data, the confidence levels obtained on a single mock data trace slightly different confidence regions. There is an indication that the Fourier domain fitting may be underestimating the confidence regions with respect to the radial binning. However, using a quantitative analysis, in terms of best fit values, no difference could be found between the two methods.

Fitting inside a larger region of the maps

In this section, the size of the region within which to fit the data was relaxed. The region of area 552 sq arcmin about the center of the map was used for testing and analysing the performance of both methods. This region of large area is denoted by the symbol **L**. Hereafter, the smaller region of 100 arcmin sq will be referred to as **S**. It is clear that within such an area corresponding to $24' \times 24'$ box size, the noise can no longer be considered stationary. However, for the sake of completeness of the discussion, the Fourier domain fitting method within such a large area of the map was used on the mock data as well. This was a check on how severe this failure of stationarity within this region is. The same mock simulated image data for the models A, B, C were used.

For each of these models, at least 10 mock images were individually fit with both methods within the larger region. To be able to further compare with the results from previous section, the same mock images were fit within the smaller area of 100 sq. arcmin. First, focusing on individual constraints

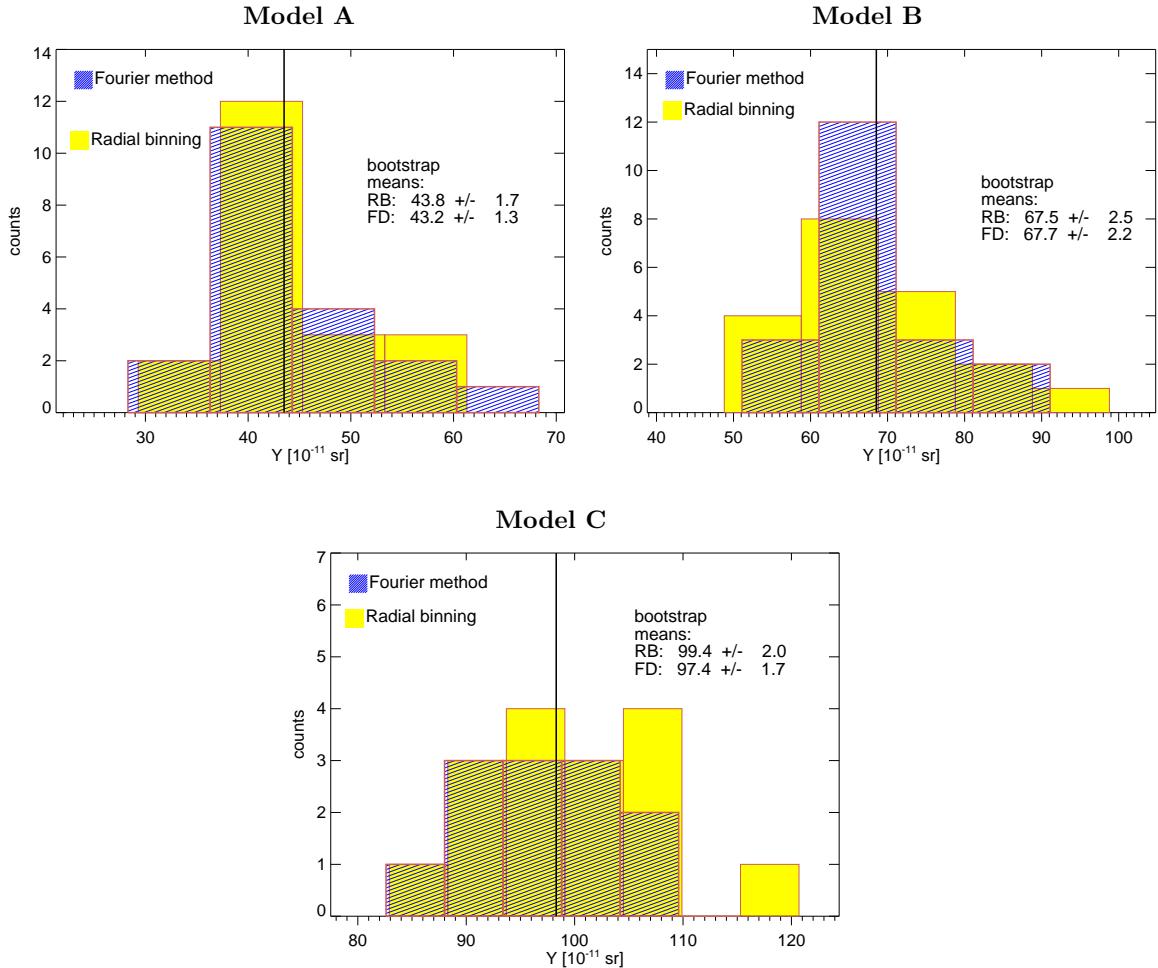


Figure 4.20: The figures A, B, C show the distribution of the best fit values for Compton-Y parameter from the two methods. Integrated Compton parameter Y for the sources with the three different input parameters given in table 4.1. The yellow shaded region is the distribution from the radial binning g method and the histogram filled with blue lines is the distribution recovered from the Fourier domain fitting method. The same set of sources are used for both cases. In all of the cases there seems to be no bias in recovering the Y parameter. The input value of the Y parameter is shown by the black solid line.

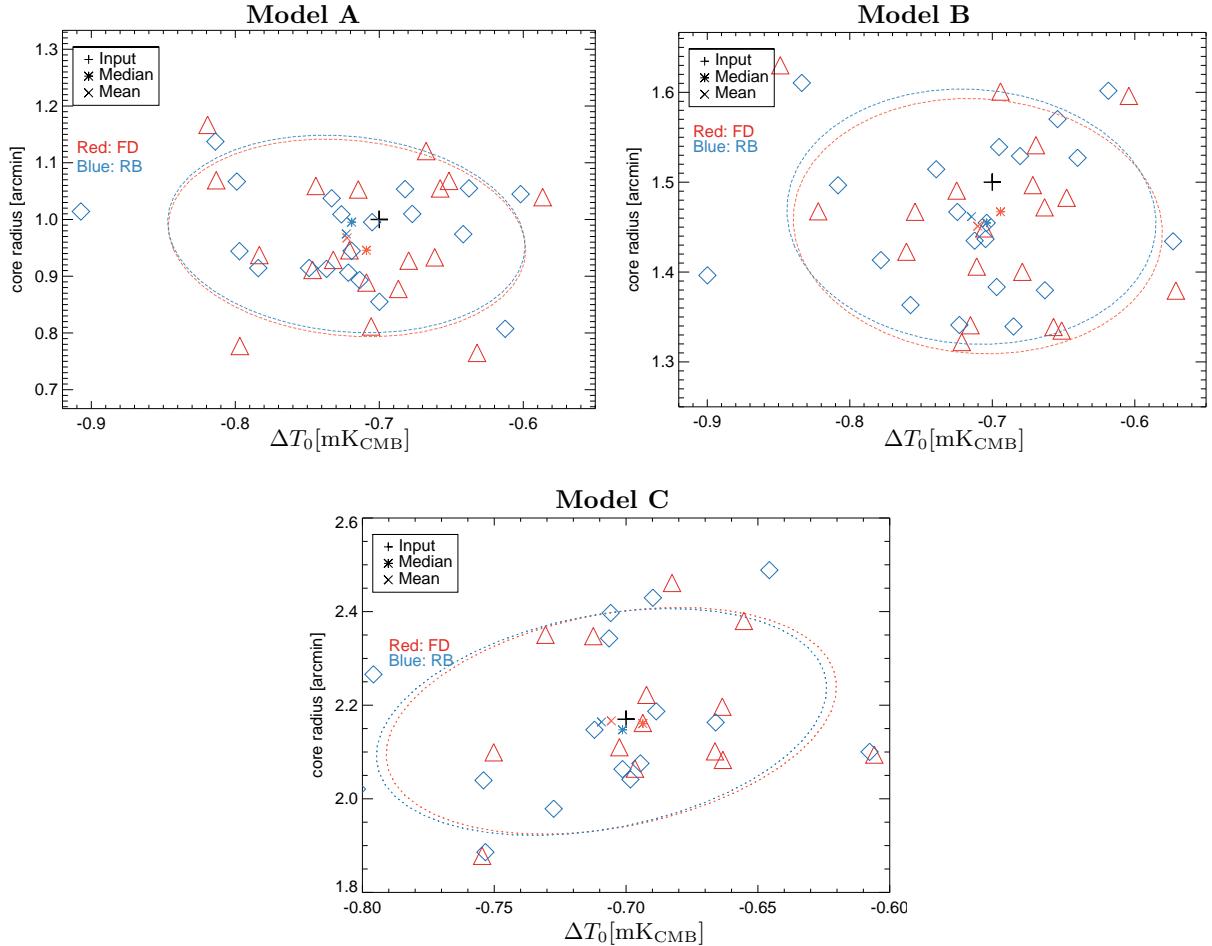


Figure 4.21: The figures **A**, **B** and **C** show the scatter plot of the recovered best fit values of the parameters (core radius and the ΔT_0) of the β model. The input values are indicated by the black cross point. The blue diamond points are the best fit values from the radial binning method and the red points are the best fit values from the Fourier domain method. The ellipses represent the 68% confidence level covariance. These scatter plots demonstrate that there is no qualitative and quantitative difference in the parameter space traced by best-fit values from both the methods.

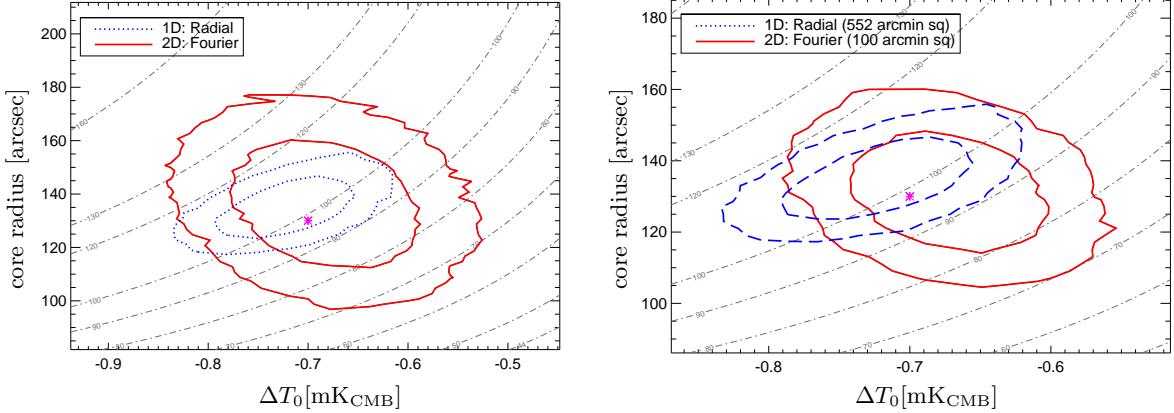


Figure 4.22: Parameter covariance of the β model from the two fitting methods and two areas of the map used to fit. *Left:* Radial binning method (blue) and Fourier domain (red) fitting applied within 552 square arcmin region. *Right:* Parameter covariance for the same example shown in left figure obtained from the radial binning method applied within 552 square arcmin region (blue) and the Fourier domain fitting applied to 100 square arcmin region (red). The Fourier domain fitting within larger area estimates larger covariance of the parameters than within a smaller area, which is counter to the expectation that the constraints should improve when adding more information. Whereas, the radial binning method when applied to this larger area of the image data constrains the parameter space more precisely than either method applied within a smaller region of the image data.

from one mock fit, Figure 4.22 shows one set of such constraints. From the confidence levels obtained from the Fourier domain fitting in region **L**, it is clear that the uncertainties on the parameters have increased in comparison to the ones obtained within a region **S**. The radial binning constraints have shrunk in comparison, which is expected given that more information has been added by the larger area considered. The failure of the Fourier domain method in such larger region is evident from one look at the power spectral density shown earlier in Figure 4.5. These diagonal elements are inversely proportional to the weights for the χ^2 . The increase in the noise covariance in region **L** is due to the non-directional aspect of the PSD. Now, from the few mock data that were fit, the measurement uncertainties on the Y parameter was considered from all four flavours of the fitting (two regions, two methods). The distribution of the uncertainties alone obtained for the integrated aperture Y parameter from the sample of mock data are shown in Figure 4.23. The fit results within the **S** region from either of the two methods show reasonable overlap in the measurement uncertainties. There is stark disagreement in the recovered measurement uncertainties in the fit using the larger area, **L**.

4.4 Conclusion & discussion

Two candidate methods were considered for estimating the global SZ signal from APEX-SZ maps. By using mock data to model the filtered APEX-SZ maps, both the methods were studied.

The results from the analysis presented in this chapter is conclusive that the radial binning method should be the preferred choice for measuring integrated Y from APEX-SZ when increasing the size of the map to fit. Even though it requires priors on cluster centre, and assume some level of azimuthal symmetry for models, it does not limit the extent of the map size that needs to be used for the fit. In

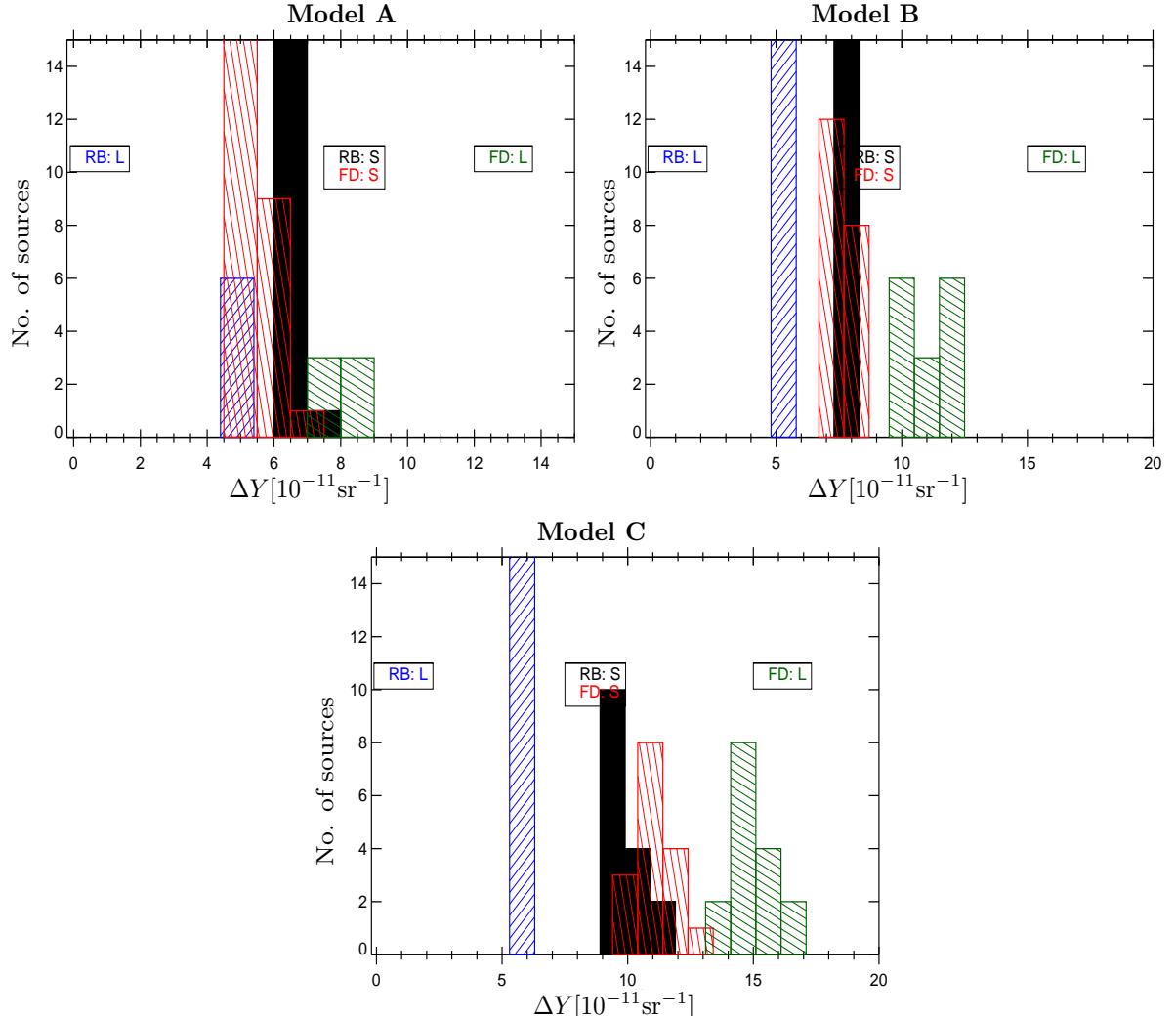


Figure 4.23: Distribution of uncertainties on the integrated Compton parameter recovered from fitting to multiple mock realizations of filtered maps. Four different ways of fitting was used. Blue histograms show the uncertainties from radial binning (RB) method applied in large area of map L. Green histograms are obtained for the Fourier doamin (FD) fitting in large area. Red histograms are obtained for FD fitting within smaller region S. Black histograms represent the results from RB method applied in samll region. On increasing the size of the area of the map used for the fitting, the uncertainties estimated from the two methods show a disagreement. The uncertainties decreases with the area of the map used for the fitting in the radial binning approach. The uncertainties are estimated to be larger in the FD method approach due to the failure of non-stationarity condition in the larger area of the map.

the Fourier domain method, the limitation of determining the size of the map upto which the noise can be considered stationary is still undesirable. Determining the optimal map size for over 40 cluster data would be required, as the integration time for each of the target varies considerably. This is a time intensive thing to pursue, given that for each cluster data numerous mock data would have to be fit using the procedure to determine the optimal size. In addition, although no bias was found in recovering the parameters using the Fourier domain fitting method when best-fit values from multiple mock fit were considered, the confidence regions of the parameters from individual fit still showed some subtle differences (when small region of fit were considered) due to strong assumption of noise stationarity. On the other hand, by simply forgoing the details on asymmetry, radial binning method provides a reliable estimate. The cluster to cluster differences arising from asymmetry etc. can be sufficiently dealt with in a statistical manner for the sample when measuring the scaling relation and this is dealt with in Chapter 6. Therefore, for measuring the global SZ signal in clusters from APEX-SZ, in the following chapter, the radial binning method will be employed.

CHAPTER 5

Results: measuring integrated Comptonization from APEX-SZ

Overview

Note: This chapter is adapted and re-produced with modifications from a recently submitted paper to the journal Monthly Notices of the Royal Academic Society titled “Weak-lensing mass calibration of the Sunyaev-Zel’dovich effect using APEX-SZ clusters” authored by A. Nagarajan, F. Pacaud, M. Sommer, M. Klein, K. Basu, F. Bertoldi, A. T. Lee, P. A. R. Ade, A. N. Bender, D. Ferrusca, N. W. Halverson, C. Horellou, B. R. Johnson, J. Kennedy, R. Kneissl, K. M. Menten, C.L. Reichardt, C. Tucker, B. Westbrook

In Chapter 4, it was concluded that the radial binning method would be a preferred choice for fitting models to APEX-SZ data as they tend to be more reliable for measuring integrated Comptonization (Y or Y_{SZ}). In this chapter, the results from fitting models to the APEX-SZ data are presented. A description of all the modelling assumptions and details of the method applied are discussed. The results are compared to the measurements from Bender et al. (2016). The measured values from this chapter is found to be statistically consistent with Bender et al. (2016) across a sample of 41 clusters. These measurements form the basis for studying the scaling relations with the APEX-SZ sample of galaxy clusters which are detailed in Chapter 6.

5.1 Generalised Navarro-Frenk-White profile fit to APEX-SZ

To model the pressure of the ICM we use the generalized Navarro-Frenk-White (gNFW) profile shown in Section 4.2.1 which is motivated by dark matter halo profiles found from simulations (Nagai, A. V. Kravtsov and Vikhlinin, 2007).

The peak signal-to-noise ratios of the APEX-SZ detections in beam smoothed maps range from ~ 16 down to non-detections. Due to scan pattern and high pass filtering, scales larger than $\sim 10'$ are not recovered in the filtered data. Due to the degeneracy in the c_{500} and R_{500} of the gNFW model and the limitations set by the data, the weak-lensing estimate of the spherical over-density radius was used. Furthermore, for the slope parameters of the gNFW profile the values were fixed to those

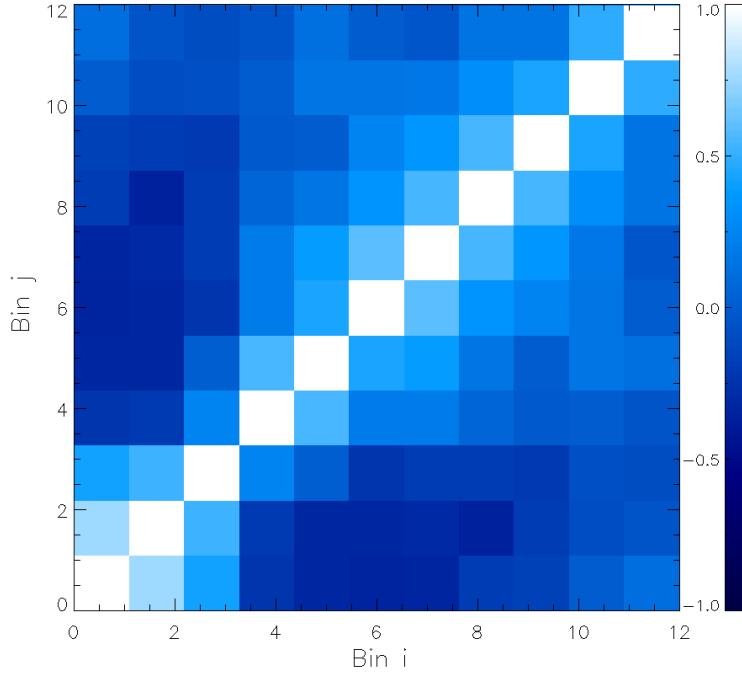


Figure 5.1: Bin-to-bin correlation matrix for the Bullet cluster field. The bin-width is $1'$.

determined in Arnaud et al. (2010), with $\{c_{500}, \alpha, \beta, \gamma\} = \{1.177, 1.0510, 5.4905, 0.3081\}$. Thus, the normalization of the gNFW profile of each cluster is the only free parameter.

In order to be consistent with the weak-lensing analysis, the centroids were fixed to the BCG centres used for the weak-lensing estimation of mass. Based on the assumed parameterisation of the gNFW profile, $Y_{\text{sph},500}$ is obtained by dividing the $Y_{\text{cyl},500}$ by a factor of 1.203.

To fit the gNFW model to the data, the data were binned about the BCG centroids used in the weak-lensing analysis, considering all data within a radius of $12'$, using a bin width of $1'$ (corresponding to the FWHM of the APEX-SZ beam) and taking pixel weights into account in the averaging. The model image is convolved with the point source transfer function to account for the finite resolution and filtering effects, as discussed in section 3.4.3, and binned in the same way as the data.

For each target, the bin-to-bin noise covariance matrix was computed from 100 noise realizations, produced by randomly inverting half the data. Because noise realizations produced in this way do not account for noise covariance components produced by astronomical signals, random realizations of primary CMB anisotropies using the Planck Collaboration and et al. (2016a) best fit CMB power spectrum were generated, these were convolved with the transfer function and the filtered CMB realizations were added to the instrument noise realisations. The noise contributions from unresolved point sources emitting synchrotron and dust emission at 150 GHz can be neglected for the APEX-SZ noise levels (Reichardt, Zahn et al., 2009). The final noise images were radially binned. The ensemble of noise realisations in each radial bin was used to compute the full bin-to-bin covariance matrix. An example correlation matrix is illustrated for the Bullet cluster in Figure 5.1. Neighbouring bins are strongly correlated due to the telescope resolution, while intermediately separated radial bins are

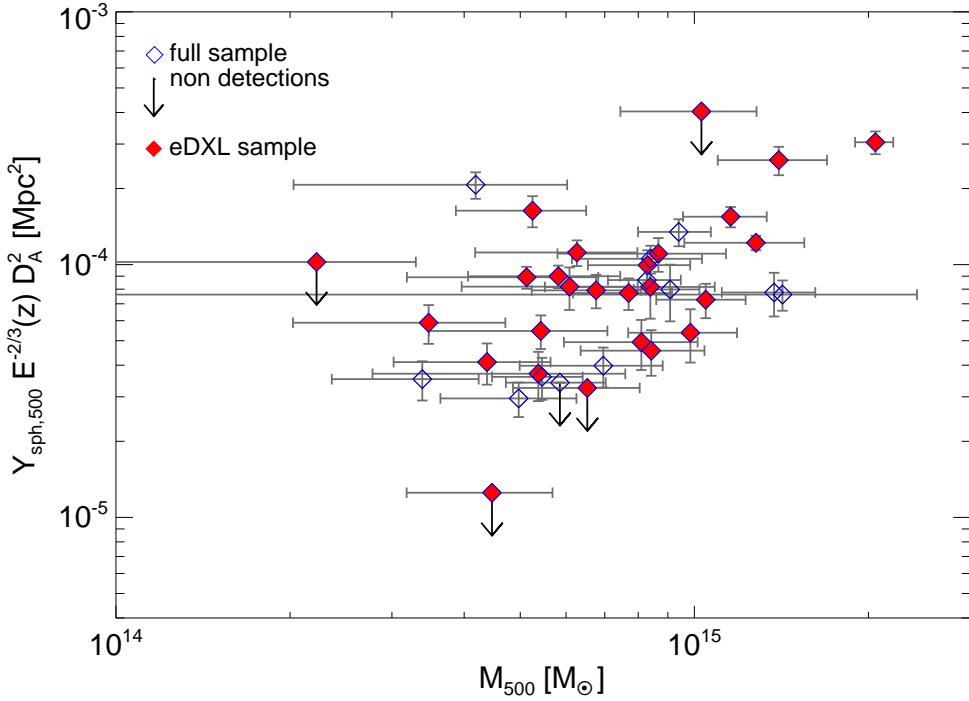


Figure 5.2: $Y_{\text{sph},500}$ measurements vs. the weak-lensing spherical mass estimates within $R_{500,\text{WL}}$. The non-detections in the integrated Compton parameter are indicated by upper limits at a 2σ level (downward arrows).

anti-correlated due to the low-order polynomial filtering applied to attenuate low-frequency noise.

The χ^2 statistic was used to define the likelihood $\mathcal{L} \propto \exp(-\chi^2/2)$ as given in Equation (4.2.3). The fitting was done using an Markov chain Monte Carlo analysis (MCMC) to estimate the confidence levels of the normalisation parameter. The $Y_{\text{SZ},500}$ for each cluster is computed using the formulation given in previous section for the recovered models. In the following section we use the MCMC approach to estimate the correlation in the measurements of $Y_{\text{SZ},500}$ and $M_{\text{WL},500}$. The measured $Y_{\text{SZ},500}$ for the full set of 39 APEX-SZ clusters are given in Table 5.1.

5.1.1 Propagation of uncertainties in R_{500} into the SZ modelling

Since by definition $M_{\text{WL},500}$ is proportional to R_{500}^3 (Equation 2.1.2), the obtained $Y_{\text{SZ},500}$ values are expected to be correlated with the weak-lensing masses because the same apertures were used for measuring both quantities.

This correlation was estimated by using the MCMC and re-fitting the gNFW profile with R_s and p_0 as free parameters. The prior on R_s was taken from the weak-lensing estimates of the R_{500} distribution via the relation given by Equation (4.5). The uncertainties in R_{500} were propagated as a prior that is a two-sided Gaussian distribution into the modelling of the SZ signal. The correlation in the measurements is determined using a Pearson correlation coefficient from the recovered distribution of $Y_{\text{SZ},500}$, $M_{\text{WL},500}$.

Table 5.1: The 27 targets of the eDXL sample. Each cluster position is taken to be the position of the BCG in the optical data. The integrated Compton-y parameters ($Y_{\text{SZ},500}$), the weak-lensing derived aperture (R_{500}) are reported.

Name	RA	Dec	redshift z	$Y_{\text{SZ},500}$ $[10^{-5} \text{ Mpc}^2]$	R_{500} [Mpc]
eDXL clusters					
A2204	16:32:46.9	+05:34:32.3	0.152	13.05 ± 0.86	$2.43^{+0.81}_{-0.81}$
RXCJ2014.8 – 2430	20:14:51.7	-24:30:22.3	0.160	0.97 ± 1.22	$2.47^{+0.82}_{-0.82}$
A1689	13:11:29.5	-01:20:27.9	0.183	32.4 ± 3.36	$2.15^{+0.72}_{-0.72}$
A2163	16:15:49.0	-06:08:41.5	0.203	17.56 ± 2.47	$2.22^{+0.74}_{-0.74}$
RXJ1504	15:04:07.5	-02:48:16.5	0.215	8.45 ± 1.66	$2.40^{+0.79}_{-0.79}$
RXCJ0532.9 – 3701	05:32:55.7	-37:01:36.0	0.275	8.68 ± 1.31	$2.26^{+0.75}_{-0.75}$
RXCJ0019.0 – 2026	00:19:08.0	-20:26:28.0	0.277	8.47 ± 1.22	$2.22^{+0.74}_{-0.74}$
RXCJ2337.6 + 0016	23:37:39.7	+00:16:17.2	0.278	8.04 ± 1.25	$2.26^{+0.75}_{-0.75}$
RXCJ0232.2 – 4420	02:32:18.6	-44:20:48.0	0.284	9.82 ± 0.98	$2.33^{+0.77}_{-0.77}$
RXCJ0437.1 + 0043	04:37:09.5	+00:43:52.1	0.284	5.44 ± 1.22	$2.22^{+0.74}_{-0.74}$
RXCJ0528.9 – 3927	05:28:53.0	-39:28:17.8	0.284	4.53 ± 0.84	$2.36^{+0.79}_{-0.79}$
RXCJ2151.0 – 0736	21:51:00.8	-07:36:31.0	0.284	3.90 ± 0.86	$2.50^{+0.83}_{-0.83}$
A2813	00:43:25.1	-20:37:01.2	0.292	10.99 ± 1.60	$2.22^{+0.74}_{-0.74}$
RXCJ0516.6 – 5430	05:16:37.6	-54:30:38.1	0.295	5.05 ± 1.03	$2.19^{+0.73}_{-0.73}$
Bullet	06:58:36.4	-55:57:19.2	0.297	12.35 ± 1.44	$2.25^{+0.75}_{-0.75}$
A2537	23:08:22.2	-02:11:31.6	0.297	6.06 ± 1.44	$2.15^{+0.72}_{-0.72}$
RXCJ0245.4 – 5302	02:45:31.3	-53:02:07.8	0.302	6.53 ± 1.14	$2.39^{+0.79}_{-0.79}$
RXCJ1135.6 – 2019	11:35:21.4	-20:19:56.6	0.305	-0.64 ± 1.02	$2.33^{+0.77}_{-0.77}$
A2744	00:14:18.5	-30:22:51.2	0.307	17.19 ± 1.59	$2.12^{+0.71}_{-0.71}$
A1300	11:31:54.2	-19:55:39.8	0.308	9.99 ± 1.02	$2.26^{+0.75}_{-0.75}$
MACSJ1115.8 + 0129	11:15:52.0	+01:29:55.0	0.348	6.18 ± 0.94	$2.22^{+0.74}_{-0.74}$
RXCJ2248.7 – 4431	22:48:44.0	-44:31:51.0	0.348	6.44 ± 20.42	$2.12^{+0.71}_{-0.71}$
RXCJ1206.2 – 0848	12:06:12.1	-08:48:03.4	0.441	12.91 ± 1.97	$2.04^{+0.68}_{-0.68}$
RXCJ2243.3 – 0935	22:43:22.8	-09:35:22.0	0.447	9.78 ± 2.46	$2.01^{+0.67}_{-0.67}$
RXJ1347 – 1145	13:47:30.6	-11:45:09.5	0.451	30.4 ± 3.92	$1.94^{+0.64}_{-0.64}$
RXCJ2214.9 – 1359	22:14:57.2	-14:00:12.3	0.483	14.53 ± 1.78	$1.97^{+0.65}_{-0.65}$
MS0451.6 – 0305	04:54:10.8	-03:00:51.4	0.539	9.98 ± 1.91	$2.01^{+0.67}_{-0.67}$

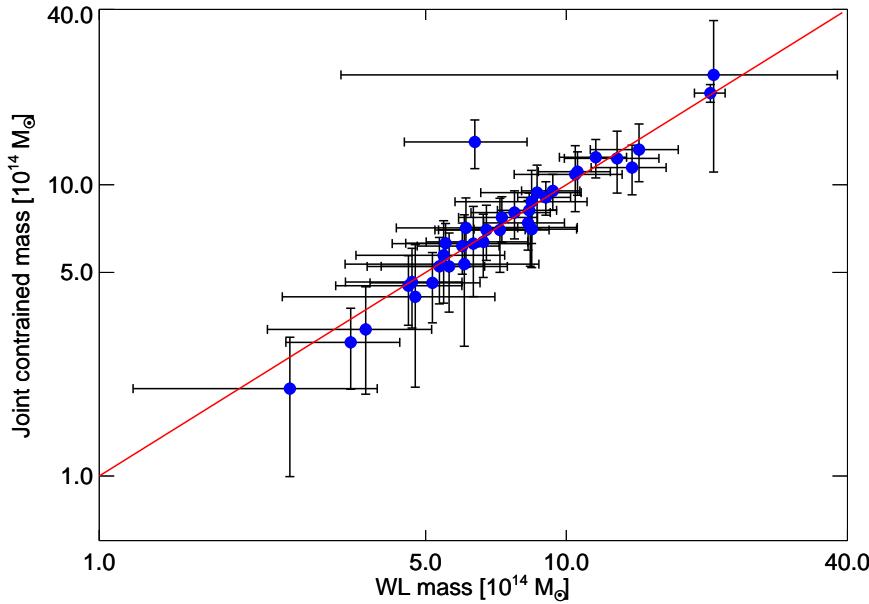


Figure 5.3: Mass estimate from joint fitting analysis vs. the weak-lensing estimate used as the prior. The joint fitting yields generally consistent mass estimate with the input prior values for all clusters with the exception for one cluster, Bullet. Bullet is observed with highest signal to noise in the APEX-SZ maps, and given that this is highly disturbed cluster there is some inconsistency in the input prior mass and the preferred estimate from joint analysis for this cluster.

Table 5.2: Same as table 5.1, but for other clusters in the full APEX-SZ sample that are not part of the X-ray selected complete sample.

Name	RA	Dec	redshift z	$Y_{\text{SZ},500}$ [10^{-5} Mpc^2]	R_{500} [Mpc]
Other clusters					
A907	09:58:22.0	-11:03:50.2	0.153	3.71 ± 0.66	$2.57^{+0.85}_{-0.85}$
A3404	06:45:29.5	-54:13:37.1	0.167	11.12 ± 1.40	$2.36^{+0.79}_{-0.79}$
A383	02:48:03.4	-03:31:45.1	0.187	1.80 ± 0.91	$2.39^{+0.79}_{-0.79}$
A520	04:54:13.7	+02:56:10.2	0.199	3.84 ± 0.73	$2.39^{+0.79}_{-0.79}$
A209	01:31:52.5	-13:36:40.7	0.206	8.55 ± 2.16	$2.29^{+0.76}_{-0.76}$
A2390	21:53:36.8	+17:41:43.7	0.228	4.91 ± 3.18	$2.25^{+0.75}_{-0.75}$
A1835	14:01:02.1	+02:52:42.6	0.253	24.25 ± 2.92	$2.19^{+0.72}_{-0.71}$
RXCJ1023.6 + 0411	10:23:39.2	+04:10:58.0	0.280	9.59 ± 1.16	$2.22^{+0.75}_{-0.74}$
XLSSC-006	02:21:45.2	-03:46:02.7	0.429	3.44 ± 0.54	$2.15^{+0.72}_{-0.71}$
MACSJ1359.2 – 1929	13:59:10.3	-19:29:24.7	0.447	2.95 ± 0.96	$2.19^{+0.73}_{-0.72}$
MACSJ1311.0 – 0311	13:11:01.8	-03:10:39.7	0.494	4.75 ± 0.86	$2.04^{+0.68}_{-0.68}$
MS1054.4 – 0321	10:56:60.0	-03:37:36.2	0.831	10.39 ± 1.42	$1.59^{+0.53}_{-0.53}$

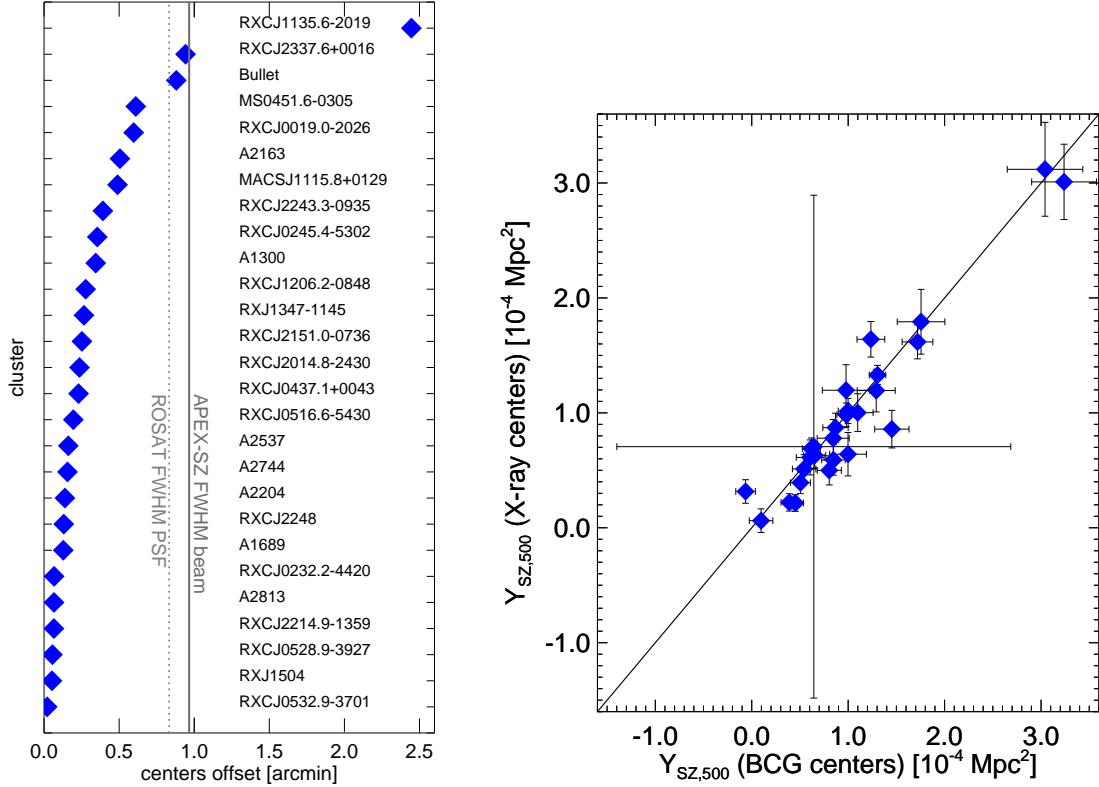


Figure 5.4: *Left:* The angular offset between the centres from ROSAT X-ray and the centres adopted for weak-lensing mass estimate are shown for the 27 eDXL clusters. 80% of the clusters have an offset between optical and X-ray centres that is smaller than half of the APEX-SZ beam. RXCJ1135 shows by far the largest offset of 2.5 arcmin. *Right:* Integrated Comptonization measured with X-ray centroid vs. optical BCG. The solid line is the 1:1 relation.

Figure 5.3 compares the input prior masses from lensing estimate and the constraint on the masses obtained by the MCMC. The red solid line marks the one-to-one relation. This illustrates that the prior values of the masses are compatible with the constraints from the data for most clusters. Bullet appears to be an exception which is not entirely surprising given that it is observed with the highest signal-to-noise in the APEX-SZ maps. Even so, the discrepancy is less than 3σ .

5.1.2 Centers for the gNFW model

The Y_{Sz} measured above used optical centres (i.e., BCG). The centroid offset between BCG and the ICM gas profile can bias the measured Y_{Sz} . In Figure 5.4, the distance between the X-ray centres of the eDXL clusters identified in ROSAT survey and the optical centres that were used to measure the weak-lensing masses are shown for all clusters. It is found that most of the clusters including merging systems like Bullet have centroid offsets in optical and X-ray at a level lower than the APEX-SZ FWHM beam. The single most extreme outlier is RXCJ1135, which is a double cluster system that appears to be in a pre-merger state. It has two dark matter peaks and a third diffuse one in between the two. The ROSAT X-ray centre lies between the two DM peaks. Measuring the Y_{Sz} signal measured at

Table 5.3: A summary of the components of the analysis used in A. Bender's work

Method	Radial Binning
Map area	circular area of 314 arcmin sq.
Model	gNFW model
free parameters	P_0 normalisation or ΔT_0
centroids	centers fixed to X-ray centroids given in Bender et al. (2016)
R_{500}	R_{500} values reported in Bender et al. (2016)

the optical centre yields a non-detection, whereas at the X-ray centre we obtain a 3σ level detection. Later in Chapter 6, these re-measured Y_{SZ} at X-ray centres will also be considered for measuring scaling relations.

5.2 Comparison with Bender et al.

To check for statistical compatibility between the integrated Comptonizations reported in Bender et al. (2016) and this work, the centroids and apertures (R_{500}) quoted in the former were used to re-measure the integrated Comptonizations for 41 clusters. This was done mainly to verify if the two pipelines, BoA and MATLAB, yield consistent measurements of the parameter. Such a compatibility test is not necessarily useful for informing whether the pipelines are accurate but rather a consistency check. Strong evidence for inconsistency would be an indication for un-trustworthy pipelines or representations of the pipelines in either or both of the work. The circular sky area about the centroids was fixed to be same as in Bender et al. (2016). A summary of the modelling choices adopted for the fitting are given in Table 5.3.

Here, the measured $Y_{\text{sph},500}$ from the previously published ones from APEX-SZ in Bender et al. (2016) and the measured ones with the independent pipeline BoA are compared. The gNFW model parameterizations, cluster centroids and R_{500} 's were kept identical to Bender et al. (2016). The model was re-fit to the BoA reduced maps using the radial binning method described in this work. The re-measured $Y_{500,\text{BoA}}$ and the literature values are plotted in the Figure 5.5 and the measurements from both work are given in Table 5.4. We have ignored relativistic corrections to the Y_{SZ} in this work, therefore, for our comparison, we modify accordingly the Y_{SZ} from Bender et al. (2016) to compute them without the relativistic corrections.

Across 41 clusters in the APEX-SZ targets that were used in the previous study, in general, there is a good agreement in the measurements. The B. C. Kelly (2007) method was used to quantify the linear relation (Equation 5.1) between the two sets of measurements by defining the relation,

$$\frac{Y_{\text{BoA}}}{10^{-10} \text{ sr}} = \alpha + \beta \frac{Y_{\text{Bender}}}{10^{-10} \text{ sr}} + \sigma_{\text{int.}} \quad (5.1)$$

The constraints obtained on $\alpha = 0.0122 \pm 0.0265$, $\sigma_{\text{int.}} = 0.056 \pm 0.031$ and $\beta = 0.961 \pm 0.039$. It should be noted that these measurements are expected to be highly correlated as they are measured from the same data sets. However, the analyses were performed independently using different pipelines. The analysis described in Section 3.4 was homogeneously applied to all clusters, in contrast to the cluster signal-to-noise optimisation adopted in the previous work by Bender et al. (2016). In the current

Table 5.4: Y_{sph} from Bender et al.⁰ and this work.

Cluster name	Bender et al. $Y_{\text{sph}} \times 10^{-10}$ sr	This work $Y_{\text{sph}} \times 10^{-10}$ sr
A2744	1.36 ± 0.25	1.96 ± 0.18
A2813	1.14 ± 0.17	1.13 ± 0.18
A209	1.96 ± 0.31	1.44 ± 0.37
RXCJ0232	1.41 ± 0.15	1.37 ± 0.14
A383	0.87 ± 0.23	0.31 ± 0.17
RXCJ0437	0.59 ± 0.12	0.53 ± 0.11
MS0451	0.48 ± 0.12	0.57 ± 0.10
A520	1.58 ± 0.22	1.63 ± 0.20
RXCJ0516	0.48 ± 0.13	0.47 ± 0.11
RXCJ0528	0.96 ± 0.14	0.71 ± 0.13
RXCJ0532	1.55 ± 0.21	1.26 ± 0.19
A3404	3.18 ± 0.44	2.82 ± 0.43
Bullet	3.86 ± 0.17	3.61 ± 0.30
A907	1.81 ± 0.28	1.53 ± 0.27
XMMXCSJ0959	0.07 ± 0.06	-0.003 ± 0.041
RXCJ1023	1.23 ± 0.17	1.42 ± 0.15
MS1054	0.40 ± 0.06	0.30 ± 0.04
MACSJ1115	0.64 ± 0.19	0.86 ± 0.12
A1300	1.39 ± 0.27	1.60 ± 0.16
RXCJ1206	1.37 ± 0.19	1.02 ± 0.17
XMMJ1230	0.06 ± 0.03	0.06 ± 0.02
RDCSJ1252	0.004 ± 0.018	0.01 ± 0.01
MACSJ1311	0.06 ± 0.07	0.26 ± 0.05
A1689	6.57 ± 0.58	4.80 ± 0.48
RXCJ1347	1.74 ± 0.22	2.03 ± 0.27
MACSJ1359	0.12 ± 0.12	0.22 ± 0.08
A1835	2.88 ± 0.37	2.98 ± 0.32
RXCJ1504	2.60 ± 0.50	2.18 ± 0.43
A2163	5.33 ± 0.82	5.05 ± 0.94
A2204	4.92 ± 0.41	5.04 ± 0.33
MACSJ1931	0.07 ± 0.21	0.26 ± 0.13
RXCJ2011	0.12 ± 0.10	0.04 ± 0.06
RXCJ2014	0.13 ± 0.48	0.30 ± 0.32
MACSJ2046	0.13 ± 0.11	0.25 ± 0.08
RXCJ2214	0.86 ± 0.11	0.82 ± 0.11
XMMXCSJ2215	0.002 ± 0.025	-0.01 ± 0.02
XMMUJ2235	-0.005 ± 0.023	0.05 ± 0.02
RXCJ2243	1.13 ± 0.15	0.78 ± 0.14
AS1077	0.55 ± 0.27	-0.05 ± 0.19
A2537	1.03 ± 0.17	0.61 ± 0.15
RXCJ2337	1.04 ± 0.21	0.91 ± 0.17

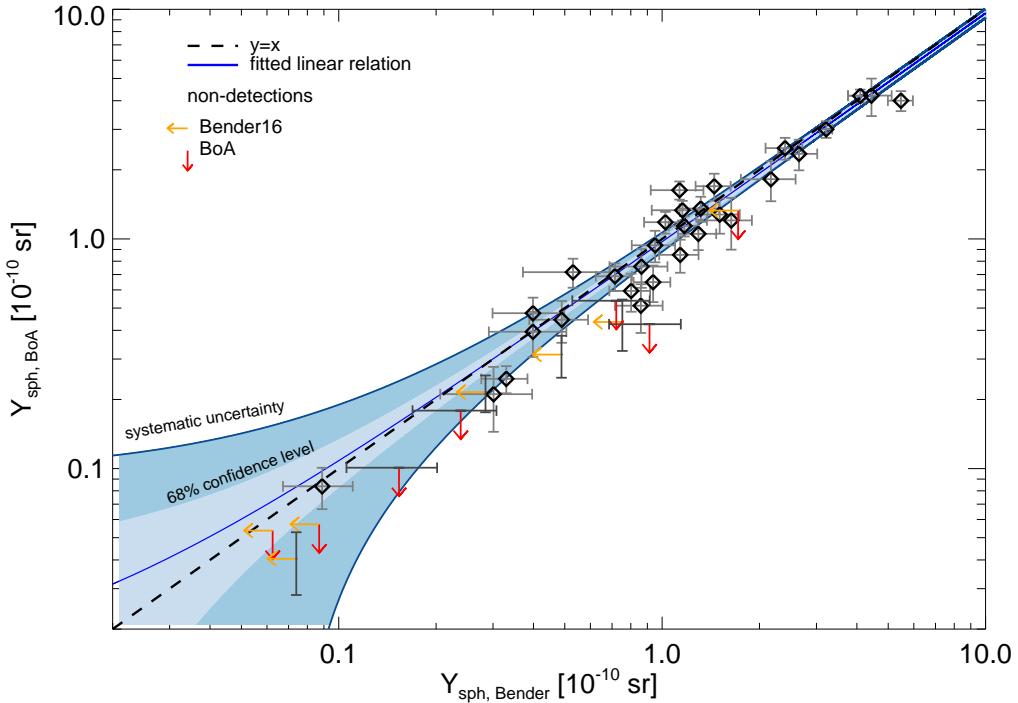


Figure 5.5: $Y_{\text{sph},500}$ measured from two pipelines (BoA and MATLAB (Bender et al. 2016)) are shown here along with their uncertainties and upper limits wherever appropriate. The correlation between the two pipelines using the Kelly method shows that the measurements are statistically consistent for the 41 clusters in the sample. The solid line is the best-fit relation and the dashed line is the one-to-one relation.

work, the point source transfer function (PST) was used to model the filtered data, whereas Bender et al. (2016) used a fixed cluster transfer function. One-to-one comparison of the two pipelines is, however, beyond the scope of this work.

5.3 Summary and discussion

In this chapter, the generalised Navaro-Frenk-White model for the ICM pressure distribution is fit to the APEX-SZ maps by using the radial binning method. The centroids for the model is first fixed to the optical centers used in the weak-lensing analysis by Klein et al. (in preparation). The model is fit with one free parameter and by using a prior on the scale radius from the lensing estimation of R_{500} . The measurement covariance between the lensing masses and the Y_{SZ} is measured by propagating the R_{500} distribution using an MCMC. The Y_{SZ} is also re-measured by changing the assumptions on the centroids ROSAT X-ray values of centres. Furthermore, the Y_{SZ} is re-computed with the same modelling assumptions as (Bender et al., 2016) and these measurements are used to then compare statistically with those measured by (Bender et al., 2016). Overall, the Y_{SZ} from this work and Bender et al. show statistical consistency.

⁰ The numbers quoted in the Table are without relativistic corrections. The actual values given in Bender et al. have included relativistic corrections.

Bender et al. (2016) provided Y_{SZ} measurements from APEX-SZ observations for specific cluster profiles with fixed centroids and fixed values of the scale radius determined from heterogeneous X-ray analysis by various literature. These Y_{SZ} measurements were obtained by using a cluster transfer function that could only be computed for one such profile for each cluster. It is, therefore, not straightforward to use the values given by Bender et al. for obtaining a cross-calibration of the SZ effect and lensing mass. To consistently re-measure the Y_{500} with the lensing prior on R_{500} for our comparison with lensing masses, we use the BoA pipeline filtered data. This allows us to also vary modelling assumptions, such as, varying the cluster profile centroids, and use them for the subsequent scaling relation analysis, as done in Chapter 6. The integrated Comptonization measured in this Chapter is, therefore, an important step and forms the basis for the mass-observable relation presented in Chapter 6.

CHAPTER 6

Weak-lensing mass calibration of the Sunyaev-Zel'dovich effect using APEX-SZ galaxy clusters

Overview

Note: A major part of this chapter was submitted for publication in the journal Monthly Notices of the Royal Academic Society in an article titled “Weak-lensing mass calibration of the Sunyaev-Zel'dovich effect using APEX-SZ galaxy clusters” by A. Nagarajan, F. Pacaud, M. Sommer, M. Klein, K. Basu, F. Bertoldi, A. T. Lee, P. A. R. Ade, A. N. Bender, D. Ferrusca, N. W. Halverson, C. Horellou, B. R. Johnson, J. Kennedy, R. Kneissl, K. M. Menten, C.L. Reichardt , C. Tucker , B. Westbrook

The use of galaxy clusters as precision cosmological probes relies on an accurate determination of their masses. However, inferring the relationship between cluster mass and observables from direct observations is difficult and prone to sample selection biases. In this work, we use weak lensing as the best possible proxy for cluster mass to calibrate the Sunyaev-Zel'dovich (SZ) effect measurements from the APEX-SZ experiment. For a well-defined (ROSAT) X-ray complete cluster sample, we calibrate the integrated Comptonization parameter, Y_{SZ} , to the weak-lensing derived total cluster mass, M_{500} . We employ a novel Bayesian approach to account for the selection effects by jointly fitting both the SZ Comptonization, $Y_{\text{SZ}} - M_{500}$, and the X-ray luminosity, $L_x - M_{500}$, scaling relations. We also account for a possible correlation between the intrinsic (log-normal) scatter of L_x and Y_{SZ} at fixed mass. We find the corresponding correlation coefficient to be $r = 0.47^{+0.24}_{-0.35}$, and at the current precision level our constraints on the scaling relations are consistent with previous works. For our APEX-SZ sample, we find that ignoring the covariance between the SZ and X-ray observables biases the normalization of the $Y_{\text{SZ}} - M_{500}$ scaling high by $1 - 2\sigma$ and the slope low by $\sim 1\sigma$, even when the SZ effect plays no role in the sample selection. We conclude that for higher-precision data and larger cluster samples, as anticipated from on-going and near-future cluster cosmology experiments, similar biases (due to intrinsic covariances of cluster observables) in the scaling relations will dominate the cosmological error budget if not accounted for correctly.

6.1 Introduction

The Λ CDM model of Big Bang cosmology predicts a hierarchical, gravity-driven scenario of structure formation in which galaxy clusters are the largest and most recently assembled quasi-virialized structures. The abundance of galaxy clusters in mass-redshift space depends on cosmological parameters, making it a sensitive probe of cosmology. To derive cosmological constraints from the surveys of galaxy cluster population, it is crucial to have an accurate mass calibration of cluster observables. The baryonic components in galaxy clusters, such as stars, galaxies, and the intra-cluster medium (ICM), are visible in a wide range of the electromagnetic spectrum. In contrast, the cold dark matter component can only be measured indirectly, e.g., through the gravitational distortion of background light, which becomes increasingly challenging to measure for galaxy clusters at higher redshifts. To relate direct observables to cluster mass, it is of great advantage that, for a scale-free initial matter power spectrum, cosmic structures evolve in a self-similar way (Kaiser, 1986). Under the simplifying assumption that the cluster ICM is in isothermal, hydrostatic equilibrium with an isothermal dark matter distribution, cluster observable global properties and the total mass of the cluster are related by simple scaling relations. These scaling relations are an essential link between cluster observables and cosmology, but also show considerable advantage in probing the thermodynamic history of the ICM (Giordini et al., 2013). There have been major efforts to empirically study the scaling relations from cluster observations, in order to constrain cosmological parameters from cluster number count measurements (e.g., Vikhlinin, Burenin et al., 2009; A. Mantz, Allen, Ebeling et al., 2010). The observational selection of representative cluster samples always relies on some directly observable quantity, such as X-ray luminosity, Sunyaev-Zel'dovich Comptonization observables or optical richness (e.g., Böhringer, Chon, Collins et al., 2013; Bleem et al., 2015; Oguri, 2014). Despite the limited size of current cluster samples, the calibration of scaling relations already emerges as the limiting factor in the error budget of number count studies of galaxy clusters (Allen, Evrard and A. B. Mantz, 2011). Mass calibration will become critical with on-going and next generation galaxy cluster surveys (SPT-3G: Benson, Ade et al. (2014), eROSITA: Merloni et al. (2012), Euclid: Laureijs et al. (2011), LSST: LSST Science Collaboration et al. (2009)), which are expected to increase sample sizes by two orders of magnitude.

Any sample of galaxy clusters is generally affected by a number of biases that depend on the underlying mass distribution, the intrinsic covariance of the cluster observables, additional measurement uncertainties and the selection method (e.g., Stanek, Evrard et al., 2006; Pacaud, Pierre et al., 2007; Vikhlinin, Burenin et al., 2009; A. Mantz, Allen, Ebeling et al., 2010). Mass functions predicted by simulations (e.g., Tinker et al., 2008) and determined from cluster surveys (e.g., Reiprich and Böhringer, 2002; Vikhlinin, A. V. Kravtsov et al., 2009) have shown the number density of clusters to be an exponentially decreasing function of cluster mass, including a trend with redshift. In the presence of scatter (intrinsic as well as that arising from measurement uncertainties), this will cause more low-mass clusters to up-scatter to a given observed mass than high mass clusters to down-scatter to that same level, thus distorting the distribution of sources in the space of observables - an effect known as Eddington bias (Eddington, 1913). These distortions are further exacerbated in the presence of sample selection thresholds that truncate the scattered distributions. In addition, depending on their distances, the selected clusters are not drawn from the same mass distribution due to the combined effect of the cosmological growth of structures, the surveyed volume and source selection thresholds - the well known Malmquist bias (Malmquist, 1920). As a consequence, samples selected by luminosity would typically be biased towards low masses and intrinsically bright sources. In the presence of a

(positive) correlation in the intrinsic scatters of the selecting mass observable and a follow-up mass observable, the follow-up observable would also be, on average, biased towards intrinsically bright sources (e.g. Allen, Evrard and A. B. Mantz (2011)). An accurate calibration of cluster scaling relations requires that these biases are controlled and corrected for.

In the work presented here, we focus on scaling relations involving the Sunyaev-Zel'dovich (SZ) effect (Sunyaev and Zeldovich, 1970, also see Chapter 2). While the recent availability of SZ-selected galaxy clusters for cosmological analysis has resulted in several precise constraints on cosmology (e.g. de Haan et al. (2016), Planck Collaboration, Ade, P. A. R., Aghanim, N. and et al. (2016) and Planck Collaboration, Ade, Aghanim and al. (2016a)), these studies largely rely on prior information on the SZ-mass calibration obtained from X-ray derived masses and/or weak-lensing masses. Thus, directly calibrating the integrated SZ Comptonization (Y_{SZ}) with cluster mass (M_{500}) has generated much interest. Weak-lensing mass estimates are best suited for calibrating cluster masses as they directly measure the line-of-sight matter distribution and do not rely on further assumptions about the physical state of matter inside clusters (like hydrostatic equilibrium or thermal pressure support). Simulations indicate that lensing masses are biased by at most a few percent (Becker and A. V. Kravtsov, 2011; Meneghetti et al., 2010; Rasia et al., 2012).

Early studies of the scaling between weak-lensing mass and SZ Comptonization suffered from either poorly defined selection criteria (e.g., Hoekstra, Mahdavi et al., 2012; Sereno and Ettori, 2015), or were limited by the availability of lensing and SZ observations (e.g., Marrone et al., 2012). Additionally, in cases where the sample selection was based on X-ray luminosities, the effects of possible correlations in the intrinsic scattering of SZ Comptonization and X-ray luminosity at fixed mass were unaccounted for (e.g., Marrone et al., 2012; A. Mantz, Allen, Morris et al., 2016).

Numerical simulations have predicted that at a given cluster mass, the dispersion of global thermodynamic properties are correlated (e.g., Truong et al., 2016; Angulo et al., 2012; Stanek, Rasia et al., 2010). In particular, these authors find a correlation in the intrinsic scatter of X-ray luminosity and integrated Comptonization in the range of 0.5–0.8. If this correlation is unaccounted for, it can bias the inferred $Y_{\text{SZ}} - M_{500}$ scaling relation for a sample that is selected on X-ray luminosities. Observationally, this correlation remains largely unconstrained.

In this work, we employ a sample of 39 galaxy clusters observed with the SZ effect using the APEX telescope (Schwan et al., 2011). To provide an accurate mass calibration of the integrated Comptonization, we measure the scaling relation of the Comptonization with weak-lensing derived masses of an X-ray selected sub-sample of 27 clusters from APEX-SZ with a well-defined selection function, henceforth referred to as eDXL. We present a Bayesian method to account for the sample selection while placing an emphasis on controlling the bias in the scaling relations due to the correlated intrinsic scatter of the selection observable (X-ray luminosity) and scaling observable (integrated Comptonization) at fixed mass.

This chapter is organized as follows: Section 6.2 summarizes the mass proxy measurements for the eDXL sample from previous chapters. Section 6.3 presents a Bayesian method for fitting scaling relations while accounting for selection effects. The results are presented in Section 6.4 and its robustness, systematics and limitations are discussed in Section 6.5. A discussion on the significance of the results is presented in Section 6.6. We offer our conclusions in Section 6.7. Unless otherwise noted, we assume a Λ CDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

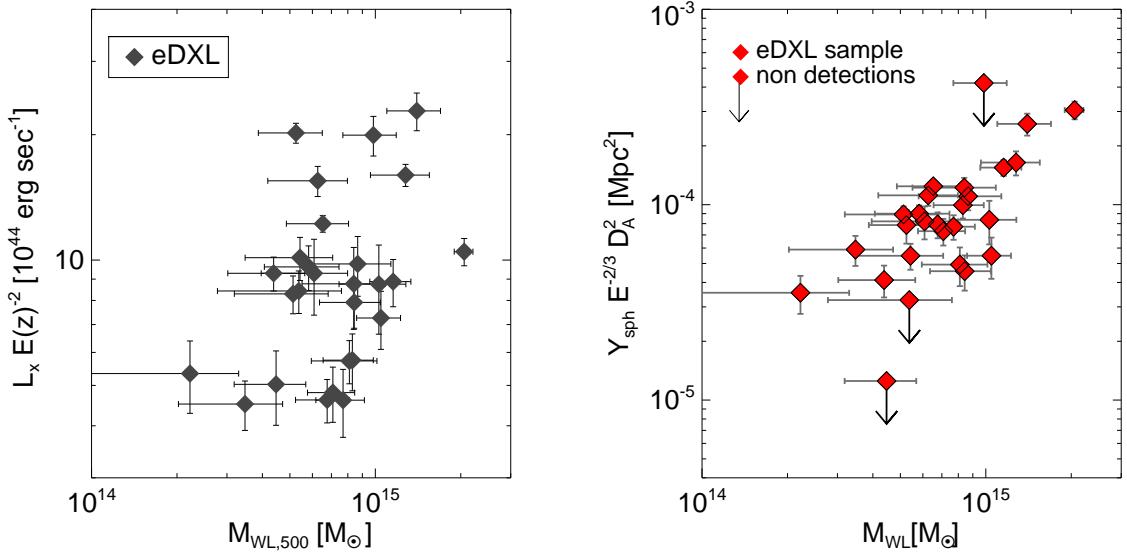


Figure 6.1: *Left:* ROSAT luminosities $L_x[0.1 - 2.4 \text{ keV}]$ and weak-lensing spherical masses for the eDXL sample with lensing follow-up. The spherical weak-lensing masses are measured within R_{500} . The luminosities are measured within a R_{500} that is independent of the weak-lensing analysis (See section 3.2.3). *Right:* The integrated Comptonization from APEX-SZ vs. weak-lensing mass estimate for the X-ray selected eDXL sample.

6.2 Mass proxy measurements

In this chapter, we use the measured weak-lensing mass estimates, the integrated Comptonization measurements from APEX-SZ, and X-ray luminosities from ROSAT (for sample selection) to study the mass calibration of the integrated Comptonization. Figure 6.1 summarizes these measured properties of the eDXL sample. The integrated Comptonization used in the analysis presented in this Chapter was obtained by assuming a gNFW profile for the electron pressure centred on the BCG as given in Chapter 5. The integrated Comptonization was measured within the same aperture as the lensing mass estimate of R_{500} . This introduces correlation in the measurements of lensing mass and integrated Comptonization. For each cluster, the measurement covariance between lensing mass and integrated Comptonization was estimated using an MCMC and the lensing prior on R_{500} (see Chapter 5, Section 5.1.1).

6.3 Method

We present a Bayesian method to account for sample selection biases in the scaling relations for the eDXL sample in which the sample selection is well-defined. Several authors have discussed using Bayesian techniques for measuring cluster scaling relations (e.g., B. C. Kelly, 2007; Andreon and Hurn, 2013; Maughan, 2014; A. Mantz, Allen, Ebeling et al., 2010; Sereno and Ettori, 2015). In this work, we apply a Bayesian formalism for measuring jointly multiple mass-observable scaling relations by accounting for a truncated selection in a measured cluster property. We differ from some

of the other work by not requiring for a model to predict the number counts of the underlying or missing population (e.g., A. Mantz, Allen, Ebeling et al., 2010) but still accounting for the shape of the underlying cluster mass function, the sample selection, the measurement uncertainties in cluster properties and masses, and the intrinsic covariances of cluster properties. The completeness of the sample allows us to compute a semi-analytical approximation for accounting for non-ignorable sample selection effects. In particular, we deal with this impact for measuring the mass scaling relations of cluster observables that do not play a role in the selection of cluster members of a sample. Our likelihood presented here bears the most similarity to the XXL likelihood used by Giles, Maughan, Pacaud et al. (2016), however, they use temperature function to model the underlying cluster population, include a scaling relation between two cluster properties with only one intrinsic scatter, their selection depends on two observables rather than measured properties, and they measure the scaling relation between cluster properties that play some role in the sample selection.

In Section 6.3.1, we outline the general framework of the method while Section 6.3.2 discusses the application of that statistical model to the eDXL sample. In Section 6.3.3, we validate our application of the statistical model for our eDXL measurements through analyses of mock data.

6.3.1 Statistical model

The key ingredients for the statistical model to determine the posterior distribution of the parameters of interest are shortly described below:

1. The mass variable, m , is the fundamental variable that describe a cluster and relate to all other observables, arranged in a vector ξ , through a scaling model $P(\xi|m, \theta)$ that is fully described by parameters θ , which needs to be determined. The full probability distribution in the mass-observable plane is obtained from the conditional probability rule:

$$P(\xi, m|\theta) = P(\xi|m, \theta)P(m). \quad (6.1)$$

2. To model the conditional probability $P(m)$, some authors leave full freedom for this function by introducing flexible parametric model (such as the multiple Gaussians of B. C. Kelly, 2007), to be constrained simultaneously in the fit. However, there is some knowledge of the cluster mass function and using it reduces degeneracies in the fit. Therefore, we use the cluster mass function as the $P(m)$. In practice, it is evaluated using the Tinker mass function (Tinker et al., 2008) in our reference cosmology, for a density contrast of $500\times\rho_c$. $P(m)$ is then proportional to the mass function $\frac{dn}{dm}(m, z)$. A proper estimation of the normalisation constant would, in general, require to set a lower limit to the cluster mass, but since we do not vary $P(m)$ in our model that estimation is not required in practice.
3. Numerical simulations and observations both demonstrated that the average mass observable scalings have power-law shapes (possibly broken power-laws when including groups) and support a Gaussian scatter in log space around these average power laws (e.g., Giodini et al., 2013; Stanek, Rasia et al., 2010; Angulo et al., 2012). The true ensemble average (over a large volume) of a global observable, $\hat{\xi}^i$, is related to mass of clusters m and redshift z as

$$\hat{\xi}^i(m, z) = \alpha^i(z)[m]^{\beta^i}, \quad (6.2)$$

where α^i and β^i are the logarithmic normalisation and slope respectively and m is the independent variable.

Deviations from a perfect power-law scaling relation are expected due to the diversity of dynamical states in galaxy cluster population, non-gravitational physics, projection effects, etc, affecting the cluster observables. The random variables ξ can be modelled as originating from a multi-variate log-normal probability density function $P(\xi|m, \Sigma, \alpha, \beta)$, where Σ is the log-normal intrinsic covariance matrix of cluster observables at fixed mass. The diagonal elements of Σ give the log-normal intrinsic scatter for a corresponding cluster observable at fixed mass, which we denote as $\sigma_{\ln \xi^i}$. The off-diagonal terms quantify the covariance of different cluster observables at fixed mass. For $i \neq j$, the covariance between the cross-terms is related by the correlation coefficient:

$$r_{ij} \equiv \frac{\Sigma_{ij}}{\Sigma_i \Sigma_j}. \quad (6.3)$$

4. Typically, we access the cluster observables through a set of observations. The measured cluster properties and mass are denoted with a tilde as $\tilde{\xi}, \tilde{m}$ respectively. The link between the true observable and its noisy estimate is provided by a measurement model $P(\tilde{\xi}, \tilde{m}|\xi, m)$.

The probability of measured cluster observables, $\tilde{\xi}$ and mass, \tilde{m} for a single cluster is

$$P(\tilde{\xi}, \tilde{m}|\theta) = \int_{\Xi} d\xi \int_0^{+\infty} dm P(\tilde{\xi}, \tilde{m}|\xi, m) P(\xi|m, \theta) P(m), \quad (6.4)$$

where Ξ is the domain in which ξ is defined.

5. The mass distribution and scaling relation model used to derive equation (6.4) refer to the whole cluster population. In practice, one can never access a pure mass selected sample and always has to deal with a censored population, were a sub-sample has been selected based on some of the observables. We here describe this selection process through a detection probability $P(\mathcal{I} = 1 | \xi_k, \phi)$, where \mathcal{I} is a boolean random variable specifying whether a k^{th} cluster was detected or not and ϕ are a number of additional model parameters that describe the selection process. The generative probability model for a cluster that passed the selection is now conditional on $\mathcal{I} = 1$ and can be expressed using Bayes theorem as:

$$P(\tilde{\xi}_k, \tilde{m}_k | \mathcal{I} = 1, \theta, \psi, \phi) = \frac{P(\mathcal{I} = 1 | \tilde{\xi}_k, \tilde{m}_k, \phi) P(\tilde{\xi}_k, \tilde{m}_k | \theta, \psi)}{P(\mathcal{I} = 1 | \theta, \psi, \phi)}. \quad (6.5)$$

The overall probability for clusters to be selected, which appears in the denominator, can be estimated by averaging the observable dependent selection probability over the global distribution of cluster observables provided by equation (6.4), i.e.:

$$P(\mathcal{I} = 1 | \theta, \psi, \phi) = \int_{\Xi} d\tilde{\xi}_k \int_{\tilde{M}} d\tilde{m}_k P(\mathcal{I} = 1 | \tilde{\xi}_k, \tilde{m}_k, \phi) P(\tilde{\xi}_k, \tilde{m}_k | \theta, \psi). \quad (6.6)$$

6. The likelihood of the scaling relation parameters given a complete set of N_{det} detected clusters

follows from equation (6.5):

$$\mathcal{L}(\theta | \tilde{\xi}_{\text{obs}}, \tilde{m}_{\text{obs}}, \psi, \phi) = \prod_{k=1}^{N_{\text{det}}} P(\tilde{\xi}_k, \tilde{m}_k | \mathcal{I} = 1, \theta, \psi, \phi), \quad (6.7)$$

where $\tilde{\xi}_{\text{obs}}$ is used to denote the full matrix of cluster observables measurements of all the detected clusters, \tilde{m}_{obs} denotes the full set of mass measurements for the detected sample of clusters and the posterior reads:

$$\mathcal{P}(\theta | \tilde{\xi}_{\text{obs}}, \tilde{m}_{\text{obs}}, \psi, \phi) = \pi(\theta) \times \mathcal{L}(\theta | \tilde{\xi}_{\text{obs}}, \tilde{m}_{\text{obs}}, \psi, \phi), \quad (6.8)$$

where $\pi(\theta)$ is the prior on the model parameters.

6.3.2 Application to the eDXL sample

We apply the method discussed in the previous subsection to our X-ray selected sample (eDXL). For this sample the class of cluster properties on which the selection function depends on is the measured X-ray luminosity (\tilde{L}_x) of clusters in the energy band 0.1–2.4 keV. Our primary goal is to measure the scaling relation of the Y_{SZ} to M_{500} using the eDXL sample. We remind here that the statistical model for the likelihood described above takes into account the impact of the sample selection function, the measurement uncertainties of cluster observables, intrinsic covariances between cluster observables at fixed mass, and the underlying cluster mass function. We describe below briefly the essential components required for the likelihood in Equation (6.5) to determine the posterior of the scaling relation parameters.

We use the mass as the fundamental variable and cluster properties (such as Y_{SZ} and L_x) as the response variables. Unbiased weak-lensing masses provide an absolute mass calibration for scaling relations and as already mentioned earlier, we expect the bias in lensing masses to be negligible as predicted by numerical simulations (e.g., Meneghetti et al., 2010; Becker and A. V. Kravtsov, 2011; Rasia et al., 2012). Thus, our weak-lensing masses are a natural choice for anchoring the cluster masses. We note here that intrinsic scatter in the lensing masses can, however, occur due to elongation and projection effects along the line-of-sight (e.g., Becker and A. V. Kravtsov, 2011; Gruen et al., 2015; Shirasaki, Nagai and Lau, 2016), which can, in turn, produce biases in measuring scaling relations if not modelled correctly (e.g., Sereno and Ettori, 2015).

In order to properly account for several sources of uncertainties and systematic effects simultaneously, we consider two ways of modelling the scaling relations. In the first model, we assume no intrinsic scatter in the weak-lensing mass, essentially making the true lensing mass same as the spherical overdensity halo mass (M_{HM} or M_{500}). We describe the corresponding set of scaling models in Section 6.3.2.1. In the second model, we assume a fixed intrinsic scatter in the true weak-lensing masses (section 6.3.2.2). In both cases, the underlying cluster mass function in the redshift-mass space is described by the Tinker halo mass function (Tinker et al., 2008). The inverse situation with either luminosity or Y_{SZ} being the independent variable would require knowledge of their number density which in turn depends on the scaling law with the total mass of the cluster. We note that the use of the mass function depends on the cosmological parameters, most prominently on σ_8 , Ω_m , Ω_Λ . By fixing these parameters, we are assuming an *a priori* perfect knowledge of the mass function. The impact of this somewhat strong assumption is mitigated by the fact that the number density of galaxy

clusters is not included in our likelihood model. We only rely on the distribution in the measurements (\tilde{M}_{WL} , \tilde{Y}_{SZ} , \tilde{L}_x) space, which only depends mildly on the shape of the cluster mass function.

The measurements \tilde{Y}_{SZ} and \tilde{M}_{WL} are drawn from a bi-variate Gaussian distribution. Incorporating this probability density as such, naturally takes into account the non-detections in the SZ and does not require any special correction to the probability density. The measured values of X-ray luminosities are treated as coming from a log-normal distribution with the log-normal uncertainty $\sigma_{\ln \tilde{L}_x}$ and is independent of other measured properties. The explicit expression of the probability densities are given in Appendix A.2. In our implementation of the likelihood, we marginalise over the true variables (X-ray luminosities, integrated Comptonizations, masses) through an MCMC.

The selection function for the eDXL sample is a Heaviside step function that depends on the observed luminosities and the applied minimum luminosity threshold, i.e., $P(\mathcal{I} = 1 | \tilde{L}_x, \tilde{L}^{\min}) = 1$ only when $\tilde{L}_x \geq \tilde{L}^{\min}$, where the thresholds correspond to the defined values in Section 3.2.1. The normalisation of the likelihood (Equation 6.5) is computed for each redshift of the eDXL sample and is dependent on the scaling parameters of the $L_x - M$ relation. This necessitates the joint modelling of multi-observable to mass scaling relations. Moreover, this joint modelling also has the advantage of considering a possible covariance between Y_{SZ} and L_x at fixed mass. The log-normal measurement uncertainty and the log-normal intrinsic scatter in X-ray luminosities allows us to analytically integrate the normalisation in Equation (6.5) over the variables L_x and \tilde{L}_x . Furthermore, the nature of the threshold cut selection gives an expression with an error function and this modulates the mass function for the sample especially at the low mass end. The explicit expression of the normalised likelihood for the eDXL sample is given in Appendix A.1. This expression of the normalisation of the likelihood remains the same for both set of scaling models discussed in Section 6.3.2.1 and 6.3.2.2.

In the subsections below, we describe the two different scaling models.

6.3.2.1 Without intrinsic scatter in lensing mass

Scaling model: The prescription for the scaling laws of the observables with the mass of a cluster is defined as

$$\frac{\hat{L}_x}{L_{\text{pvt}}} E(z)^{-2} = A_{\text{LM}} \left(\frac{M_{\text{HM},500}}{M_{\text{pvt}}} \right)^{B_{\text{LM}}} ; \quad (6.9)$$

$$\frac{\hat{Y}_{\text{SZ},500}}{Y_{\text{pvt}}} E(z)^{-2/3} = A_{\text{SZ}} \left(\frac{M_{\text{HM},500}}{M_{\text{pvt}}} \right)^{B_{\text{SZ}}} , \quad (6.10)$$

where $M_{\text{HM},500}$ is the spherical halo mass or the true total mass of a galaxy cluster and, where the pivot values for luminosities, masses, and SZ Compton parameters are $L_{\text{pvt}} = 8.77 \times 10^{44} \text{ erg s}^{-1}$, $M_{\text{pvt}} = 7.084 \times 10^{14} M_{\odot}$, $Y_{\text{pvt}} = 7.93 \times 10^{-5} \text{ Mpc}^2$ respectively. The pivot values reflect the median values of the measurements $\tilde{L}_x E(z)^{-2}$, $\tilde{M}_{\text{WL},500}$, and $\tilde{Y}_{\text{SZ},500} E(z)^{-2/3}$ across the eDXL sample. We choose these values to minimise the degeneracy in measuring the normalisation and slope of the scaling relations. The above scaling power-law are modelled with log-normal intrinsic scatter in L_x and Y_{SZ}

at fixed mass with correlation parameter r . The intrinsic covariance matrix is given as follows:

$$\begin{pmatrix} \sigma_{\ln L_x}^2 & r\sigma_{\ln L_x}\sigma_{\ln Y_{\text{SZ}}} \\ r\sigma_{\ln L_x}\sigma_{\ln Y_{\text{SZ}}} & \sigma_{\ln Y_{\text{SZ}}}^2 \end{pmatrix}, \quad (6.11)$$

where $\sigma_{\ln L_x}$, $\sigma_{\ln Y_{\text{SZ}}}$ are the log-normal intrinsic scatters in L_x and Y_{SZ} at fixed mass, respectively, and r is the correlation coefficient.

In this model, we anchor the halo masses to the lensing masses by a one-to-one scaling of true lensing mass, M_{WL} , to halo mass, M_{HM} , by setting $M_{\text{HM}} = M_{\text{WL}}$.

The redshift evolutions of the scaling relations are power-law of time evolution of the Hubble parameter, $E(z)$. We use the logarithmic self-similar slope for the evolution in the $Y_{\text{SZ}}-M_{500}$ and L_x-M_{500} relations. Throughout the analysis, we keep them fixed. We fix the logarithmic slope of the redshift evolution of the L_x-M_{500} relation to the self-similar evolution value for soft-band luminosities (Ettori, 2015). This slope is shallower than the self-similar slope of bolometric luminosities and is confirmed by other authors (Vikhlinin, Burenin et al. (2009) and Sereno and Ettori (2015)). Additionally, we choose uniform priors in the interval $(0, \infty)$ for the parameter set, $\{A_{\text{LM}}, B_{\text{LM}}, A_{\text{SZ}}, B_{\text{SZ}}\}$. The priors for the intrinsic scatters ($\sigma_{\ln Y_{\text{SZ}}}$, $\sigma_{\ln L_x}$) are uniform in the interval $[0.02, \infty)$ and we place an uniform prior on the correlation parameter r in the open interval $(-1, 1)$.

6.3.2.2 With intrinsic scatter in lensing mass

To take into account a possible scatter in lensing masses, we add a scaling law between the lensing mass and true spherical overdense mass and model the lensing mass observable to scatter from the halo mass with a dispersion. This additional scaling is given below:

$$\frac{\hat{M}_{\text{WL},500}}{M_{\text{pvt}}} = A_{\text{WL}} \left(\frac{M_{\text{HM},500}}{M_{\text{pvt}}} \right)^{B_{\text{WL}}}, \quad (6.12)$$

where the normalization (A_{WL}) and the slope (B_{WL}) of the relation are both fixed to unity.

The scatter in the lensing mass from the true halo mass is predicted to be log-normal and of the level of 20–23% for the massive clusters of $M_{500} \geq 1.4 \times 10^{14} M_{\odot}$ in the redshift range of 0.25–0.50 (Becker and A. V. Kravtsov, 2011). The constraints from observations are consistent with the predictions (e.g., A. Mantz, von der Linden et al., 2015; Sereno and Ettori, 2015). Since we lack the statistical power to constrain the dispersion in lensing mass observable, we use this prior to fix the intrinsic scatter. We, however, model it to be a Gaussian dispersion of $0.20 M_{\text{HM},500}$. Introducing this lensing scatter in our modelling requires a marginalisation over the true lensing mass, M_{WL} . Since the measurement uncertainties on the lensing mass itself is from a bi-variate Gaussian distribution, our intuitive choice of a Gaussian intrinsic scatter in true lensing mass observable simplifies the marginalisation over these additional variables. Therefore, the marginalisation over these variables to relate $M_{\text{HM},500}$ to $\tilde{M}_{\text{WL},500}$ is done analytically in our implementation fully taking into account the measurement covariances between lensing masses and integrated Comptonizations (the calculations are outlined in Appendix A.2.3).

It is understood that not accounting for an intrinsic scatter in lensing masses can bias the estimate of the scaling parameters (Sereno and Ettori, 2015; Sereno, Ettori and Moscardini, 2015; Gruen et al.,

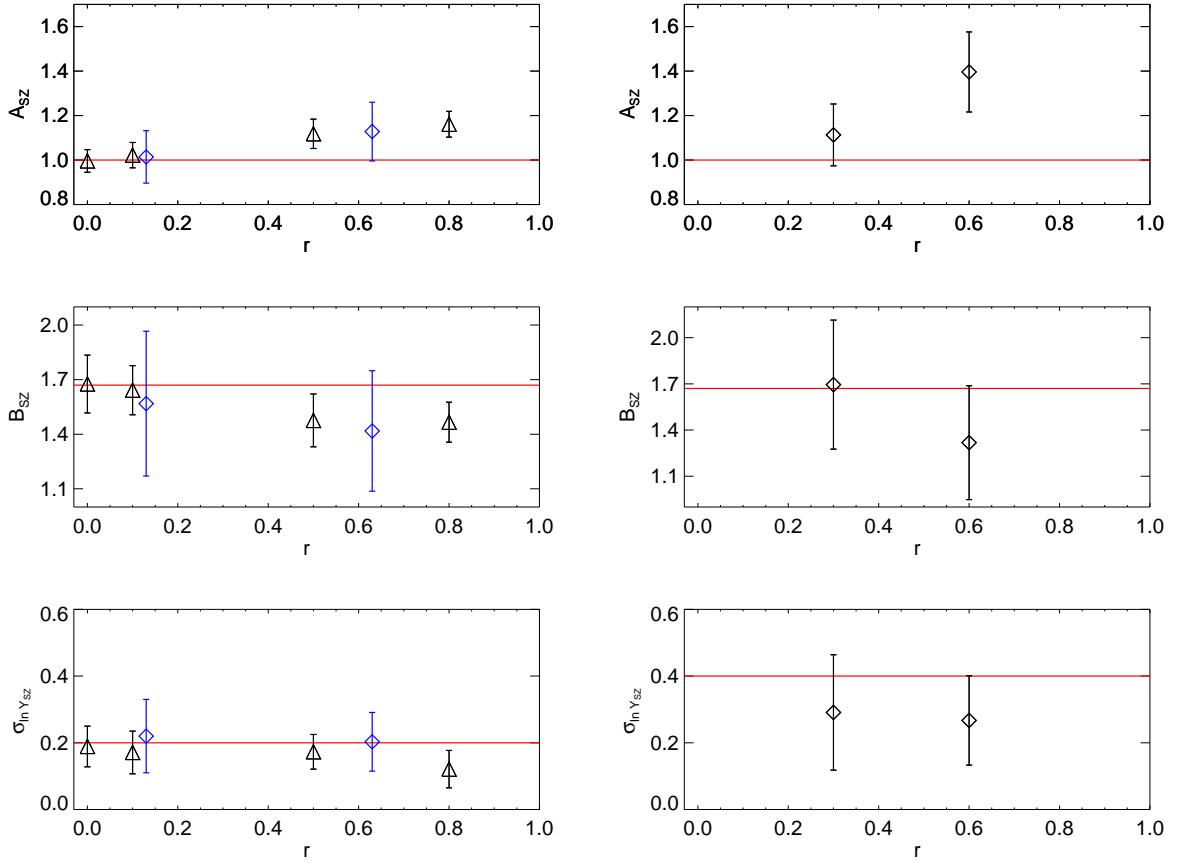


Figure 6.2: The recovered mean parameters from mock data realisations are shown here for the $Y_{\text{SZ}} - M_{500}$ relation. The red lines mark the input values of the scaling parameters. The error bar on the recovered mean value represents the uncertainty level from single set of mock sample analysis. *Left:* The black triangular symbols correspond to results from mock data with 10% measurement uncertainties. The blue diamond symbols correspond to results for mock data with realistic uncertainties (See text). *Right:* The recovered mean parameters for different input scaling relations, where $\sigma_{\ln L_x}$ and $\sigma_{\ln Y_{\text{sz}}}$ were increased to 0.6 and 0.4 respectively. We simulated the mock data with realistic uncertainties. The bias we see in the normalisation for realistic mock samples indicates that even for cluster sample size of 30 and with realistic measurements, ignoring the correlation show significant bias.

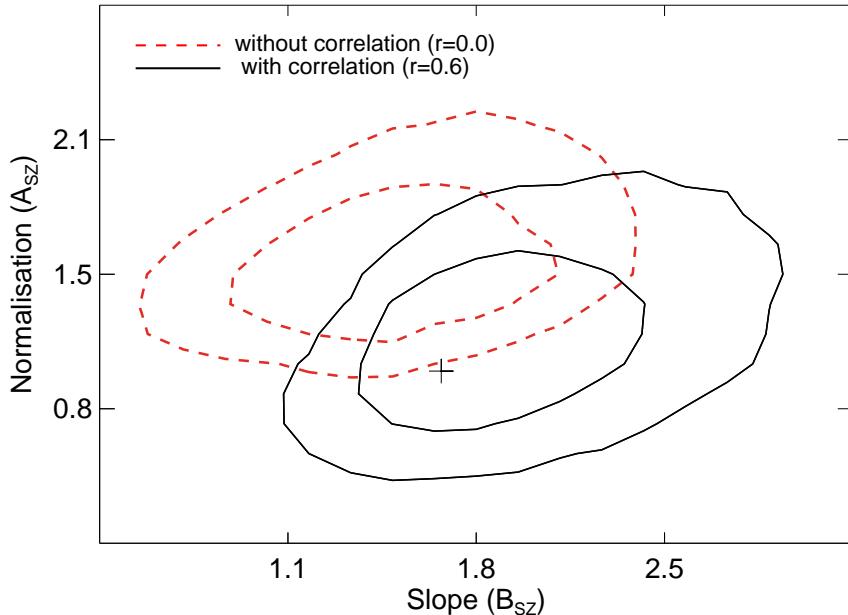


Figure 6.3: An example of 68% and 95% confidence levels of the normalisation and slope parameters of the $Y_{SZ} - M_{500}$ scaling relation fit to a mock sample. Red dashed contours: $r = 0$, and the black solid contours: $r = 0.6$. The mock sample was generated using realistic measurement errors and with a correlation ($r = 0.6$) in intrinsic scatters of L_x and Y_{SZ} . The input value of the $Y_{SZ} - M_{500}$ relation is shown as cross.

2015). But including such a scatter also requires a consideration of the correlations between the lensing scatter and the intrinsic scatters of other cluster observables at fixed mass. Due to limitations set by our measurements and sample size, we are forced to fix this scatter to 20% (Gaussian) and do not marginalise over this scatter. We assume zero correlations in the intrinsic covariances of lensing mass observable with other observables at fixed mass. Therefore, we give this model here as a consideration of the impact of such a scatter in our lensing masses on the scaling parameters and in this work, we follow the model given in Section 6.3.2.1 as our fiducial model.

6.3.3 Tests with simulations

Based on reported findings from simulations (e.g., Truong et al., 2016; Stanek, Rasia et al., 2010) that thermodynamic observables are correlated at fixed masses, we study the impact of the correlated scatters and selection biases on scaling relations of these observables. It is important and crucial to understand this impact for inferring deviations in self-similar scaling. In the following sections, we present results based on mock datasets that mimic the behaviour of our eDXL observations and measurements. We also test for more precise measurements and present a detailed description of the results in Appendix A.3, which is useful for a look-up of the level of bias in possible future surveys and follow-up studies that may have more precise measurements. We address two key issues relevant for our analysis for the eDXL sample:

1. The presence of correlated intrinsic scatters of luminosities and Compton-Y at fixed masses and the sample selection is based on the measured luminosities. We parametrize this correlation in

scatter with ‘ r ’. We present this discussion in Section 6.3.3.1.

2. Second, weak-lensing masses have an intrinsic scatter due to projection effects, we consider the impact of ignoring this information in our analysis. We present this discussion in Section 6.3.3.2.

6.3.3.1 Correlation in the scattering of selecting observable and follow-up observable at a fixed mass

In order to understand the level of bias that can occur in the recovered scaling relations if the correlation in the scatters of selecting observable and follow-up observable is ignored, we simulate sets of mock data with measurement uncertainties. Our mock samples have 30 clusters, which is similar in size to the eDXL sample. We generate mock samples with three observables: the independent variable (e.g., M_{500}), and two response variables (e.g., L_x , Y_{SZ}) modelled using a power law relation with the independent variable. The selection is on one of the response variables, namely, L_x . The samples were generated at a median redshift of 0.3 using the Tinker mass function (Tinker et al., 2008).

Realistic measurement uncertainties We assume different values of the correlation coefficient r between 0.0 and 1.0 in the intrinsic scatters of the two response variables as input for generating mock samples. We test with simulated measurements of 30% uncertainty in M_{500} and 25% on the Y_{SZ} . These measurement uncertainties reflect the median relative uncertainties of our eDXL mass and mass proxy measurements. Simulation studies report a positive correlation in the intrinsic scatters of L_x and Y_{SZ} at fixed mass in the range of 0.5-0.9 (Stanek, Rasia et al., 2010; Angulo et al., 2012; Truong et al., 2016). We choose values of correlation on the lower end (ranging between 0.1 and 0.6) for our mock samples to test when the impact starts becoming significant. We fit numerous realisations of mock data sets for each input relation using the method prescription in section 6.3.2.1 with r set to zero. The measured average recovered scaling relation parameters for the $Y_{\text{SZ}} - M_{500}$ relation are plotted in Figure 6.2 for different set of input scaling relations. The bias in the normalisation of the $Y_{\text{SZ}} - M_{500}$ relation for a set of input relations with 20% and 40% intrinsic scatters in Y_{SZ} and L_x is about 1σ . For an eDXL like sample with larger intrinsic scatters ($\sigma_{\ln L_x} = 0.6$, $\sigma_{\ln Y_{\text{SZ}}} = 0.4$ were used based on recovered values in Section 6.4), we find that the normalisation of the $Y_{\text{SZ}} - M_{500}$ relation is significantly biased at the level of 2σ . An example of recovered normalisation and slope of the $Y_{\text{SZ}} - M_{500}$ relation from a mock eDXL-like data set is shown in Figure 6.3. The input scaling relations lies outside the 95% confidence level of the recovered parameter space when ignoring the correlation in the scatters. On fitting with the assumed correlation, the recovered parameter space is consistent with the input values within 68% confidence.

Besides the normalisation, for different input relations we observe the average slope is almost 1σ shallower than the input. We also observe an under-estimation of the intrinsic scatter $\sigma_{\ln Y_{\text{SZ}}}$ for the eDXL-like data set. The full table of results is summarized in the Appendix A.3 in the Table A.1. A summary of results from this Section and Appendix A.3.1 is plotted in Figure 6.2 showing the means and standard deviations of the recovered modes for each scaling parameter of the $Y_{\text{SZ}} - M_{500}$ relation.

6.3.3.2 Intrinsic scatter in lensing masses

Here we present the analysis using mock data where we scatter the true lensing mass from the halo mass using a log-normal scatter of 20%. As done in previous Section, we generate mock samples with the realistic measurement uncertainties and eDXL-like scaling parameters. We consider the following two cases:

1. *Without correlation in the intrinsic scatters of L_x and Y_{SZ} :* We introduce un-correlated intrinsic scatters in the L_x and Y_{SZ} observables for the mock samples. We fit numerous realisations of mock samples with the model given in Section 6.3.2.1 (i.e., ignoring the lensing scatter) with $r = 0$. From the analysis of mock data with realistic measurements, we find that the recovered values of the scaling parameters show bias values less than 0.7σ . For the more precise measurements, we find a bias in the slope of the $Y_{\text{SZ}} - M$ relation and the intrinsic scatter in Y_{SZ} at fixed mass to be $\sim 1.4\sigma$. The other scaling parameters show less than 1σ bias.
2. *With correlation in the intrinsic scatters of L_x and Y_{SZ} :* Most importantly for our case, we test the impact of the presence of an intrinsic scatter in the lensing masses and simultaneously having correlated intrinsic scatters in L_x and Y_{SZ} at fixed mass. Therefore, we inject a correlation of 0.6 in the intrinsic scatters of L_x and Y_{SZ} for our mock samples. We fit the scaling relations with the same procedure as done previously by fitting scaling models with $r = 0$ and without intrinsic scatter in lensing mass. We observe a total bias of 2.7σ in the normalisation parameter of $Y_{\text{SZ}} - M_{500}$ relation. The slope of this relation is found to be biased low by 0.34 ($\sim 1\sigma$) from the input value. We fit the mock samples again but fixing the r value to 0.6. From the mean recovered scaling parameters of mock samples, we find that the bias in the normalisation reduces to $+0.5\sigma$. The slope parameter is now lowered by 0.17 (0.9σ) from the input value. This level of bias in the slope occurring due to the scatter in weak-lensing mass is consistent with the findings and discussion given in Sereno and Ettori (2015). The results from this Section are summarised in Table A.2.

6.4 Results

We jointly fit the three observables (L_x , Y_{SZ} , M_{500}) of the eDXL sample to the $L_x - M_{500}$ and $Y_{\text{SZ}} - M_{500}$ scaling relations. The main purpose of fitting the $L_x - M_{500}$ relation is to account for the sample selection.

In this section, we present the results of these fits under progressively less conservative assumptions on intrinsic scatter of weak-lensing masses. First, in Section 6.4.1, we present fits with correlated scatter in L_x and Y_{SZ} (allowing the correlation parameter, r , to vary, and fixing it) while ignoring the intrinsic scatter in the lensing masses. In Section 6.4.2, we also add the expected intrinsic scatter in weak-lensing masses.

6.4.1 Including correlated intrinsic scatters in Y_{SZ} and L_x at fixed mass

We fit the $L_x - M_{500}$ and $Y_{\text{SZ}} - M_{500}$ relations using the model described in Section 6.3.2.1. As discussed earlier, we include a correlation coefficient parameter in the intrinsic scatters of luminosity and Comptonization at fixed mass. We marginalise over the correlation parameter, r , allowing it to vary

Table 6.1: Results of the scaling relations analysis for the eDXL sample using the method described in Section 6.3. The medians and 68% confidence levels of the marginalised distributions are quoted. Centroids for gNFW model fits were fixed to the optical centres (BCG) or X-ray (X) values obtained from the ROSAT survey. The results in the first row (shaded) will be considered as our fiducial result.

Priors			Recovered parameters						
Centroid	r	$\sigma_{WL HM}$	$L_x - M_{500}$ scaling parameters				$Y_{SZ} - M_{500}$ scaling parameters		
			r	A_{LM}	B_{LM}	$\sigma_{\ln L_x}$	A_{SZ}	B_{SZ}	$\sigma_{\ln Y_{SZ}}$
BCG	$\in (-1, 1)$	-	$0.47^{+0.24}_{-0.35}$	$0.32^{+0.17}_{-0.15}$	$1.59^{+0.33}_{-0.27}$	$0.75^{+0.19}_{-0.16}$	$0.86^{+0.18}_{-0.21}$	$1.51^{+0.28}_{-0.24}$	$0.36^{+0.13}_{-0.12}$
BCG	fixed	-	(0.0)	$0.36^{+0.18}_{-0.16}$	$1.58^{+0.32}_{-0.25}$	$0.70^{+0.19}_{-0.15}$	$1.06^{+0.09}_{-0.10}$	$1.33^{+0.21}_{-0.22}$	$0.33^{+0.12}_{-0.10}$
BCG	fixed	-	(0.5)	$0.32^{+0.17}_{-0.15}$	$1.60^{+0.33}_{-0.27}$	$0.74^{+0.19}_{-0.15}$	$0.81^{+0.13}_{-0.13}$	$1.54^{+0.24}_{-0.22}$	$0.37^{+0.11}_{-0.10}$
BCG	$\in (-1, 1)$	$0.2M_{HM}$	> -0.51 (at 16%)	$0.41^{+0.16}_{-0.17}$	$1.71^{+0.37}_{-0.29}$	$0.65^{+0.19}_{-0.14}$	$0.97^{+0.16}_{-0.19}$	$1.67^{+0.34}_{-0.27}$	$0.19^{+0.14}_{-0.09}$
BCG	fixed	$0.2M_{HM}$	(0.0)	$0.43^{+0.16}_{-0.18}$	$1.69^{+0.40}_{-0.28}$	$0.63^{+0.19}_{-0.14}$	$1.00^{+0.11}_{-0.12}$	$1.64^{+0.30}_{-0.27}$	$0.17^{+0.14}_{-0.08}$
BCG	fixed	$0.2M_{HM}$	(0.5)	$0.37^{+0.15}_{-0.15}$	$1.71^{+0.37}_{-0.29}$	$0.68^{+0.17}_{-0.13}$	$0.88^{+0.13}_{-0.13}$	$1.78^{+0.30}_{-0.26}$	$0.18^{+0.16}_{-0.09}$
X	$\in (-1, 1)$	-	$0.49^{+0.23}_{-0.33}$	$0.33^{+0.16}_{-0.15}$	$1.63^{+0.35}_{-0.26}$	$0.74^{+0.20}_{-0.15}$	$0.70^{+0.18}_{-0.21}$	$1.73^{+0.36}_{-0.34}$	$0.48^{+0.14}_{-0.11}$

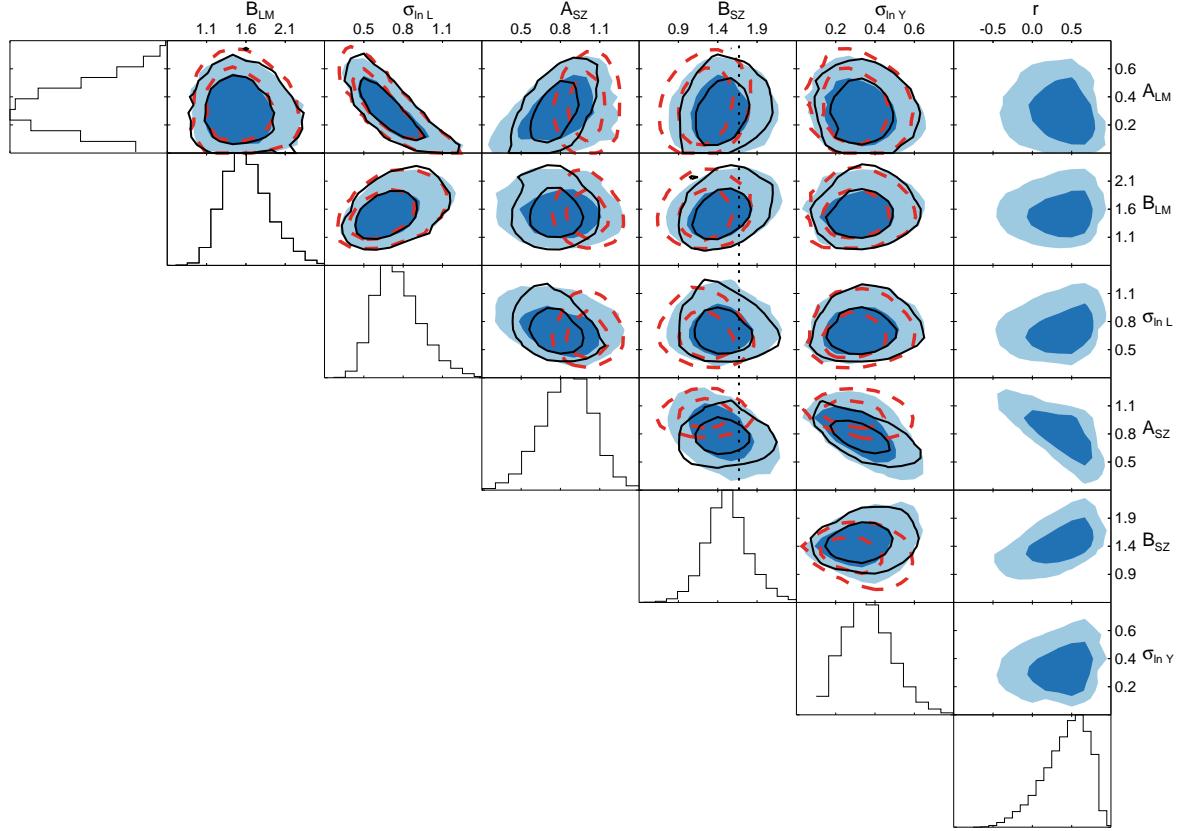


Figure 6.4: Marginalized posterior distributions of the parameters of the scaling models, shown as 1σ and 2σ confidence levels. Colour shaded contours are from the full joint likelihood fit including the correlated intrinsic scatter r with a uniform prior such that $-1 < r < 1$. Red contours indicate the marginalized confidence levels with $r = 0$. The black contours indicate the marginalized confidence levels with $r = 0.5$. The histograms show the marginalised distribution of each recovered scaling parameter when we allow r to vary. The vertical dotted line corresponds to the self-similar expectation of the $Y_{\text{SZ}} - M_{500}$ scaling slope.

between -1 and $+1$. The result is summarised in Table 6.1. Including correlated intrinsic scatters in Comptonization and luminosity at fixed mass results in a slope of $1.51^{+0.31}_{-0.22}$ in the $Y_{\text{SZ}} - M_{500}$ scaling relation, fully consistent with self-similarity. For the correlation between intrinsic scatters of luminosity and Comptonization we find $r = 0.47^{+0.24}_{-0.35}$. Approximately 90% of the posterior distribution prefers a positive correlation. The marginalized posterior distributions are shown in Figure 6.4. The correlation parameter, r , correlates the strongest with the SZ normalisation A_{SZ} (anti-correlation) but also with the slope B_{SZ} (positive correlation). Ignoring the correlation between intrinsic scatters of luminosity and Compton-Y at fixed mass (i.e., $r = 0$) results in a $Y_{\text{SZ}} - M_{500}$ scaling relation with a recovered slope of $1.33^{+0.21}_{-0.22}$, marginally shallower than what is expected from self-similarity (1.67) and the normalisation found is higher by 1σ . The uncertainties in the recovered scaling relation are lower when r is set to a fixed value (either 0.0 or 0.5). If one indeed uses the prior of ignoring the correlation in scatter completely (as would be the case using a method similar to that of B. C. Kelly (2007)), the bias in the normalisation of the $Y_{\text{SZ}} - M_{500}$ relation is on the order of $\sim 2\sigma$. A similar level of bias was found in our analysis of mock data sets in Section 6.3.3.1. We note that on applying a method similar to B. C. Kelly (2007) to measure $Y_{\text{SZ}} - M_{500}$ relation yield

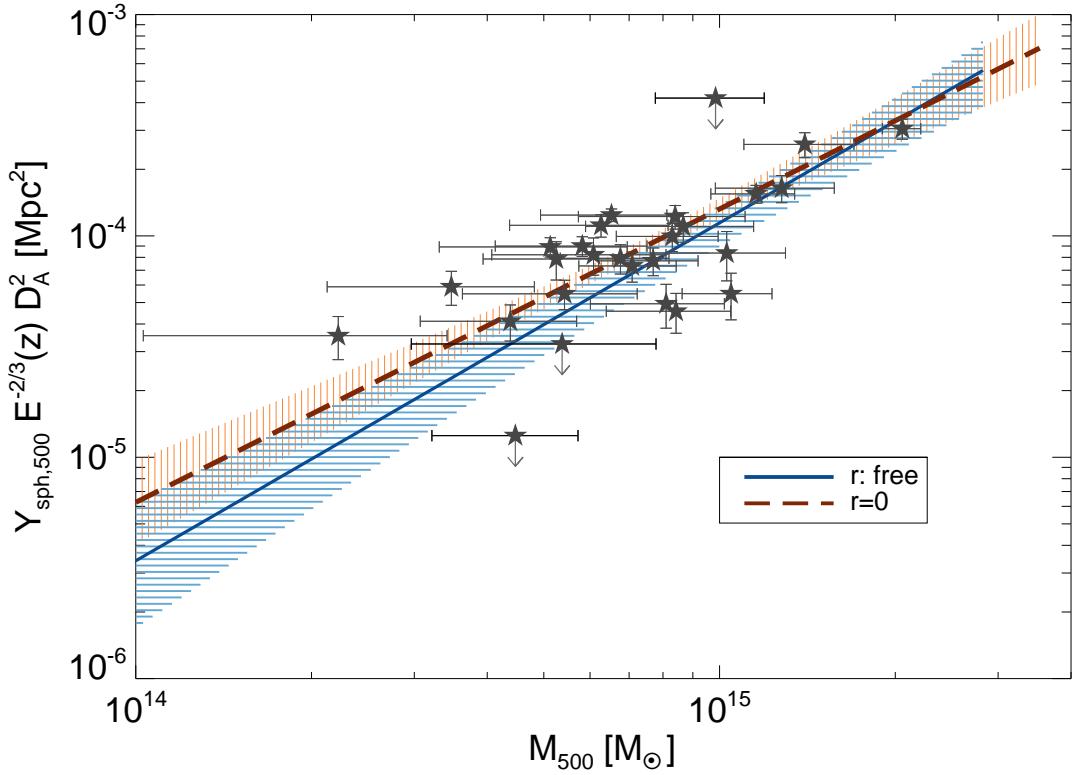


Figure 6.5: The best-fit $Y_{\text{SZ}} - M_{500}$ scaling relation including correlated intrinsic scatter is shown as the solid blue line. The red dashed line shows the fitted scaling relation assuming no correlated scatter. The red and blue line-filled regions denote 68% confidence. Assuming un-correlated intrinsic scatters in luminosity and Comptonization finds a higher normalisation and shallower slope for the $Y_{\text{SZ}} - M_{500}$ relation. The measured weak-lensing masses and integrated Comptonizations plotted here are same as for the eDXL sample shown in Figure 5.2.

constraints that are similar to our biased constraints from ignoring the correlation in intrinsic scatters of L_x and Y_{SZ} at fixed mass. We also fit the scaling using a generalisation of ordinary least square methods, such as the bivariate correlated errors and intrinsic scatters (BCES) method (Akritas and Bershady, 1996). Using the orthogonal relation with M_{500} as the independent variable (i.e., BCES with $Y_{\text{SZ}}|M_{500}$), the estimated relation gives the same normalisation as our biased estimate (i.e., when $r = 0$), but a shallower slope of 1.13, which is 1σ lower than our $r = 0$ fit and as well as the multiple Gaussian mixture fit (e.g., B. C. Kelly, 2007).

In Figure 6.4, we compare the results of this analysis to the case of no correlation in the intrinsic scatters. The marginal change in the $Y_{\text{SZ}} - M_{500}$ scaling parameters is illustrated in Figure 6.5, where it becomes evident that the bias from setting $r = 0$ is more prominent at the low-mass end. Table 6.1 summarises the results for fitting with different assumptions.

The normalization of the $L_x - M_{500}$ relation, A_{LM} , shows a strong anti-correlation with the intrinsic scatter $\sigma_{\ln L_x}$ in luminosity at fixed mass, with a Pearson correlation coefficient of -0.81. Our recovered normalisation of the $L_x - M_{500}$ relation is $0.32^{+0.17}_{-0.15}$, and the slope is $1.59^{+0.33}_{-0.27}$ for our fiducial analysis with varying r parameter. From Figure 6.4 and Table 6.1, we can observe that the $L_x - M_{500}$ relation

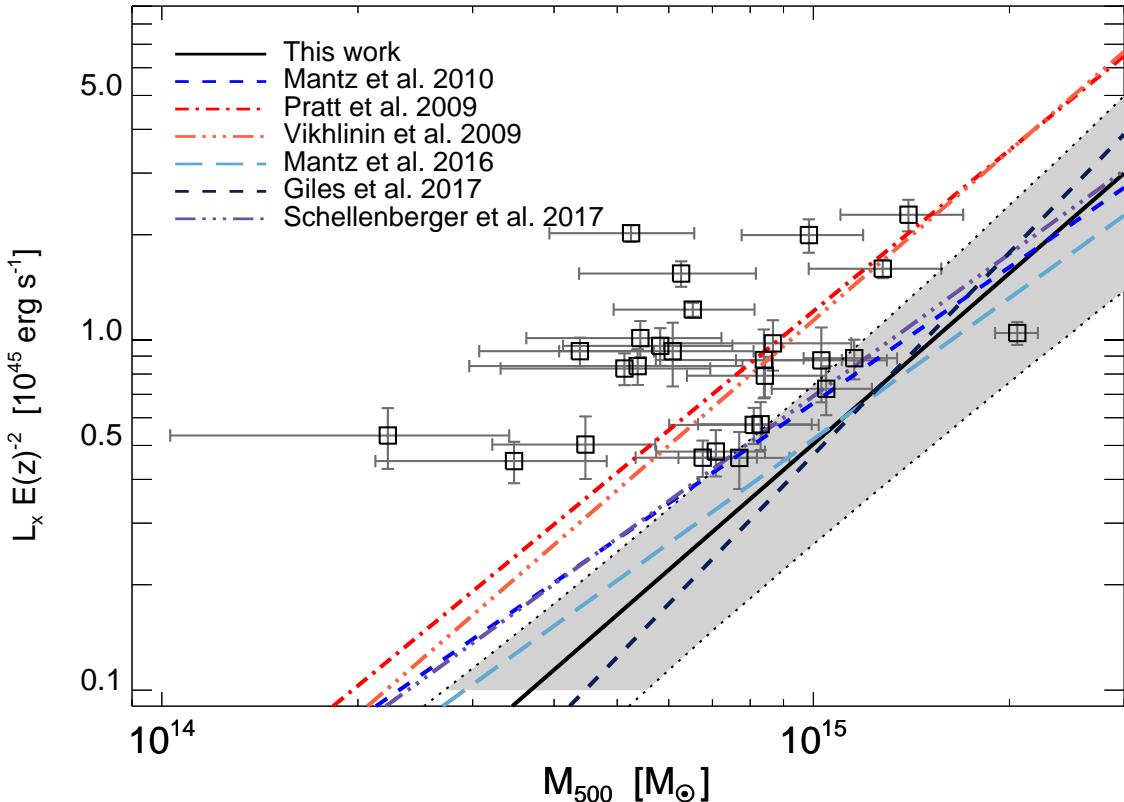


Figure 6.6: Luminosity-mass relation: The luminosities and our scaling relations are measured in the energy band 0.1 – 2.4 keV. The scaling relations are represented for the luminosity-mass relation for the same energy band. The grey shaded region represents the 68 % confidence level of our scaling relation. The measured values are generally up-scattered from the best-fit relation. This is due to the Eddington & Malmquist biases in the sample which is corrected for in the scaling relation determination through our Bayesian analysis. At the median mass of the sample, our constraints mean a correction factor of 1/3 to a biased estimate from generalizations of ordinary least square methods (e.g., BCES). Previous literature measurements of the luminosity-mass relation for the luminosity in the same energy band of 0.1 – 2.4 keV are plotted for comparison with our constraints. This is discussed in Section 6.6.1.

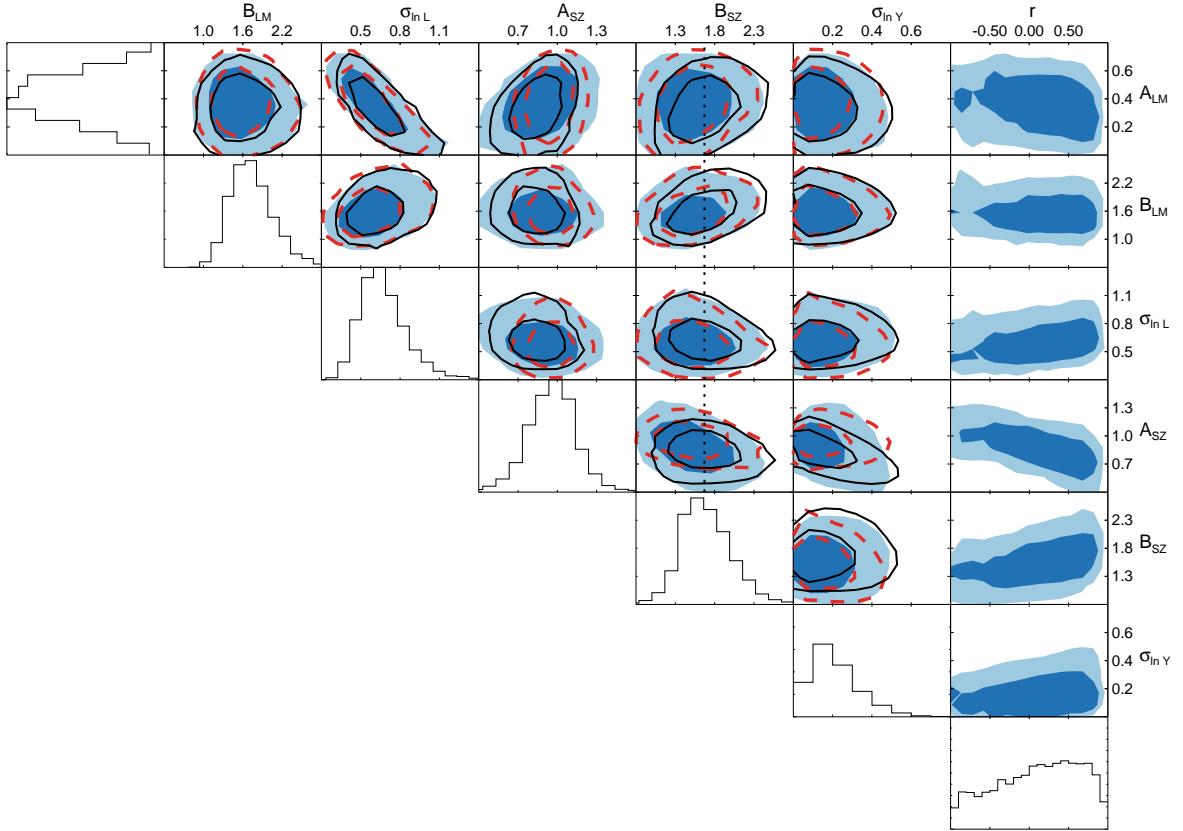


Figure 6.7: Same as Figure 6.4 but for $\sigma_{WL|HM} = 0.20M_{HM}$. The red contours are marginalized recovered confidence level for $r = 0.0$, and black contours for $r = 0.5$. The histograms show the marginalised distribution of parameters recovered from varying r .

constraints are unaffected by the correlation parameter r . This relation with its 68% confidence levels is shown in Figure 6.6. An estimate of the scaling relation obtained from a bivariate correlated errors and intrinsic scatters (BCES) method (Akritas and Bershady, 1996) gives a normalisation at M_{pvt} of $0.98 \times L_{\text{pvt}}$ (using both a bisector, and an orthogonal $L_x|M_{500}$ fit). This normalisation is $\sim 4\sigma$ away from our best-fit relation at the pivot mass of the eDXL sample.

A further discussion on the constraints obtained on the $Y_{\text{SZ}} - M_{500}$ and $L_x - M_{500}$ relations is given in Section 6.6.1. The results summarised here with r as a free parameter will be considered as our fiducial result.

6.4.2 Including uncorrelated intrinsic scatter in the weak-lensing masses

We include an intrinsic scatter term in the model by adding the scaling relation between true halo mass and weak-lensing mass given in Section 6.3.2.2. We marginalise over the true weak-lensing masses analytically for a Gaussian scattered lensing masses with a 20% dispersion. The implementation is detailed in Appendix A.2.3. We assume the bias in the lensing mass to be negligible. Due to the limited sample size, we forgo fitting and marginalisation of the percentage scatter of lensing masses

w.r.t halo mass ($\sigma_{WL|HM}$). Sereno and Ettori (2015) marginalised the scatter for a much larger sample and were able to constrain its value at approximately 20% log-normal scatter.

First, we consider the scenario with no correlation between the intrinsic scatters of L_x and Y_{SZ} (i.e., $r = 0$). While the normalisation of the $Y_{SZ}-M_{500}$ scaling is comparable to the case where the scatter in lensing mass was ignored, we find a steeper slope of $1.64^{+0.30}_{-0.27}$, which is a 1σ increase from 1.33 found in the previous section for $r = 0$. Fixing r to the mean value of 0.5 recovered from the previous subsection, the $Y_{SZ}-M_{500}$ slope increases marginally ($\sim 0.5\sigma$) from 1.64 to $1.78^{+0.30}_{-0.26}$, while the normalization decreases by $\sim 1\sigma$.

Finally, we carry out the analysis allowing all parameters, including r , to vary. The result is summarised in Table 6.1. We note that the data do not have the leverage to constrain r in this case, as is evident from Figure 6.7. We quote a lower limit of -0.51 for r with 84% of the distribution lying above this limit. The marginalised posterior distributions are shown in Figure 6.7 for all of the cases discussed here.

Assuming a 20% Gaussian intrinsic scatter in weak-lensing mass, the intrinsic scatters $\sigma_{\ln L_x}$ and $\sigma_{\ln Y_{SZ}}$ are both reduced, by 10% and 16% respectively. The normalisation, A_{LM} , being anti-correlated with $\sigma_{\ln L_x}$, increases by 0.4σ . These differences in the L_x-M_{500} relation with respect to constraints obtained in Section 6.4.1 are marginal, however, this trend of increased normalisation and decrease in intrinsic scatter is consistent with the demonstrated effect due to scatter in weak-lensing mass (Gruen et al., 2015).

6.4.3 Correlated intrinsic scatter: interpretation from residuals

We examine the distribution of residuals in $\log(L_x)$ and $\log(Y_{SZ})$ obtained for our best-fit scaling relations in Section 6.4.1. For each cluster in the sample that is a detection in APEX-SZ, the residual is computed at fixed lensing mass. We predict the 68% and 95% confidence regions using Monte-Carlo realisations.

For this purpose, we generate population of masses from Tinker mass-function and scatter the masses with the measurement uncertainty in the lensing mass. Additionally, we generate other observables including the luminosities using our best-fit scaling relations and covariance matrix (using the Equation 6.11 with $r = 0.47$). The observables are scattered with their measurement uncertainties. The procedure for generating cluster observables is similar to the one described Appendix A.3. For each cluster at a given redshift, we generate 6000 realisations of cluster observables that would make the selection of the eDXL sample. The distribution of the generated residuals and their 68% and 95% confidence levels are shown in Figure 6.8 for individual clusters. The measured residual for each cluster is indicated in the same. We combine all the measured residuals in the residual plane which is shown in the lower panel of Figure 6.8. The distribution of residuals show a positive alignment with a Pearson correlation coefficient of 0.73. The generated residuals from Monte-Carlo simulations are combined together in the residual plane for 24 clusters. The model prediction of 68% and 95% confidence levels of the residuals in Figure 6.8 show that the distribution of residuals are consistent with our model prediction.

Next, we repeat the exercise for the $r = 0$ recovered mean scaling relation. The residuals for the best-fit scaling relations are plotted in Figure 6.9. The predicted 68% and 95% confidence levels for $r = 0$ are represented as contours, where the 95% confidence encompasses all of the residuals. We

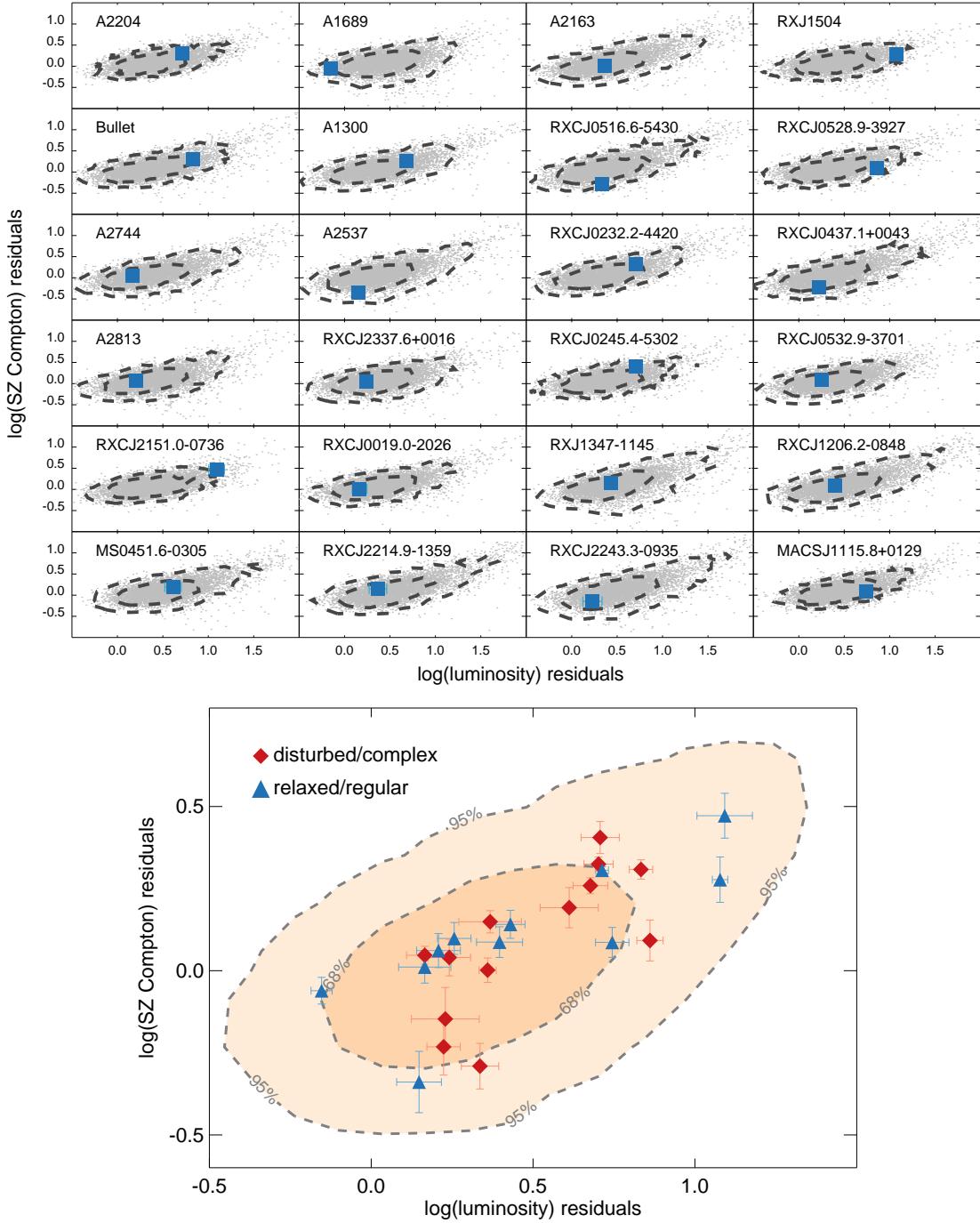


Figure 6.8: *Top panel:* Luminosity and Comptonization residual data (blue square) after subtracting the best-fit $Y_{\text{SZ}}-M_{500}$ and L_x-M_{500} scaling relations for each of the 24 clusters that were detected with APEX-SZ. The contours show the 1σ and 2σ confidence level prediction from random realisations of simulated residuals (grey points) for the best-fit intrinsic covariance model. *Lower panel:* distribution of residuals of 24 clusters, showing a positive correlation. The Pearson correlation coefficient in residuals is 0.73. The contours represent the 1σ and 2σ confidence levels obtained by combining the distributions shown in the top panel.

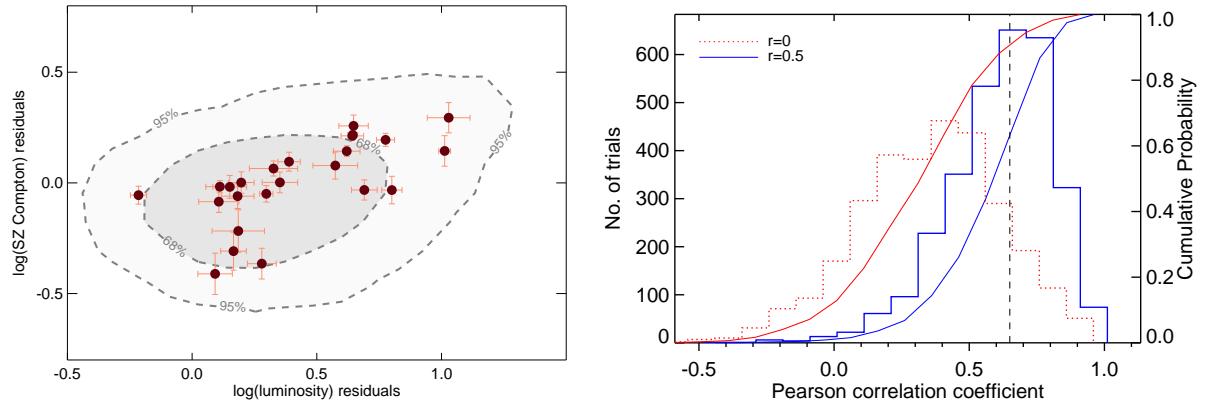


Figure 6.9: *Left:* Same as Figure 6.8, however, the residuals correspond to best-fit relations for the $r = 0$ covariance model. *Right:* Distributions of Pearson correlation coefficients in simulated realisations of 24 cluster residuals with un-correlated ($r = 0$) and correlated ($r = 0.5$) intrinsic scatters. The vertical dashed line indicates the measured Pearson correlation coefficient (0.65) in the residual data.

find a positive alignment in the residuals with a Pearson correlation coefficient of 0.65. This being still positively aligned, we use 3000 mock random realisations for the $r = 0$ model prediction of 24 cluster residuals and compute the Pearson correlation. We iterate the process with $r = 0.5$. Both distributions of Pearson coefficients are shown in Figure 6.9. Finding a strong correlation in the residual appears to be less likely when there is no intrinsic correlation. However, it does not altogether rule out the $r = 0$ value as 10% of the distribution lies above 0.65. This reflects our weak constraint on r .

6.5 Robustness and limitations of the analysis

We now examine the robustness of the scaling relation analysis from the last section to potential modelling variations (Section 6.5.1, 6.5.2, 6.5.3 and 6.5.4), systematic errors (Section 6.5.5 and 6.5.6) and data selection choices (Section 6.5.7).

6.5.1 Redshift evolution of scaling relations

We assumed a self-similar evolution in the $Y_{\text{SZ}} - M_{500}$ relation for our analysis. In order to check the validity of this assumption for our measurements, we split the sample into three redshift ranges: 0.15 to 0.22, 0.27 to 0.31, 0.31 to 0.55, consisting of five, 15 and seven clusters respectively. We then fit the joint scaling relations with all parameters but for A_{SZ} fixed to best fit values from Section 6.4.1. The recovered medians and 68% confidence levels of A_{SZ} are $0.80^{+0.15}_{-0.12}$, $0.81^{+0.08}_{-0.08}$, $0.90^{+0.16}_{-0.13}$ in the low, median and high redshift bins respectively. The normalisation is consistent within statistical errors in all three redshift ranges of the sample.

We change the redshift evolution slope from self-similarity ($-2/3$) to 0 based on the best fit value of the slope of the redshift evolution found in Sereno and Ettori (2015), even though they do not find this deviation from self-similar evolution to be greater than 68% confidence level. Assuming zero slope for the redshift evolution (i.e. no redshift evolution in the $Y_{\text{SZ}} - M_{500}$ relation) increases

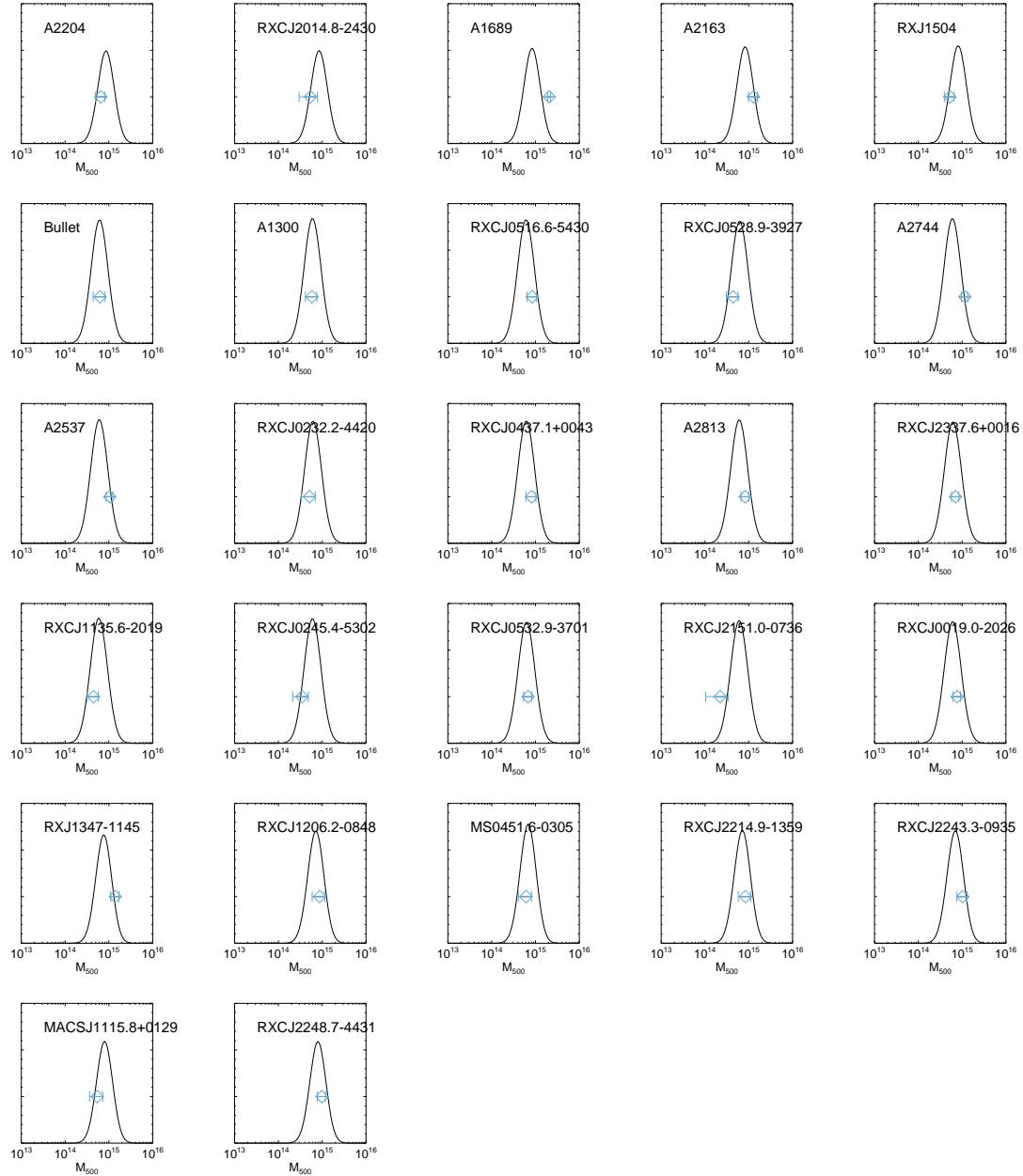


Figure 6.10: For each cluster we compare the measured mass estimate to the probability distribution function prediction from our modelling.

the normalisation by 9%. The other parameters, including the correlation in intrinsic scatters, are consistent with results obtained in Section 6.4.1.

6.5.2 Treatment of completeness of the eDXL sample

We check the effect of varying completeness of the sample in the luminosity-redshift plane. Our model assumes high completeness for the sample, which enables us to apply an analytical integration of the normalisation of the likelihood model described in Section 6.3. *A posteriori*, we predict the cluster number count using our best-fit values of the L_x - M_{500} relation from our fiducial result, the eDXL sample selection function and the same mass function used for our modelling. Our model predicts a total number of clusters to be 27. We re-compute the prediction of the cluster number counts, this time considering the completeness in the luminosity-redshift plane. Using the same model as before, we predict a sample size of 24 clusters. This suggests an average completeness $\approx 89\%$.

6.5.3 Consistency check of measured masses and model prediction

We examine the best-fit L_x - M_{500} scaling relation prediction of the mass distribution for each redshift of the sample. We compare this probability distribution function with the measured weak lensing mass and check whether there is any inconsistency between the two. The model prediction of mass distribution for each cluster in the sample is shown in Figure 6.10 along with the measured weak-lensing mass estimate for comparison. Overall, there appears to be good agreement between the prediction and measurements. We do not find any hint of inconsistency.

6.5.4 Additional covariances in the scatters of mass observables

Our scaling models treat the intrinsic scatter in lensing mass at fixed mass as being independent of scatter in the thermodynamic observables. Dark matter simulations (Shirasaki, Nagai and Lau, 2016; Angulo et al., 2012) found a correlation in the range of 0.6-0.9 in intrinsic scatter of integrated Comptonization and weak-lensing mass. Penna-Lima et al. (2016) were unable to constrain this correlation for a sample with a size similar to the one used in this work, and given our lack of statistical power to constrain any more free parameters, we have ignored this correlation in the present work. Since our dominant source of bias is expected to come from the selection in luminosity, any correlation in Compton-Y and lensing mass (intrinsic) scatters would be a second order effect.

We note that the scatter in X-ray luminosities is sensitive to the physical processes near the core of the cluster. The scatter in weak-lensing mass, however, is more affected by averaging over projected structures over a larger area. Thus, we expect a weaker correlation in the scatters of L_x and weak-lensing masses. The correlation in the scatters of lensing mass with luminosity was found to be 0.41 from dark matter simulations (Angulo et al., 2012).

6.5.5 Impact of mis-centring gNFW profile

In Section 6.4, we analysed the scaling relations for measurements of Y_{SZ} with gNFW model centred on BCG. Alternatively, in Section 3.2.3 we obtained X-ray centroids for eDXL clusters from ROSAT.

Here, we use the re-computed Compton-Y's with these centroids as described in Section 5.1.2 and fit the joint scaling relations with these measurements.

When using X-ray centroids rather than optical (BCG) centroids to determine the integrated Comptonizations, the resulting scaling relation parameters differ only marginally from the results of Section 6.4.1. The recovered medians and 68% confidence levels are given in the last row of Table 6.1. The main difference is an increase in the intrinsic scatter of integrated Comptonization at fixed mass. This is indeed expected considering that the accuracy of the centroid estimation from ROSAT is limited. We find that all deviations in the scaling parameters are within 1σ confidence with respect to the ones obtained in Section 6.4.1. This is not surprising considering that 80% of the eDXL clusters show less than half of the APEX-SZ beam offset in the two centres.

The angular distance of the offset between optical and X-ray centroids is plotted for 27 eDXL clusters in Figure 5.4. The cluster RXCJ1135.6–2019 shows the largest offset (approximately 2.5 arc minutes). The estimated $Y_{\text{sph},500}$ for this cluster at the X-ray centre produces a detection, but a non-detection at the optical centre as already discussed in Section 3.2.3. We provide further details on the impact of this cluster on the scaling relation in Section 6.5.7.

6.5.6 Systematic uncertainty in weak-lensing mass estimates

As discussed in Section 3.3 of Chapter 3, the weak-lensing mass estimates have a combined systematic uncertainty of approximately 8%. This level of systematic uncertainty is comparable to those reported in other weak-lensing estimates (e.g., D. E. Applegate et al., 2014), although the contributions and different sources of systematics are not identical. In our modelling, we have not accounted for this additional uncertainty in the lensing masses. To the first order, shifting the masses by $\pm 8\%$ can be propagated via the scaling model as a shift in the normalisation A_{SZ} , δA_{SZ} , relative to A_{SZ} for a fixed slope ($B_{\text{SZ}} = 1.51$). The relative uncertainty, $\frac{\delta A_{\text{SZ}}}{A_{\text{SZ}}}$, comes out to be +0.14 and -0.10 for 8% underestimation and overestimation respectively. These uncertainty levels are comparable to the confidence levels obtained for the normalisation parameter. We note that this estimate ignores the effect of the mass function, however, this is expected to be a secondary effect.

6.5.7 Treatment of outliers

We explore the stability of our constraints to potential outliers. We identify a couple of such cluster measurements and drop them completely or replace them with alternate measurements from different modelling assumptions or literature. In the following series of tests, we find that the cluster mass measurement of A1689 has the strongest impact on the $L_{\text{x}} - M_{500}$ relation at $\sim 1\sigma$ level. This, in turn, affects our $Y_{\text{SZ}} - M_{500}$ relation at $\sim 1\sigma$. Since our cluster sample is X-ray selected and the variations in the constraints are at most $\sim 1\sigma$, we do not have a reason to exclude this cluster from our analysis. We describe these details below.

In our sample, the weak-lensing mass estimate of A1689 not only places it as the most massive cluster, it also has the most precise measurement in mass. From Figure 6.6, we observe that A1689 is the only cluster that is down-scattered in the $L_{\text{x}} - M_{500}$ relation. We investigate how our scaling relation constraints are driven by this cluster. We drop it altogether from our Bayesian fitting and re-perform the joint analysis with 26 cluster measurements. The most notable change we observe a

decrease of almost 0.7σ in the intrinsic scatter of the $L_x - M_{500}$ relation, and this being anti-correlated with the A_{LM} , increases the latter by $\sim 1.1\sigma$. We observe an increase of $\sim 0.8\sigma$ increase in the normalisation of the $Y_{SZ} - M_{500}$ relation, and the slopes B_{SZ} , B_{LM} are estimated to be steeper by $\sim 0.7\sigma$. The correlation parameter r is lowered marginally by 0.4σ , with over 77% of the distribution still preferring a more positive or greater than zero correlation. In order to disentangle the effect of differences in the $L_x - M_{500}$ relation impacting the changes in the $Y_{SZ} - M_{500}$ relation, we fix the $L_x - M_{500}$ relation and the r parameter to the best-fit values obtained from Section 6.4.1 and re-fit the $Y_{SZ} - M_{500}$ relation by including and excluding the cluster in our analysis. We recover identical scaling relation parameters for the $Y_{SZ} - M_{500}$ relation irrespective of whether A1689 is included or excluded in the analysis. We infer that the steeper B_{SZ} , and higher A_{SZ} are found mainly due to the impact the mass measurement of A1689 has on the $L_x - M_{500}$ relation.

The weak-lensing mass estimates across the sample quoted in this paper have used the Bhattacharya et al. (2013) $c - M_{200}$ relation to break the degeneracy in concentration and mass. The intrinsic scatter in the relation has been ignored for the purpose of obtaining the mass estimates. In the NFW fitting with two free parameters without using the $c - M_{200}$ relation, we find the mass estimate to be 2σ lower with $M_{500} = 17.2^{+1.4}_{-1.4} \times 10^{14} M_\odot$. A detailed study of the cluster using strong lensing, weak-lensing as well as triaxiality information finds a cuspier core and a lower mass (Umetsu, Sereno et al., 2015). There is general consensus that the cluster is elongated along the line of sight (Sereno, Ettori and Baldi, 2012; Sereno, Umetsu et al., 2013; Limousin et al., 2013; Umetsu, Sereno et al., 2015). We replace our lensing mass estimate for this cluster with the spherical mass estimate obtained from the detailed analysis in Umetsu, Sereno et al. (2015) of $M_{500c} = 12.6 \pm 1.9 \times 10^{14} M_\odot$ and re-measure the corresponding $Y_{SZ,500}$ in the same aperture. We perform the joint fitting with the updated mass and integrated Comptonization for A1689 and fit for all the seven free parameters. The recovered scaling relations are indifferent to dropping A1689 from the fit with the exception that the intrinsic scatter in SZ is lowered to 28% (decreases by 7%). We find the slope, B_{SZ} , to be $1.73^{+0.34}_{-0.32}$, which is steeper than found in Section 6.4.1 by 0.6σ . We fix the $L_x - M_{500}$ relation to our best fit results as done previously and find the difference in slope and normalisation reduce to 0.45σ and 0.3σ respectively. In conclusion, the marginal changes in the $Y_{SZ} - M_{500}$ relation mainly occur due to the $L_x - M_{500}$ relation preferring higher normalisation, steeper slope and lower intrinsic scatter when A1689 is not as massive.

In addition, we investigate the non-detection in SZ for RXCJ1135.6 – 2019. As discussed in Section 6.5.5, the SZ measurement at the BCG centre is a non-detection versus a 3σ detection at the X-ray centre. We drop this cluster and re-fit the scaling relations by allowing the r parameter to float. The parameter that shows the most significant change is B_{SZ} preferring a shallower value of $1.38^{+0.24}_{-0.26}$. This change is well within the confidence levels (0.5σ) of our fit results in Section 6.4.1. We replace the Y_{SZ} measurement from BCG centre to X-ray centre and re-fit the relations and find the slope, B_{SZ} , to be $1.42^{+0.27}_{-0.23}$ which remains shallower by 0.42σ while rest of the scaling parameters are almost identical to the results obtained in Section 6.4.1. The impact of this cluster measurement is marginal and at most on the slope of the $Y_{SZ} - M_{500}$ relation.

As a final step, we replace the mass of A1689 by the value obtained in Umetsu, Sereno et al. (2015), the corresponding SZ measurement for this cluster within this aperture as done earlier and replace the SZ measurement of RXCJ1135.6 – 2019 with the one obtained using the X-ray centroid. We jointly fit the seven free scaling parameters and find that the slope B_{SZ} is $1.59^{+0.31}_{-0.28}$ which effectively implies that the two outliers cancel each other's effects on the recovered slope. The other scaling parameters, namely, A_{SZ} and A_{LM} increase by 0.8σ , 1σ respectively. The intrinsic scatter $\sigma_{\ln L_x}$ and $\sigma_{\ln Y_{SZ}}$ are

lowered by 1σ and 0.5σ , respectively. The correlation parameter r is lowered by 0.3σ . The values of all scaling parameters except for B_{SZ} are similar to the ones obtained while fitting with lower mass measurement of A1689 and leaving RXCJ1135.6 – 2019 as a non-detection. Fixing the $L_{\text{x}}-M_{500}$ relation to the best fit values in Section 6.4.1 reduces the 0.8σ difference in A_{SZ} to 0.3σ with respect to our best-fit result. In conclusion, the most dominant effect on the scaling relation comes from A1689 which has been shown to mainly influence the $L_{\text{x}}-M_{500}$ relation.

6.6 Discussion

In this Section, we discuss the implications of our constraints on the scaling relations and the correlation coefficient of intrinsic scatters in Y_{SZ} and L_{x} . First, we compare our recovered relations to previous work, which is summarised in Section 6.6.1. In Section 6.6.2, we discuss the origin of the correlation in intrinsic scatters and finally, we discuss the implications for future cosmological studies in Section 6.6.3.

6.6.1 Comparison to literature

We compare the scaling relation fit results from Section 6.4.1 to other published works in Figures 6.11 and 6.12. The scaling between M_{500} and $Y_{\text{SZ},500}$ has been inferred in several previous publications using weak-lensing masses (e.g., Marrone et al., 2012; Hoekstra, Mahdavi et al., 2012; Hoekstra, Herbonnet et al., 2015; Sereno and Ettori, 2015) and using X-ray mass proxies (e.g., Andersson et al., 2011; Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al., 2014; A. Mantz, Allen, Morris et al., 2016). The comparison with the former set of works is shown in Figure 6.11 and with latter is shown in Figure 6.12. In both cases, there is general agreement in the scaling relation in the mass ranges of our eDXL sample with three exceptions (Marrone et al., 2012; Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al., 2014; de Haan et al., 2016). There are numerous differences other than the measurement of mass itself, such as, in the sample selection and follow-up statistics between this work and the above mentioned literature. We note that at lower mass end there is mild disagreement with Marrone et al. (2012) and de Haan et al. (2016). At higher mass end there is some stark disagreement with Marrone et al. (2012), de Haan et al. (2016) and Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al. (2014). The former two publications (i.e., Marrone et al., 2012; de Haan et al., 2016) estimate a significantly steeper slope for the mass scaling of the integrated Comptonization than the self-similar value for the slope. The relation given in Marrone et al. (2012) used for comparison is the one including A383 in their sample, and while dropping it they find a shallower slope that is consistent with self-similar slope and this scaling relation is more consistent within our uncertainties. In addition to this, their lensing masses from Okabe, Takada et al. (2010) have increased by 9 – 20% according to Okabe and G. P. Smith (2016). For comparison with de Haan et al. (2016), we use their $Y_{\text{x}}-M_{500}$ relation, where Y_{x} is the X-ray equivalent of Y_{SZ} . Their constraints on the $Y_{\text{x}} - M_{500}$ relation are a byproduct of a full cosmological analysis. The disagreement with Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al. (2014) is due to their relation preferring a higher normalisation. This discrepancy could arise due to differences in the type of mass estimates used for the calibration, and in the approaches adopted for mitigating selection biases. In Figure 6.12, we also show the scaling relation prediction from numerical simulation of Sembolini et al. (2013) which indicates a higher normalisation than our constraints. A summary of slope values obtained from different literature and

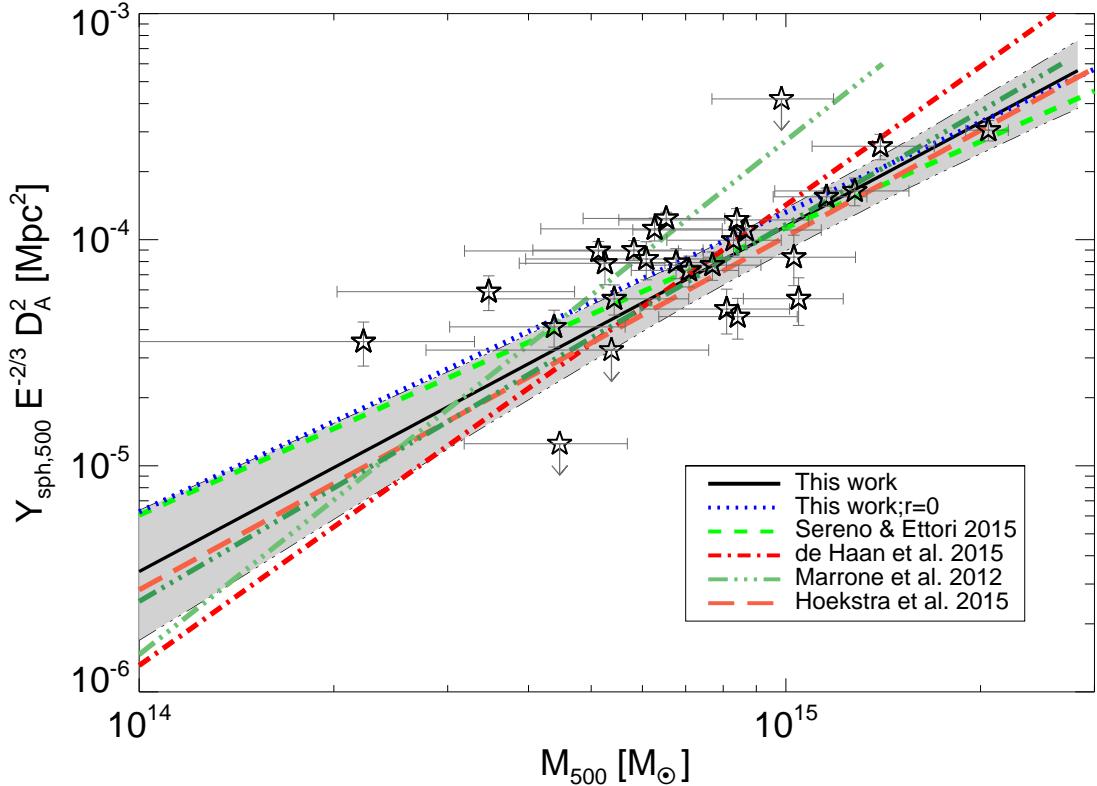


Figure 6.11: The best-fit Y_{SZ} – M_{500} scaling relation is shown as the solid black line, with the 68% confidence level indicated by the grey shaded region. The constraints shown here were obtained by ignoring the weak-lensing intrinsic scatter. The scaling relations found in other works that used weak-lensing mass to calibrate total mass within R_{500} are overplotted.

this work is presented in Figure 6.13. For comparison, we also include results using R_{2500} . We note that our recovered slope is in disagreement with Czakon et al. (2015), who find a shallower slope for the $Y_{\text{SZ},2500}$ – $M_{x,2500}$ relation. Whereas, our biased estimate of slope when $r = 0$ agrees well with the shallower estimate.

The ROSAT luminosities used in our work are model dependent (Section 3.2.3), and suffer from low signal to noise and poor de-blending capabilities, in particular for resolving out AGNs. Therefore, they are used in this work as a purely phenomenological description for the selection and not much importance should be given to possible discrepancy with other, more involved works. Nevertheless, we compare the L_x – M_{500} relation obtained in this work with other literature and they are shown in Figure 6.6. Whenever possible we use scaling relations obtained directly from the ROSAT luminosities in a similar band. For comparing with Giles, Maughan, Dahle et al. (2017), their relation measured for the ROSAT luminosities were used. To compare with Vikhlinin, Burenin et al. (2009), their L_x – M_{500} relation was modified to give the relation for luminosities in the same energy band as used in this work, i.e., 0.1 – 2.4 keV. Similarly, the relation used for comparing with Pratt et al. (2009) was for the luminosity in the 0.1-2.4 keV energy band (from Table B.2 in Appendix B of Pratt et al., 2009). Our L_x – M_{500} relation is in good agreement with A. Mantz, Allen, Ebeling et al. (2010), A. Mantz, Allen, Morris et al. (2016), Giles, Maughan, Dahle et al. (2017), Schellenberger and Reiprich (2017a).

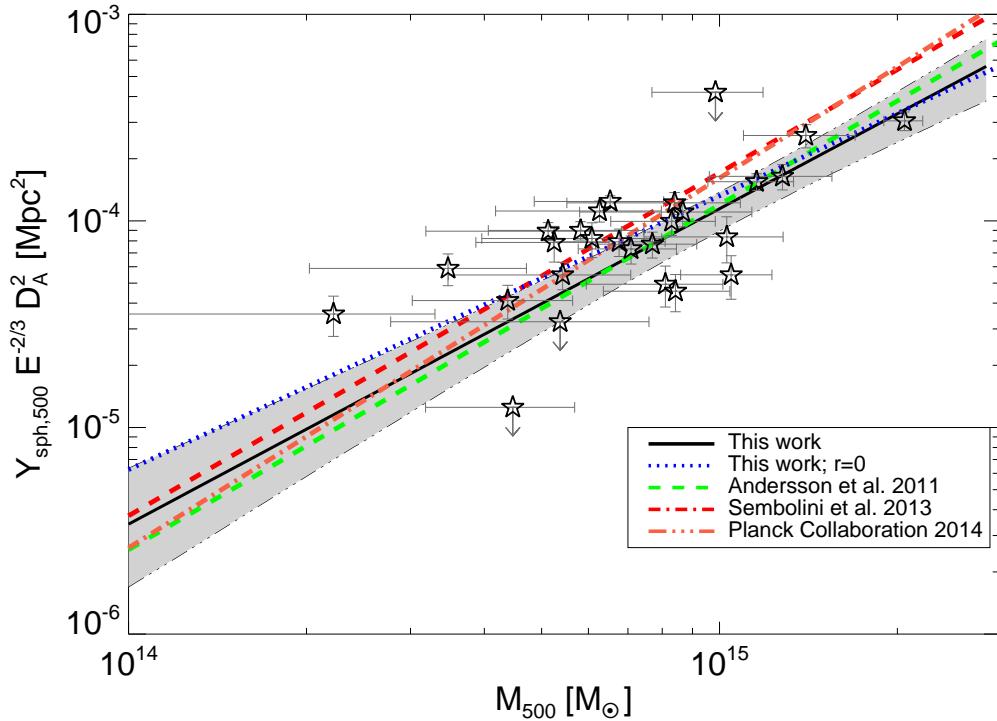


Figure 6.12: Same as Figure 6.11 but comparing to previous works using mass proxies other than weak-lensing mass. The prediction from simulation work of Sembolini et al. (2013) is also shown for comparison.

All of these work have accounted for their sample selection in a manner equivalent to ours. We observe that there is a notable offset in our measured relation with respect to the best-fit of Vikhlinin, Burenin et al. (2009) and Pratt et al. (2009). Possible sources of this offset could arise due to different approaches used by these authors for accounting for the sample selection biases. The difference also could be due to different radius adopted for the luminosity measurements. In addition to these differences, the masses used in these publications for the calibration were primarily done with either hydrostatic masses or from Y_x with the exception of A. Mantz, Allen, Morris et al. (2016), who used some weak lensing masses in addition. However, the slope is found to be consistent with all of the previous work (e.g., Vikhlinin, Burenin et al., 2009; Pratt et al., 2009; A. Mantz, Allen, Ebeling et al., 2010; A. Mantz, Allen, Morris et al., 2016; Sereno and Ettori, 2015; Giles, Maughan, Dahle et al., 2017; Schellenberger and Reiprich, 2017a) within 68% confidence level of the slope (B_{LM}). The intrinsic scatter in L_x , $\sigma_{\ln L_x}$, recovered is $0.75^{+0.21}_{-0.15}$ which is consistent with the findings of Giles, Maughan, Dahle et al. (2017), whereas it is higher than what was found by A. Mantz, Allen, Ebeling et al. (2010), A. Mantz, Allen, Morris et al. (2016), Vikhlinin, Burenin et al. (2009), Schellenberger and Reiprich (2017a).

6.6.2 Origin of correlation in intrinsic scatters of L_x and Y_{SZ} at fixed mass

Our constraint on the correlation parameter (r) suggests a positively correlated scattering of luminosity and SZ with 84% of the marginalised distribution of this correlation lying above 0.12. In Section 6.4.3, we confirmed the consistency of our modelling by examining the residual data. In Figure 6.8,

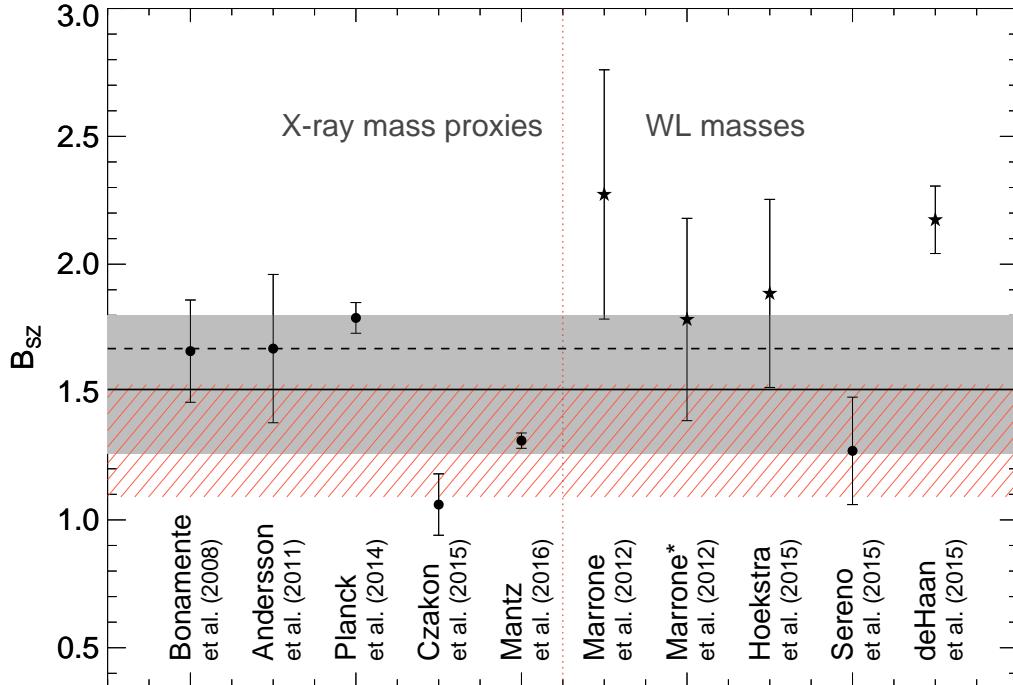


Figure 6.13: Comparison of constraints on the slope of $Y_{\text{SZ},\Delta} - M_\Delta$ scaling relation between the literature and this work. The black solid line is our best-fit slope value and the corresponding 1σ confidence is represented as the grey shaded region. The dashed line represents the self-similar slope value. The 1σ confidence on the slope constraints obtained with $r = 0.0$ is indicated by the line filled region. We inverted the slope constraint for certain works which originally provided results for $1/B_{\text{SZ}}$; these are represented as stars. The asterisk on Marrone et al. (2012) represents their result without A383.

we indicate cluster morphologies as relaxed or disturbed. The classification was determined from using the centroid shift w parameter in units of R_{500} from weak-lensing estimate. The procedure for centroid shift calculation closely follows the method given in Wei\ss{}mann et al. (2013). We make a cut in w at 0.0088. The clusters with centroid shifts less than this value are classified as relaxed or regular. The rest are determined to be disturbed. From the distribution of the residuals, we do not find an indication of them being morphologically segregated.

Since both SZ and luminosity probe the same ICM, they are expected to be correlated. Soft-band luminosity for massive clusters are essentially tracers of gas mass density and the Comptonization is sensitive to the product of gas mass density and temperature. The correlation in scattering of luminosity and SZ observable could arise more due to fluctuations in gas mass fraction, as found in the pre-heating model of simulated clusters by Stanek, Rasia et al. (2010). For the pre-heating scenario they report the correlation in intrinsic scatters to be as strong as 0.88 between the gas mass fraction and the Comptonization and followed by a correlation of 0.78 in scatters of integrated Comptonization and bolometric luminosity. In this case, the scatter in luminosity and SZ is largely driven by the scatter in the gas mass fraction. While Stanek, Rasia et al. (2010) used bolometric

luminosity that has stronger temperature dependence than the soft-band luminosity of the ROSAT survey for massive clusters, we can extrapolate that the scatter in soft-band luminosity is indeed dominated by gas mass fraction. Furthermore, more recent work in Truong et al. (2016) in their simulation with AGN contributions find a correlation of 0.67 in intrinsic scatters of Y_x and gas mass and 0.53 in scatters of bolometric luminosity and Y_x .

Our constraint on r is consistent with these predictions from simulations and is likely to originate from the scatter of gas mass in the ICM.

6.6.3 Impact on cluster based cosmological studies

The mass-observable calibration plays a decisive role in cosmological studies of the cluster population. Despite recent progress, it still dominates the error budget of current analyses (Planck Collaboration, Ade, Aghanim and al., 2016a). We consider here the implications of our new estimate of the $Y_{\text{SZ}} - M_{500}$ relation on future SZ cluster surveys such as SPT-3G (Benson, Ade et al. (2014)), AdvACT (Henderson et al. (2016)) and the Simons Observatory¹.

In this context, we focus on a simplified model for the SPT-3G 2500 deg² survey. We take a fixed detection threshold of $Y_{\text{SZ},\text{lim}} = 0.88 \times 10^{-5} \text{ Mpc}^2$, chosen to match the predicted SPT-3G number counts (5000 clusters) of Benson, Ade et al. (2014). We ignore measurement errors for simplicity.

We first estimate the effect of the statistical errors of our $Y_{\text{SZ}} - M_{500}$ on the constraining power for such a survey. To do so, we compute the predicted number of clusters for the same cosmology but using the scaling relation parameters of each point in our MCMC chain. The resulting distribution is extremely broad with a 68% confidence interval spanning the range 1000 - 7800. Although our data set is one of the largest available samples with detailed SZ measurements and reliable lensing masses, the combined error of all the free parameters in our modelling outweighs the statistical errors for such a survey (essentially Poisson errors given the large survey volume). This shows that a tremendous calibration effort is required in order to really take advantage of the next generation of SZ surveys. A common procedure to circumvent this problem is to improve the precision of the mass calibration by resorting to constraints derived from X-ray observations (de Haan et al., 2016; Planck Collaboration, Ade, Aghanim and al., 2016a). However, the propagation of X-ray information to the SZ observables also involves a number of modelling assumptions, whose systematic effects are likely to dominate the error budget once larger cluster samples are available.

Now, focusing on the average number of detected clusters, no strong claim can be made on revised expectations for the upcoming surveys since, within the uncertainties, our measurements are compatible with several of the previously available calibrations. However, our work revealed the importance of some modelling assumptions on the final results, in particular the impact of the correlation between the SZ Comptonization and X-ray emission at fixed mass on scaling relations relying simultaneously on X-ray and SZ observables. In the case of our imaginary survey, the predicted number of clusters would raise from 5000 to ~ 21000 if, instead of our best fit scaling relation, one used the best fit obtained without including the covariance in the scatter of L_x and Y_{SZ} . We stress here that this shift purely originates from differences in the modelling assumptions and not from the statistical uncertainties inherent to our sample. The simulations described in Appendix A.3 show, for instance, that even with much smaller statistical errors, similar deviations in the best fit parameters are expected

¹ <http://simonobservatory.org/>

if the covariance between L_x and Y_{SZ} is neglected while modelling a correlated population. This demonstrates the necessity to always consider such a covariance term when X-ray and SZ observations are mixed. A detailed understanding of the nature of this covariance from both numerical simulations and other observations would definitely help in constraining the correlation and add relevant priors to the analysis of future surveys.

6.7 Summary & Conclusions

We study the statistical correlations between three galaxy cluster observables: the integrated Comptonization, Y_{SZ} , the X-ray luminosity, L_x , and the weak lensing mass, M_{500} . Special attention is given to the sample selection bias and the correlation between the intrinsic scatter of X-ray luminosity and SZ Comptonization.

1. We construct a complete sample of 30 clusters (eDXL) from the ROSAT X-ray survey catalogues. We obtain SZ effect measurements from the APEX-SZ experiment and lensing follow-up observations for 27 of these clusters. The global completeness of this sample is $\sim 90\%$.
2. We present and implement a Bayesian analysis method that allows to control the sample selection bias while fitting for the mass-observable scaling relations. In the statistical formalism, we account for the impact of the measurement uncertainties of cluster properties, the shape of the cluster mass function, intrinsic covariances of cluster observables at fixed mass and the selection function of the cluster sample.
3. We jointly constrain the $Y_{\text{SZ}}-M_{500}$ and L_x-M_{500} scaling relations, while accounting for the correlation between the intrinsic scatter of Y_{SZ} and L_x observables. The constraint on the correlation coefficient, $r = 0.47^{+0.24}_{-0.35}$, is weak, but suggests a positive correlation, with 84% of the marginalised distribution above $r = 0.12$. The slope of the $Y_{\text{SZ}}-M_{500}$ relation is found to be consistent with the self-similar expectation. At the current precision level, we find a general consistency in the relation with previous work. However, we note a marginally lower normalisation than given by Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al. (2014) at the pivot mass value of our sample. Our L_x-M_{500} relation is measured using ROSAT luminosities for the sole purpose of accounting for the sample selection function. We find a slope for the L_x-M_{500} relation that is steeper than the self-similar slope value as noted by previous literature. Our constraints on the relation is found to be consistent with previous literature that account for sample selection in a manner equivalent to the one used in this work.
4. We perform post-predictive checks that yield a consistent picture of the robustness of our modelling of the completeness of the sample, the measured correlation in intrinsic scatters and the mass distribution. We vary our modelling assumptions on the redshift evolution of the $Y_{\text{SZ}}-M_{500}$ relation, and the gNFW centroids for measuring Y_{SZ} . We find our constraints to be stable within the confidence levels. In addition, we re-measure the scaling relation parameters by treating a couple of cluster measurements as possible outliers in our modelling. Among these, we find that removing the A1689 measurements altogether or replacing its mass by a lower estimate from literature to dominantly affect the calibration of the L_x-M_{500} relation at $\sim 1\sigma$ and thereby also affects the $Y_{\text{SZ}}-M_{500}$ relation at the same level. However, due to the selection in X-ray luminosities that yield a complete sample, we have no reason to remove this cluster from the analysis.

5. Using mock data emulating the eDXL mass proxy measurements, and assuming more precise measurements of the global observables, we show that ignoring the intrinsic covariance of cluster observables can lead to significant biases (depending on the strength of the correlation) in the measured mass-observable scaling relation, even when that particular observable plays no role in the sample selection.
6. In our eDXL/APEX-SZ sample, neglecting the correlation (r) of the intrinsic scatter between the integrated Comptonization and X-ray luminosity biases the normalisation (high by $1\text{--}2\sigma$) and slope (low by 1σ) of the $Y_{\text{SZ}} - M_{500}$ scaling relation.

Cosmological studies using galaxy cluster scaling relations are currently limited by the systematic uncertainties in those assumed scaling laws. We show that, for future cluster-based cosmology experiments, the biases induced in the measured scaling relations by not accounting for the intrinsic covariance of cluster observables can dominate the error budget of cosmological analyses. It is prudent therefore to explicitly account for these covariance terms in the modeling of the scaling relations for future cluster-cosmology experiments.

Summary & discussion

7.1 Summary and significance

In Chapter 2, galaxy clusters were introduced as a cosmological tool capable of probing of the nature of dark energy and the sum of neutrino masses. It presented the mass calibration of cluster observables as one of the current limitations of using the abundance and growth of galaxy clusters for precision cosmology. One of the sources of biases in mass calibration is due to the biases in the sample selection of galaxy clusters themselves. Additionally, the correlations in the intrinsic covariance of cluster observables propagates the selection biases in one observable-mass plane to other cluster observable-mass planes.

In this work, a calibration of the integrated Comptonization defined within an aperture of R_{500} is measured by anchoring masses to weak-lensing mass measurements. An X-ray selected sample with well-defined sample selection criteria and almost complete follow-up in SZ and lensing is used for the mass calibration of the SZ effect. The measurement of the mass-calibration is done employing a Bayesian method to mitigate sample selection biases, including the effect of correlation in intrinsic scatter of SZ and X-ray luminosity. The consideration of such correlations in measurements of scaling properties is done by few authors (e.g., A. Mantz, Allen, Ebeling et al., 2010). Still, it has been largely ignored by others (e.g., Marrone et al., 2012) or some authors have relied on strong priors (e.g., Czakon et al., 2015). This is the first time that the $Y_{\text{SZ}} - M_{500}$ relation is calibrated by marginalising over the correlation in scatters of SZ and X-ray luminosity at fixed mass, including sample selection, the underlying mass function and using weak-lensing masses as the anchor for absolute mass scale. The posterior of the correlation parameter prefers a positive and non-zero value at 1σ confidence. This is of great interest as the $Y - M$ scaling normalisation and slope show degeneracies with the correlation parameter. The $Y_{\text{SZ}} - M_{500}$ scaling relation is found biased by almost by 2σ , when the correlation is completely ignored. Previous authors based on A. V. Kravtsov, Vikhlinin and Nagai (2006) have argued that this correlation should be small. However, recent numerical simulations present a different picture suggesting a strong correlation among cluster ICM observables. The measured value of this correlation, r , from this work is consistent with the predictions from simulations (Stanek, Rasia et al., 2010; Angulo et al., 2012; Truong et al., 2016), however, the obtained constraints are weak.

This work demonstrates that even with a sample size of ~ 30 clusters and the moderate measurement uncertainties, typical of most current measurements of cluster properties, the systematic effect in

measuring scaling relations due to the correlation in intrinsic scatters of cluster observables can be significant at $1-2\sigma$ level. It is also demonstrated with the help of mock samples with precise measurements, the systematic effect will be more acute.

In addition to the correlation between SZ and luminosity at fixed mass, this work also considered the intrinsic scatter in lensing masses (Becker and A. V. Kravtsov, 2011). This intrinsic scatter tends to bias the slope of the $Y-M$ relation and is also important to consider in the statistical framework as suggested by Sereno and Ettori (2015). A 20% Gaussian scatter is assumed and the calibration is measured along with the correlation parameter between SZ and luminosity at fixed mass. In general, a steeper slope is preferred by such a modelling, however, the constraints on the correlation parameter is destroyed due to limited precision on mass estimates. Some authors have suggested a possible correlation in lensing masses and SZ at fixed mass (Shirasaki, Nagai and Lau, 2016). Due to limited precision of the measurements of cluster observables, this correlation is not attempted to be measured in this work.

One of the significant conclusion of this work is the implications of the systematic bias in $Y_{\text{SZ}}-M$ scaling for future cluster cosmological analyses. It is demonstrated that for near-future cluster surveys (e.g., SPT-3G Benson, Ade et al., 2014), employing the unbiased and biased calibration of cluster observable predict vastly different cluster counts, 5000 and ~ 21000 respectively. A bias in the cluster counts will find different geometries and energy budget for the Universe.

7.2 Scope of this work

The Eddington and Malmquist biases are widely common in extra-galactic surveys of astrophysical objects. The Bayesian approach taken in this work is widely applicable for measuring a relationship between observables and a covariate variable, for which there is some knowledge of the underlying covariate distribution. The method accounts for non-ignorable selection effects for a well-defined sample in measuring such relationships of some observables with a covariate variable. For any complete sample with a complete follow-up in the measurements of properties of astrophysical objects, these relationships between observable and covariate can be measured reliably with the approach outlined in Chapter 6.

In particular, the Bayesian method employed in this work along with the analytical modulation of the cluster mass function used to approximate the sample selection by an observable threshold can be used and generalised for measuring scaling properties of cluster observables (those that play a role in the sample selection, and as well for those that do not) by using other samples, such as, LoCuSS (G. P. Smith et al., 2016), REXCESS (Pratt et al., 2009), 400d (Burenin et al., 2007; Vikhlinin, Burenin et al., 2009). The method may also be easily extended to calibrate different types of masses and measure the so called *mass bias* between X-ray masses and lensing masses. This mass bias is crucial for current cosmological analyses that have historically relied on X-ray mass scaling of observables (e.g., Vikhlinin, A. V. Kravtsov et al., 2009; Benson, de Haan et al., 2013; de Haan et al., 2016). An extension to the work done in this thesis is to measure this mass bias between the Planck masses measured using their mass calibration (Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al., 2014) and lensing mass for the eDXL sample. Current measurements of this bias between Planck masses and lensing masses have yielded inconsistent values between several authors (Penna-Lima et al., 2016). Most of these works have used samples with ill-defined selection or have used a χ^2 method to measure the bias. With the sample introduced in this work and along

with the Bayesian method employed here, there is scope for an accurate measurement of this bias.

7.3 Limitations of this work

Although, in Chapter 4 and 5, care is taken in measuring the Y_{SZ} from APEX-SZ, the Compton- Y is not measured with relativistic corrections. On average for the eDXL sample this would require 5% level correction factor. The availability of homogeneous X-ray analysis can provide temperatures that could help determine the corrections on cluster-to-cluster basis. The measured relations are limited by the 8% systematic uncertainties in the lensing masses, which is currently ignored in the statistical analysis. The constraints on correlation coefficient of the intrinsic scatter between L_{x} and Y_{SZ} is weak and this work did not pursue constraining or accounting for the correlations between other observables. Better X-ray data will be useful in further understanding the origin of the correlation between SZ and X-ray luminosity at fixed mass and confirm with more confidence the value of this correlation.

Bibliography

- Abazajian, K. N. et al. (2016), ArXiv e-prints, arXiv: 1610.02743.
- Abell, G. O. (1958), *ApJS* **3** 211.
- Abell, G. O., H. G. Corwin and R. P. Olowin (1989), *ApJS* **70** 1.
- Akritas, M. G. and M. A. Bershady (1996), *ApJ* **470** 706, eprint: astro-ph/9605002.
- Allen, S. W., A. E. Evrard and A. B. Mantz (2011), *ARA&A* **49** 409, arXiv: 1103.4829 [astro-ph.CO].
- Andersson, K. et al. (2011), *ApJ* **738**, 48 48, arXiv: 1006.3068 [astro-ph.CO].
- Andreon, S. (2015), *A&A* **582**, A100 A100, arXiv: 1509.00589.
- Andreon, S. and M. Hurn (2013), Statistical Analysis and Data Mining: The ASA Data Science Journal, Vol. 9, Issue 1, p. 15-33 6 15, arXiv: 1210.6232 [astro-ph.IM].
- Angulo, R. E. et al. (2012), *MNRAS* **426** 2046, arXiv: 1203.3216.
- Applegate, D. E. et al. (2014), *MNRAS* **439** 48, arXiv: 1208.0605.
- Arnaud, M. et al. (2010), *A&A* **517**, A92 A92, arXiv: 0910.1234 [astro-ph.CO].
- Bartelmann, M. (1996), *A&A* **313** 697, eprint: arXiv:astro-ph/9602053.
- Bartelmann, M. and M. Maturi (2017), Scholarpedia **12** 32440, arXiv: 1612.06535.
- Basu, K. et al. (2010), *A&A* **519**, A29 A29, arXiv: 0911.3905.
- Becker, M. R. and A. V. Kravtsov (2011), *ApJ* **740**, 25 25, arXiv: 1011.1681 [astro-ph.CO].
- Bender, A. N. et al. (2016), *MNRAS* **460** 3432.
- Benson, B. A., P. A. R. Ade et al. (2014), “SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope”, *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, vol. 9153, Proceedings of SPIE 91531P, arXiv: 1407.2973 [astro-ph.IM].
- Benson, B. A., T. de Haan et al. (2013), *ApJ* **763**, 147 147, arXiv: 1112.5435 [astro-ph.CO].
- Bhattacharya, S. et al. (2013), *ApJ* **766**, 32 32, arXiv: 1112.5479 [astro-ph.CO].
- Birkinshaw, M. (1999), *Physics Reports* **310** 97, eprint: astro-ph/9808050.
- Bleem, L. E. et al. (2015), *ApJS* **216**, 27 27, arXiv: 1409.0850.
- Bocquet, S. et al. (2016), *MNRAS* **456** 2361, arXiv: 1502.07357.
- Böhringer, H., G. Chon, C. A. Collins et al. (2013), *A&A* **555**, A30 A30.
- Böhringer, H., G. Chon and P. P. Kronberg (2016), *A&A* **596**, A22 A22.
- Böhringer, H., G. Chon, J. Retzlaff et al. (2017), *AJ* **153**, 220 220, arXiv: 1704.06489.
- Böhringer, H., P. Schuecker et al. (2004), *A&A* **425** 367, eprint: astro-ph/0405546.
- Böhringer, H., W. Voges et al. (2000), *ApJS* **129** 435, eprint: astro-ph/0003219.
- Bonamente, M. et al. (2008), *ApJ* **675**, 106-114 106, arXiv: 0708.0815.
- Bond, J. R., L. Kofman and D. Pogosyan (1996), *Nature* **380** 603, eprint: astro-ph/9512141.
- Borgani, S. and A. Kravtsov (2011), *Advanced Science Letters* **4** 204, arXiv: 0906.4370 [astro-ph.CO].
- Bower, R. G., J. R. Lucey and R. S. Ellis (1992), *MNRAS* **254** 601.
- Boylan-Kolchin, M. et al. (2009), *MNRAS* **398** 1150, arXiv: 0903.3041 [astro-ph.CO].
- Bryan, G. L. and M. L. Norman (1998), *ApJ* **495**.1 80, ISSN: 0004-637X, arXiv: 9710107 [astro-ph].

Bibliography

- Burenin, R. A. et al. (2007), ApJS **172** 561, eprint: astro-ph/0610739.
- Carlstrom, J. E., G. P. Holder and E. D. Reese (2002), ARA&A **40** 643, eprint: astro-ph/0208192.
- Cavaliere, A. and R. Fusco-Femiano (1976), A&A **49** 137.
- Czakon, N. G. et al. (2015), ApJ **806**, 18 18, arXiv: 1406.2800.
- Davis, M. and P. J. E. Peebles (1983), ApJ **267** 465.
- de Haan, T. et al. (2016), ArXiv e-prints, arXiv: 1603.06522.
- Dietrich, J. P. et al. (2017), ArXiv e-prints, arXiv: 1711.05344.
- Dobbs, M. et al. (2006), New Astronomy Reviews **50** 960.
- Ebeling, H., A. C. Edge, S. W. Allen et al. (2000), MNRAS **318** 333, eprint: astro-ph/0003191.
- Ebeling, H., A. C. Edge, H. Bohringer et al. (1998), MNRAS **301** 881, eprint: astro-ph/9812394.
- Eddington, A. S. (1913), MNRAS **73** 359.
- Erben, T. et al. (2001), A&A **366** 717, eprint: astro-ph/0007021.
- Ettori, S. (2015), MNRAS **446** 2629, arXiv: 1410.8522.
- Felten, J. E. et al. (1966), ApJ **146** 955.
- Ferraro, S. et al. (2016), Phys. Rev. D **94**.12, 123526 123526, arXiv: 1605.02722.
- Gelman, A. et al. (2014), *Bayesian Data Analysis*.
- Giles, P. A., B. J. Maughan, H. Dahle et al. (2017), MNRAS **465** 858, arXiv: 1510.04270.
- Giles, P. A., B. J. Maughan, F. Pacaud et al. (2016), A&A **592**, A3 A3, arXiv: 1512.03833.
- Giodini, S. et al. (2013), Space Sci. Rev. **177** 247, arXiv: 1305.3286.
- Gladders, M. D. and H. K. C. Yee (2000), AJ **120** 2148, eprint: astro-ph/0004092.
- Gonzalez, A. H. et al. (2015), ApJLett **812**, L40 L40, arXiv: 1509.01989.
- Gruen, D. et al. (2015), MNRAS **449** 4264, arXiv: 1501.01632.
- Gull, S. F. and K. J. E. Northover (1976), Nature **263** 572.
- Gunn, J. E. and J. R. Gott III (1972), ApJ **176** 1.
- Güsten, R. et al. (2006), A&A **454** L13.
- Hallman, E. J. et al. (2007), ApJ **665** 911, arXiv: 0705.0531.
- Halverson, N. W. et al. (2009), ApJ **701** 42, arXiv: 0807.4208.
- Hand, N. et al. (2012), Physical Review Letters **109**.4, 041101 041101, arXiv: 1203.4219.
- Hasselfield, M. et al. (2013), Journal of cosmology and Astrophysics **7**, 008 008, arXiv: 1301.0816 [astro-ph.CO].
- Henderson, S. W. et al. (2016), Journal of Low Temperature Physics **184**.3 772, ISSN: 1573-7357.
- Hennig, C. et al. (2017), MNRAS **467** 4015, arXiv: 1604.00988.
- Hoekstra, H. (2007), MNRAS **379** 317, arXiv: 0705.0358.
- Hoekstra, H., R. Herbonnet et al. (2015), MNRAS **449** 685, arXiv: 1502.01883.
- Hoekstra, H., A. Mahdavi et al. (2012), MNRAS **427** 1298, arXiv: 1208.0606.
- Hogg, D. W. (1999), ArXiv Astrophysics e-prints, eprint: astro-ph/9905116.
- Ilbert, O. et al. (2009), ApJ **690** 1236, arXiv: 0809.2101.
- Israel, H. et al. (2012), A&A **546**, A79 A79, arXiv: 1112.4444.
- Itoh, N., Y. Kohyama and S. Nozawa (1998), ApJ **502** 7, eprint: astro-ph/9712289.
- Jee, M. J. et al. (2012), ApJ **747**, 96 96, arXiv: 1202.6368.
- Jenkins, A. et al. (2001), MNRAS **321** 372, eprint: astro-ph/0005260.
- Kaiser, N. (1986), MNRAS **222** 323.
- Kaiser, N., G. Squires and T. Broadhurst (1995), ApJ **449** 460, eprint: astro-ph/9411005.
- Kelly, B. C. (2007), ApJ **665** 1489, arXiv: 0705.2774.
- Kelly, P. L. et al. (2014), MNRAS **439** 28, arXiv: 1208.0602.
- Klein, M. (2014), “Weak lensing measurements for the APEX-SZ cluster survey”, 3848, PhD Thesis: University of Bonn.

- Koester, B. P. et al. (2007), ApJ **660** 239, eprint: astro-ph/0701265.
- Komatsu, E. et al. (2011), ApJS **192**, 18 18, arXiv: 1001.4538 [astro-ph.CO].
- Kravtsov, A. V., A. Vikhlinin and D. Nagai (2006), ApJ **650** 128, eprint: astro-ph/0603205.
- Laureijs, R. et al. (2011), ArXiv e-prints, eprint: 1110.3193.
- Limousin, M. et al. (2013), Space Sci. Rev. **177** 155, arXiv: 1210.3067.
- Lin, Y.-T. and J. J. Mohr (2004), ApJ **617** 879, eprint: astro-ph/0408557.
- LSST Science Collaboration et al. (2009), ArXiv e-prints, arXiv: 0912.0201 [astro-ph.IM].
- Lynden-Bell, D. (1967), MNRAS **136** 101.
- Malmquist, G. K. (1920), Meddelanden fran Lunds Astronomiska Observatorium Serie II **22** 3.
- Mantz, A. B. et al. (2015), MNRAS **446** 2205, arXiv: 1407.4516.
- Mantz, A., S. W. Allen, H. Ebeling et al. (2010), MNRAS **406** 1773, arXiv: 0909.3099.
- Mantz, A., S. W. Allen, R. G. Morris et al. (2016), ArXiv e-prints, arXiv: 1606.03407.
- Mantz, A., A. von der Linden et al. (2015), MNRAS **446** 2205, arXiv: 1407.4516.
- Marriage, T. A. et al. (2011), ApJ **737**, 61 61, arXiv: 1010.1065.
- Marrone, D. P. et al. (2012), ApJ **754**, 119 119, arXiv: 1107.5115.
- Martino, R. et al. (2014), MNRAS **443** 2342, arXiv: 1406.6831.
- Maughan, B. J. (2014), MNRAS **437** 1171, arXiv: 1212.0858.
- Maughan, B. J. et al. (2012), MNRAS **421** 1583, arXiv: 1108.1200.
- Meneghetti, M. et al. (2010), A&A **514**, A93 A93, arXiv: 0912.1343.
- Merloni, A. et al. (2012), ArXiv e-prints, arXiv: 1209.3114 [astro-ph.HE].
- Mulroy, S. L. et al. (2014), MNRAS **443** 3309, arXiv: 1407.1767.
- Nagai, D., A. V. Kravtsov and A. Vikhlinin (2007), ApJ **668** 1, eprint: astro-ph/0703661.
- Narayan, R. and M. Bartelmann (1999), “Gravitational lensing”, *Formation of Structure in the Universe*, ed. by A. Dekel and J. P. Ostriker 360.
- Navarro, J. F., C. S. Frenk and S. D. M. White (1995), MNRAS **275** 720, eprint: astro-ph/9408069.
- (1996), ApJ **462** 563, eprint: astro-ph/9508025.
- (1997), ApJ **490** 493, eprint: astro-ph/9611107.
- Nord, M. (2009), “Observations of the Sunyaev-Zel’dovich Effect towards Clusters of Galaxies with the APEX Telescope”, 1873, PhD Thesis: University of Bonn.
- Nord, M. et al. (2009), A&A **506** 623, arXiv: 0902.2131 [astro-ph.CO].
- Oguri, M. (2014), MNRAS **444** 147, arXiv: 1407.4693.
- Okabe, N. and G. P. Smith (2016), MNRAS **461** 3794, arXiv: 1507.04493.
- Okabe, N., G. P. Smith et al. (2013), ApJLett **769**, L35 L35, arXiv: 1302.2728.
- Okabe, N., M. Takada et al. (2010), PASJ **62** 811, arXiv: 0903.1103.
- Pacaud, F., N. Clerc et al. (2016), A&A **592**, A2 A2, arXiv: 1512.04264.
- Pacaud, F., M. Pierre et al. (2007), MNRAS **382** 1289, arXiv: 0709.1950.
- Peebles, P. J. E. (1993), *Principles of Physical Cosmology*.
- Penna-Lima, M. et al. (2016), ArXiv e-prints, arXiv: 1608.05356.
- Percival, W. J. (2005), A&A **443** 819, eprint: astro-ph/0508156.
- Perlmutter, S. et al. (1999), ApJ **517** 565, eprint: astro-ph/9812133.
- Phillips, P. R. (1995), ApJ **455** 419.
- Piffaretti, R. et al. (2011), A&A **534**, A109 A109, arXiv: 1007.1916.
- Plagge, T. et al. (2010), ApJ **716** 1118, arXiv: 0911.2444.
- Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al. (2014), A&A **571**, A20 A20.
- Planck Collaboration, Ade, P. A. R., Aghanim, N. and et al. (2016), A&A **594** A13.

Bibliography

- Planck Collaboration, P. A. R. Ade, N. Aghanim and et al. (2016a), *A&A* **594**, A24 A24, arXiv: [1502.01597](#).
– (2016b), *A&A* **594**, A27 A27, arXiv: [1502.01598](#).
- Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud et al. (2011), *A&A* **536**, A8 A8, arXiv: [1101.2024](#).
- Planck Collaboration and et al. (2016a), *A&A* **594** A11.
– (2016b), *A&A* **586**, A140 A140, arXiv: [1504.03339](#).
- Planck Collaboration, Ade, P. A. R., Aghanim, N. et al. (2016), *A&A* **594** A17.
- Postman, M. et al. (2012), *ApJS* **199**, 25 25, arXiv: [1106.3328](#).
- Pratt, G. W. et al. (2009), *A&A* **498** 361, arXiv: [0809.3784](#).
- Press, W. H. and P. Schechter (1974), *ApJ* **187** 425.
- Rasia, E. et al. (2012), *New Journal of Physics* **14**.5, 055018 055018, arXiv: [1201.1569](#).
- Reichardt, C. L., B. Stalder et al. (2013), *ApJ* **763**, 127 127, arXiv: [1203.5775](#).
- Reichardt, C. L., O. Zahn et al. (2009), *ApJ* **701** 1958, arXiv: [0904.3939](#) [astro-ph.CO].
- Reichert, A. et al. (2011), *A&A* **535**, A4 A4, arXiv: [1109.3708](#).
- Reiprich, T. H., K. Basu et al. (2013), *Space Sci. Rev.* **177** 195, arXiv: [1303.3286](#).
- Reiprich, T. H. and H. Böhringer (2002), *ApJ* **567** 716, eprint: [astro-ph/0111285](#).
- Riess, A. G. et al. (1998), *AJ* **116** 1009, eprint: [astro-ph/9805201](#).
- Romero, C. et al. (2016), ArXiv e-prints, arXiv: [1608.03980](#).
- Rozo, E. et al. (2010), *ApJ* **708** 645, arXiv: [0902.3702](#) [astro-ph.CO].
- Rykoff, E. S. et al. (2016), *ApJS* **224**, 11, arXiv: [1601.00621](#).
- Saliwanchik, B. R. et al. (2015), *ApJ* **799**, 137 137, arXiv: [1312.3015](#).
- Salvati, L., M. Douspis and N. Aghanim (2017), ArXiv e-prints, arXiv: [1708.00697](#).
- Sarazin, C. L. (1988), *X-ray Emission from Clusters of Galaxies*.
- Sayers, J., S. R. Golwala et al. (2011), *ApJ* **728**, 39 39, arXiv: [1010.1798](#) [astro-ph.CO].
- Sayers, J., T. Mroczkowski et al. (2013), *ApJ* **778**, 52 52, arXiv: [1312.3680](#) [astro-ph.CO].
- Sayers, J. et al. (2013), *ApJ* **768**.2 177, ISSN: 0004-637X, arXiv: [1211.1632](#).
- Schellenberger, G. and T. H. Reiprich (2017a), *MNRAS* **469** 3738, arXiv: [1705.05842](#).
– (2017b), *MNRAS* **471** 1370, arXiv: [1705.05843](#).
- Schrabback, T., D. Applegate et al. (2016), ArXiv e-prints, arXiv: [1611.03866](#).
- Schrabback, T., T. Erben et al. (2007), *A&A* **468** 823, eprint: [arXiv:astro-ph/0606611](#).
- Schwan, D. et al. (2011), *Review of Scientific Instruments* **82**.9 091301, arXiv: [1008.0342](#) [astro-ph.IM].
- Sembolini, F. et al. (2013), *MNRAS* **429** 323, arXiv: [1207.4438](#).
- Sereno, M., G. Covone et al. (2017), *MNRAS* **472** 1946, arXiv: [1703.06886](#).
- Sereno, M. and S. Ettori (2015), *MNRAS* **450** 3675.
- Sereno, M., S. Ettori and A. Baldi (2012), *MNRAS* **419** 2646, arXiv: [1109.2732](#).
- Sereno, M., S. Ettori and L. Moscardini (2015), *MNRAS* **450** 3649, arXiv: [1407.7869](#).
- Sereno, M., K. Umetsu et al. (2013), *Astronomische Nachrichten* **334** 445.
- Sheth, R. K., H. J. Mo and G. Tormen (2001), *MNRAS* **323** 1, eprint: [astro-ph/9907024](#).
- Sheth, R. K. and G. Tormen (1999), *MNRAS* **308** 119, eprint: [astro-ph/9901122](#).
- Shirasaki, M., D. Nagai and E. T. Lau (2016), *MNRAS* **460** 3913, arXiv: [1603.08609](#).
- Shull, J. M., B. D. Smith and C. W. Danforth (2012), *ApJ* **759**, 23 23, arXiv: [1112.2706](#).
- Smith, G. P. et al. (2016), *MNRAS* **456** L74, arXiv: [1511.01919](#).
- Smoot, G. F. et al. (1992), *ApJLett* **396** L1.
- Soergel, B. et al. (2016), *MNRAS* **461** 3172, arXiv: [1603.03904](#).
- Stanek, R., A. E. Evrard et al. (2006), *ApJ* **648** 956, eprint: [astro-ph/0602324](#).

- Stanek, R., E. Rasia et al. (2010), ApJ **715** 1508, arXiv: [0910.1599](#).
- Sunyaev, R. A. and Y. B. Zeldovich (1970), Ap&SS **7** 20.
- Sunyaev, R. A. and Y. B. Zel'dovich (1970), Comments on Astrophysics and Space Physics **2** 66.
- (1972), Comments on Astrophysics and Space Physics **4** 173.
- Tinker, J. et al. (2008), ApJ **688**, 709-728 709, arXiv: [0803.2706](#).
- Truong, N. et al. (2016), ArXiv e-prints, arXiv: [1607.00019](#).
- Tyson, J. A., F. Valdes and R. A. Wenk (1990), ApJLett **349** L1.
- Umetsu, K., E. Medezinski et al. (2014), ApJ **795**, 163 163, arXiv: [1404.1375](#).
- Umetsu, K., M. Sereno et al. (2015), ApJ **806**, 207 207, arXiv: [1503.01482](#).
- Vanderlinde, K. et al. (2010), ApJ **722** 1180, arXiv: [1003.0003](#).
- Vikhlinin, A., R. A. Burenin et al. (2009), ApJ **692** 1033, arXiv: [0805.2207](#).
- Vikhlinin, A., A. V. Kravtsov et al. (2009), ApJ **692** 1060, arXiv: [0812.2720](#).
- von der Linden, A. et al. (2014), MNRAS **439** 2, arXiv: [1208.0597](#).
- Wang, S. et al. (2005), Physical Review Letters **95**.1, 011302 011302, eprint: [astro-ph/0505390](#).
- Weißmann, A. et al. (2013), A&A **549**, A19 A19, arXiv: [1210.6445](#) [astro-ph.CO].
- Wen, Z. L., J. L. Han and F. S. Liu (2012), ApJS **199**, 34 34, arXiv: [1202.6424](#) [astro-ph.CO].
- Yang, H.-Y. K., S. Bhattacharya and P. M. Ricker (2010), ApJ **725** 1124, arXiv: [1010.0249](#).
- Zhang, Y.-Y. et al. (2006), A&A **456** 55, eprint: [astro-ph/0603275](#).
- Zwicky, F. (1933), Helvetica Physica Acta **6** 110.
- Zwicky, F. et al. (1961), *Catalogue of galaxies and of clusters of galaxies, Vol. I.*

APPENDIX A

Appendix: Mass calibration and statistical models

A.1 Derivation of the normalised likelihood for the eDXL sample

Denoting measured variables with tilde, symbols \tilde{M}_{WL} , \tilde{Y}_{SZ} , and \tilde{L}_x represent the measured observables, namely, weak-lensing mass, SZ Compton parameter, and soft-band luminosities of a given cluster respectively. $\sigma_{\tilde{M}_{\text{WL}}}$, $\sigma_{\tilde{Y}_{\text{SZ}}}$ and $\sigma_{\tilde{L}}$ represent their uncertainties respectively.

We assume that the ROSAT measurement scatter is log-normal, making it easier to compute analytically the normalisation of the likelihood. Therefore, in addition to the log-normal intrinsic scatter in the luminosities, there is a log-normal distributed measurement scatter. We model the log-normal scatters in the L and Y as correlated intrinsic scatters with a correlation coefficient parameter, r . The selection function for eDXL cluster at a given redshift is

$$P(\mathcal{I} = 1 | \tilde{L}, L^{\min}) = 1 \text{ when } \tilde{L} \geq L^{\min} \quad (\text{A.1})$$

otherwise, $P(\mathcal{I} = 1 | \tilde{L}, L^{\min}) = 0$. The full normalized likelihood is given as

$$\mathcal{L} = \frac{P(\tilde{M}, \tilde{Y}, \tilde{L} | \theta)}{\int d\tilde{L} P(\mathcal{I} = 1 | \tilde{L}, \tilde{L}^{\min}) P(\tilde{M}, \tilde{Y}, \tilde{L} | \theta)}. \quad (\text{A.2})$$

Since measurements on luminosities are independent of other measured quantities, the normalisation can be reduced to $\int d\tilde{L} P(\mathcal{I} = 1 | \tilde{L}, \tilde{L}^{\min}) P(\tilde{L} | \theta)$ or $\int d\ln \tilde{L} P(\mathcal{I} = 1 | \ln \tilde{L}, \ln \tilde{L}^{\min}) P(\ln \tilde{L} | \theta)$. We consider,

$$\begin{aligned} P(\ln \tilde{L} | \theta) &= \int d\ln L P(\ln \tilde{L} | \ln L) P(\ln L | \theta) = \iint d\ln L dM P(\ln \tilde{L} | \ln L) P(\ln L | M, \theta) P(M) \\ &= \int dM P(\ln \tilde{L} | \theta, M) P(M). \end{aligned} \quad (\text{A.3})$$

Assuming a log-normal distribution for measured luminosity distribution and the log-normal intrinsic

scatter in the scaling law, the probability density, $P(\ln \tilde{L}|M, \theta)$, is given by

$$P(\ln \tilde{L}|M, \theta) = \frac{1}{\sqrt{2\pi(\sigma_{\ln \tilde{L}}^2 + \sigma_{\ln L}^2)}} \exp\left(-\frac{1}{2} \frac{(\ln \tilde{L} - \ln \hat{L}(M))^2}{\sigma_{\ln \tilde{L}}^2 + \sigma_{\ln L}^2}\right). \quad (\text{A.4})$$

To fully obtain the normalisation of the likelihood, we compute:

$$\int_0^\infty dM \int_{-\infty}^{+\infty} d\ln \tilde{L} P(\mathcal{I} = 1 | \ln \tilde{L}, \ln \tilde{L}^{\min}) P(\ln \tilde{L}|M, \theta) P(M), \quad (\text{A.5})$$

which reduces to $\int_0^{+\infty} dM \int_{\ln L^{\min}}^{+\infty} d\ln \tilde{L} P(\ln \tilde{L}|M, \theta) P(M)$.

Substituting for $P(\ln \tilde{L}|M, \theta)$ from Equation (A.4) and integrating the above expression gives the final normalised likelihood as:

$$\mathcal{L} \propto \prod_{i=1}^{N_{\det}} \frac{P(\tilde{M}_i, \tilde{Y}_i, \tilde{L}_i | \theta)}{\int_0^{+\infty} \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\ln \tilde{L}_i^{\min} - \ln \hat{L}(M'_i)}{\sqrt{2(\sigma_{\ln \tilde{L}_i}^2 + \sigma_{\ln L_i}^2)}} \right) \right] P(M'_i) dM'_i} \quad (\text{A.6})$$

The erf in the denominator arises due to the Heaviside step function used in the selection of luminosities. The denominator of the likelihood gives the probability of including the cluster in the sample. The function $L(M_i)$ is given by the relation in Equation (6.9). The detailed description of probabilities related to Y_{SZ} and M_{WL} are given in appendix A.2. The integrations over the nuisance parameters, namely, the true underlying values of the observables are computed via an MCMC by marginalising over the true observable variables. At each step of the MCMC, the normalization varies with the $L - M$ scaling parameters, which are left free in our fitting. Hence, the likelihood is re-normalised at every step in the MCMC.

A.2 Probabilities

A.2.1 Measurement probabilities

Measurement probability on weak-lensing and SZ observables: The conditional probability that $\tilde{M}_{\text{WL}}, \tilde{Y}_{\text{SZ}}$ are measured given true values, M_{WL} and Y_{SZ} , is denoted by $p(\tilde{M}_{\text{WL}}, \tilde{Y}_{\text{SZ}}|M_{\text{WL}}, Y_{\text{SZ}})$. We model the measurement probability distributions on weak-lensing masses and the Compton-y as multi-variate Gaussian probability distribution functions. The equation is as follows:

$$P(\tilde{M}_{\text{WL}}, \tilde{Y}_{\text{SZ}}|M_{\text{WL}}, Y_{\text{SZ}}) = \left[\frac{1}{2\pi\sigma_{\tilde{M}_{\text{WL}}}\sigma_{\tilde{Y}_{\text{SZ}}}(\sqrt{1-\rho^2})} \right] F(\tilde{M}_{\text{WL}}, \tilde{Y}_{\text{SZ}}|M_{\text{WL}}, Y_{\text{SZ}}), \quad (\text{A.7})$$

where

$$F(\tilde{M}, \tilde{Y}|M, Y) = \exp\left(-\frac{1}{2}\frac{1}{(1-\rho^2)}\left\{\left(\frac{\tilde{M}-M}{\sigma_{\tilde{M}}}\right)^2 + \left(\frac{\tilde{Y}-Y}{\sigma_{\tilde{Y}}}\right)^2 - 2\rho\frac{(\tilde{Y}-Y)(\tilde{M}-M)}{\sigma_{\tilde{M}}\sigma_{\tilde{Y}}}\right\}\right)$$

and ρ is the correlation in the measurement uncertainties in the two observables.

Measurement probability of luminosities $P(\tilde{L}_x|L)$ is assumed to be log-normal with lognormal measurement scatter $\sigma_{\ln \tilde{L}} = \sigma_{\tilde{L}}/\tilde{L}_x$.

A.2.2 Scaling model probabilities

Log-normal probability distribution of the Y_{SZ} and L_x at fixed mass is given by,

$$\begin{aligned} P(Y_{\text{SZ}}, L_x|M, \theta) &= \frac{1}{2\pi\sigma_{\ln L_x}\sigma_{\ln Y_{\text{SZ}}}Y_{\text{SZ}}L_x\sqrt{1-r^2}} \\ &\times \exp\left\{-\frac{1}{2}\frac{1}{(1-r^2)}\left[\left(\frac{\ln Y_{\text{SZ}} - \ln \hat{Y}(M)}{\sigma_{\ln Y_{\text{SZ}}}}\right)^2 + \left(\frac{\ln L_x - \ln \hat{L}(M)}{\sigma_{\ln L_x}}\right)^2\right.\right. \\ &\quad \left.\left.- 2r\frac{(\ln Y_{\text{SZ}} - \ln \hat{Y}(M))(\ln L_x - \ln \hat{L}(M))}{\sigma_{\ln L_x}\sigma_{\ln Y_{\text{SZ}}}}\right]\right\}. \end{aligned} \quad (\text{A.8})$$

A.2.3 Intrinsic scatter in weak-lensing mass

Here we calculate the marginalisation over the true weak-lensing mass $M_{\text{WL},500}$ for the scaling model described in Section 6.3.2.2. We assume the true weak-lensing mass is unbiased and scatters from the halo mass M_{HM} with Gaussian distribution such that the dispersion is proportional to halo mass. The probability distribution of measured lensing masses is given by a Gaussian probability:

$$P(M_{\text{WL}}|M_{\text{HM}}) = \frac{1}{\sqrt{2\pi}\sigma_{M_{\text{WL}}|M_{\text{HM}}}} \exp\left[-0.5\left(\frac{M_{\text{WL}} - M_{\text{HM}}}{\sigma_{M_{\text{WL}}|M_{\text{HM}}}}\right)^2\right], \quad (\text{A.9})$$

where $\sigma_{M_{\text{WL}}|M_{\text{HM}}} = 0.2M_{\text{HM}}$.

To calculate $P(\tilde{M}_{\text{WL}}|M_{\text{HM}})$, we compute the integral $\int M_{\text{WL}} P(\tilde{M}_{\text{WL}}|M_{\text{WL}})P(M_{\text{WL}}|M_{\text{HM}})$. Since in our case, the measurement probability of weak-lensing mass and the SZ observable measurements are correlated, we should take into account the cross-terms to compute the final integral. Consider the integral,

$$\int dM_{\text{WL}} \exp\left[-0.5\left(\frac{M_{\text{WL}} - M_{\text{HM}}}{\sigma_{M_{\text{WL}}|M_{\text{HM}}}}\right)^2\right] F(\tilde{M}_{\text{WL}}, \tilde{Y}_{\text{SZ}}|M_{\text{WL}}, Y_{\text{SZ}}), \quad (\text{A.10})$$

where $F(\tilde{M}_{\text{WL}}, \tilde{Y}_{\text{SZ}} | M_{\text{WL}}, Y_{\text{SZ}})$ has the same meaning as given in equation (A.7). By defining,

$$A = \frac{1}{2} \left\{ \left[\frac{1}{(1 - \rho^2) \sigma_{\tilde{M}_{\text{WL}}}^2} \right] + \left[\frac{1}{\sigma_{M_{\text{WL|HM}}}^2} \right] \right\}, \quad (\text{A.11})$$

$$B = \frac{1}{2} \left\{ \frac{1}{1 - \rho^2} \left[\frac{-2\tilde{M}_{\text{WL}}}{\sigma_{\tilde{M}_{\text{WL}}}^2} + \frac{2\rho(\tilde{Y}_{\text{SZ}} - Y_{\text{SZ}})}{\sigma_{\tilde{M}_{\text{WL}}} \sigma_{\tilde{Y}_{\text{SZ}}}} \right] + \left[\frac{-2M_{\text{HM}}}{\sigma_{M_{\text{WL|HM}}}^2} \right] \right\}, \quad (\text{A.12})$$

and

$$C = \frac{1}{2} \left\{ \left[\left(\frac{\tilde{M}_{\text{WL}}}{\sigma_{\tilde{M}_{\text{WL}}}} \right)^2 + \left(\frac{\tilde{Y}_{\text{SZ}} - Y_{\text{SZ}}}{\sigma_{\tilde{Y}_{\text{SZ}}}} \right)^2 - \frac{2\rho\tilde{M}_{\text{WL}}(\tilde{Y}_{\text{SZ}} - Y_{\text{SZ}})}{\sigma_{\tilde{M}_{\text{WL}}} \sigma_{\tilde{Y}_{\text{SZ}}}} \right] \right. \quad (\text{A.13})$$

$$\left. + \left[\frac{M_{\text{HM}}^2}{\sigma_{M_{\text{WL|HM}}}^2} \right] \right\}, \quad (\text{A.14})$$

the integral (A.10) reduces to

$$\int_0^{+\infty} dx e^{-Ax^2 - Bx - C} = \frac{1}{2} \sqrt{\frac{\pi}{A}} e^{\frac{B^2}{4A} - C} \left[1 - \text{erf} \left(\frac{B}{2\sqrt{A}} \right) \right], \quad (\text{A.15})$$

where M_{WL} is the running x variable in the above integral.

A.3 Tests with simulated data

We analyse with the help of simulated mock data sets, the impact of ignoring correlated scatters in response variables at a fixed independent variable. The two response variables are used synonymously with L_x , Y_{SZ} for the sake of consistency. However, the response variables are inter-changeable when the selection is on some other observable such as temperature, Y_{SZ} . The presented analysis can be extended to scaling relations with other thermodynamic properties as long as the selection is known and the independent variable is calibrated with mass of a cluster.

A.3.1 Mock data with correlated intrinsic scatters

A.3.1.1 Generating mock samples

We use the Tinker mass function (Tinker et al., 2008) to generate a sample of galaxy clusters whose underlying distribution in terms total masses in R_{500} is $\propto p(M_{500}, z)$. The mass function used here was for the cosmology where $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $H_0 = 70$ km/s/Mpc, $\sigma_8 = 0.82$, $\Omega_b = 0.045$, and slope of the primordial power-spectrum, $n_{pk} = -1.0$. Using a set of input scaling relation parameters for the $L_x - M_{500}$ and the $Y_{\text{SZ}} - M_{500}$ relations (Equations (6.9) and (6.10)), we generate true values of Y_{SZ} and L_x observables including a log-normal intrinsic covariance at fixed mass as per Equation (6.11). Subsequently, Y_{SZ} , M_{500} are further scattered with Gaussian measurement scatters assuming some relative percentage uncertainties. L_x is scattered with a log-normal measurement uncertainty. The

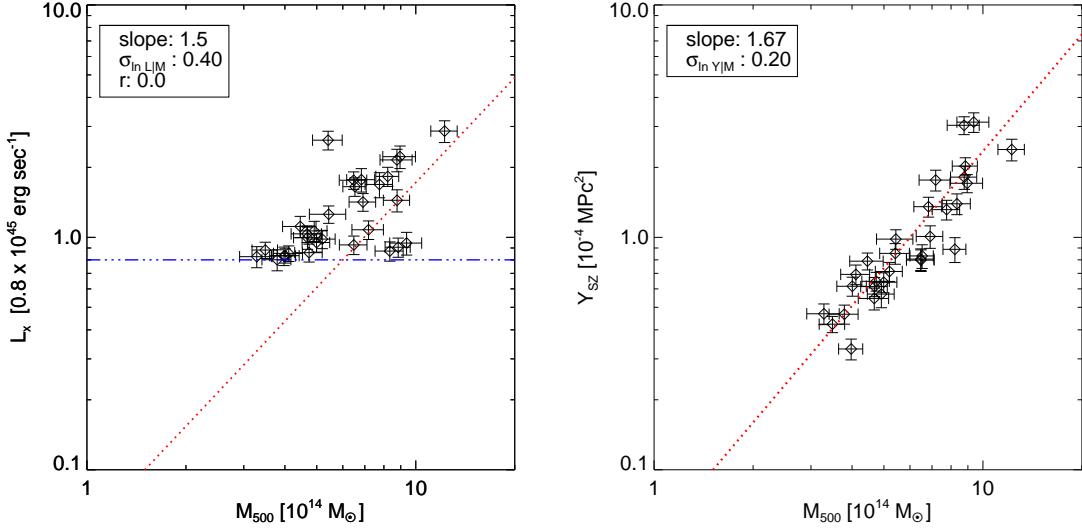


Figure A.1: An example set of simulated sample with their mock measurements at redshift $z = 0.30$ is shown here. The sample was selected by applying threshold selection (blue horizontal line) in measured luminosities. The simulations of mock data included a correlated intrinsic scatters in the two response variables (L_x, Y_{SZ}) with the correlation coefficient set to $r = 0.0$.

mock cluster sample is then built based on the mock measured luminosities and an applied luminosity threshold. We consider three cases:

1. there are statistically small measurement uncertainties $\sim 10\%$ on all observables.
2. the measurement uncertainties are larger as in the eDXL measurements.
3. the intrinsic scatters on response variables are larger and keeping the measurement uncertainties similar to realistic measurement uncertainties as in point (ii). We call this eDXL like samples due to the recovered scatters found in our analysis. The input scaling relation is shown in Table A.1.

The mock sample size is kept fixed to 30 clusters at redshift 0.3 (median redshift of eDXL sample). Mock data were generated from the input scaling relations given in Table A.1. For these 30 clusters, the Y_{SZ} was simultaneously generated using a correlated intrinsic scatter in Y_{SZ} and L_x using the `mrandomn` IDL routine.

Small measurement uncertainties The sample with “small” errors were generated with 10% measurement scatter. We simulate the datasets consisting of 30 clusters with 10% measurement scatter in each observable which is more precise than our APEX-SZ measurements. We assume a set of input scaling relations parameters for the $L_x - M_{500}$ and $Y_{\text{SZ}} - M_{500}$. An example set of mock dataset is shown in Figure A.1 for uncorrelated ($r = 0$) intrinsic scatters at fixed mass. Figure A.2 shows mock data set of clusters for input correlation, $r = 0.8$.

Realistic measurement uncertainties The samples with “real” errors were generated by using

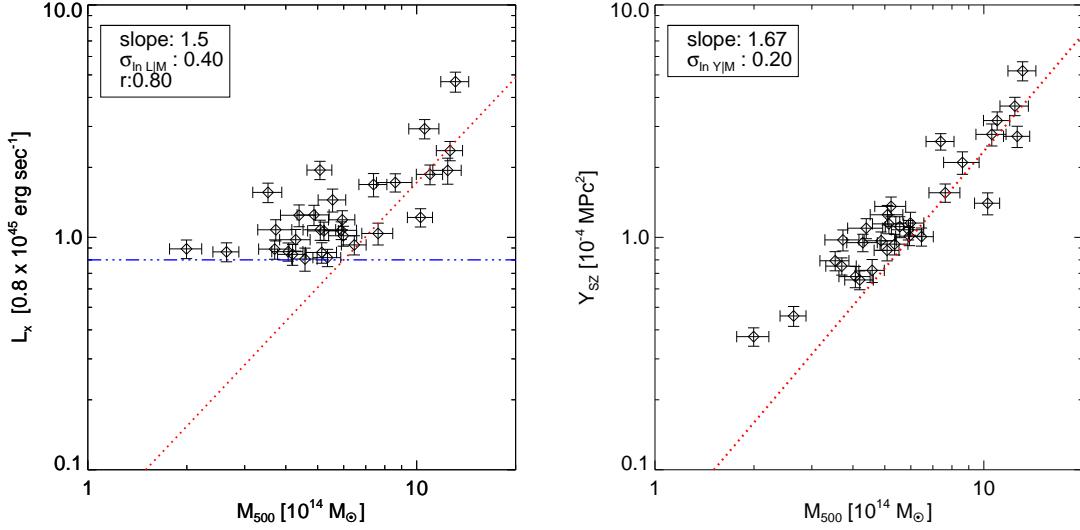


Figure A.2: An example set of simulated sample with their mock measurements at redshift $z = 0.30$ is shown here. The sample was selected by applying threshold selection (blue horizontal line) in measured luminosities. The simulations of mock data included a correlated intrinsic scatters in the two response variables (L_x , Y_{SZ}) with the correlation coefficient set to $r = 0.80$.

the median of $\sigma_{\tilde{M}_{WL}}/\tilde{M}_{WL}$, $\sigma_{\tilde{Y}_{SZ}}/\tilde{Y}_{SZ}$, $\sigma_{\tilde{L}_x}/\tilde{L}_x$ and keeping this ratio for the generated sample. For realistic measurement uncertainties, we use $L_x - M_{500}$, $Y_{SZ} - M_{500}$ two sets of scaling parameters. The first is identical input relations as used in simulating mock samples with small measurement uncertainties. For the second set of scaling parameters, we inject higher intrinsic scatters in L_x and Y_{SZ} . The values of these scaling parameters were chosen based on the eDXL sample fit values from Section 6.4.1.

Numerous realisations of mock samples for all of the above cases were generated for a range of input values of r . The analysis of these mock samples is presented in the next Section.

A.3.1.2 Analysis of scaling relations fit to mock data

For each input relations and r values given in Table A.1, we fit each mock sample with the model described in Section 6.3.2.1 and the normalised likelihood given in Appendix A.1. For all cases, we set $r = 0$ while fitting mock data. This is to investigate the level of bias in scaling parameters if the r parameter is ignored in the joint fit of the $L_x - M_{500}$ and $Y_{SZ} - M_{500}$ relations. The results are given in Table A.1. The means reported are an average of recovered modes of scaling relation parameters from fit to different mock samples (typically ~ 35 - 60 samples). The uncertainty quoted on each parameter is the standard deviation of the sample distribution of recovered modes. This corresponds to an uncertainty in the scaling relation parameter from fitting a single set of mock cluster sample. Below, we define relative bias and its significance.

$$\text{relative bias} \times 100 = \frac{(\text{recovered} - \text{input})}{\text{input}} \% \quad (\text{A.16})$$

Table A.1: The bias in recovered parameters of the scaling relations for the simulated cluster samples with different input correlation coefficient, r , values. The variable, r , refers to the correlation in the intrinsic scattering of L_x and Y_{SZ} at fixed mass. The bias in the SZ scaling parameters grows significant with strong correlation coefficient. The relative bias and significance quoted here were calculated using the equations (A.16) and (A.17)

r	A_{LM}	B_{LM}	$\sigma_{\ln L}$	A_{SZ}	B_{SZ}	$\sigma_{\ln Y_{\text{SZ}}}$
input	0.80	1.50	0.40	1.0	1.67	0.20
small measurement uncertainties						
0.0	0.789 ± 0.137	1.471 ± 0.238	0.400 ± 0.096	0.996 ± 0.051	1.676 ± 0.159	0.189 ± 0.061
rel. bias	- 1.4 %	-2.0 %	0.0 %	-0.4 %	0.4 %	-5.3 %
d	- 0.08 σ	-0.12 σ	0.00 σ	-0.09 σ	0.04 σ	- 0.17 σ
0.1	0.832 ± 0.117	1.481 ± 0.198	0.392 ± 0.078	1.022 ± 0.057	1.642 ± 0.135	0.171 ± 0.064
rel. bias	4.0 %	-1.3 %	- 2.0 %	2.2 %	- 1.6 %	- 14.5 %
d	0.27 σ	-0.10 σ	- 0.10 σ	0.39 σ	- 0.20 σ	- 0.46 σ
0.5	0.822 ± 0.110	1.535 ± 0.213	0.398 ± 0.081	1.118 ± 0.066	1.476 ± 0.145	0.173 ± 0.052
rel. bias	2.8 %	2.4 %	- 0.5 %	11.8 %	- 11.6 %	- 13.5 %
d	0.20 σ	0.16 σ	-0.03 σ	1.79 σ	- 1.34 σ	- 0.62 σ
0.8	0.861 ± 0.129	1.525 ± 0.210	0.350 ± 0.091	1.161 ± 0.058	1.466 ± 0.110	0.121 ± 0.056
rel. bias	7.7 %	1.7 %	- 12.5 %	16.1 %	- 12.2 %	- 39.3 %
d	0.48 σ	0.12 σ	- 0.55 σ	2.76 σ	- 1.91 σ	- 1.40 σ
realistic measurement uncertainties						
0.1	0.926 ± 0.190	1.516 ± 0.240	0.360 ± 0.123	1.014 ± 0.118	1.568 ± 0.398	0.22 ± 0.110
rel. bias	15.7 %	1.1 %	- 10.6 %	1.4 %	- 6.1 %	12.7 %
d	0.66 σ	0.07 σ	- 0.34 σ	0.12 σ	- 0.26 σ	0.23 σ
0.6	1.014 ± 0.133	1.438 ± 0.298	0.310 ± 0.076	1.128 ± 0.132	1.418 ± 0.331	0.203 ± 0.088
rel. bias	26.8 %	- 4.1 %	- 23.2 %	12.8 %	- 15.1 %	1.6 %
d	1.61 σ	- 0.21 σ	- 1.23 σ	0.97 σ	- 0.76 σ	0.04 σ
eDXL-like samples						
input	0.30	1.50	0.60	1.0	1.67	0.40
$r = 0.3$	0.37 ± 0.12	1.61 ± 0.29	0.54 ± 0.11	1.11 ± 0.14	1.70 ± 0.42	0.29 ± 0.17
rel. bias	23.3 %	7.3 %	- 0.1	11.0 %	1.8 %	- 27.5 %
d	0.58 σ	0.38 σ	- 0.55 σ	0.79 σ	0.07 σ	- 0.65 σ
$r = 0.6$	0.440 ± 0.150	1.428 ± 0.297	0.453 ± 0.167	1.396 ± 0.180	1.318 ± 0.370	0.267 ± 0.134
rel. bias	46.7 %	- 4.8 %	- 24.5 %	39.6 %	- 21.1 %	- 49.6 %
d	0.94 σ	- 0.24 σ	- 0.88 σ	2.20 σ	- 0.95 σ	- 0.99 σ

$$\text{Detection level of the bias} = d = \frac{\text{relative bias}}{\frac{\Delta_{\text{recovered}}}{\text{input}}} \quad (\text{A.17})$$

The relative bias and significance are quoted in Table A.1. The significance here is a measure of an average bias level w.r.t. the uncertainties on the recovered parameter for any cluster sample.

A.3.2 Mock data with intrinsic scatter in the lensing mass

The procedure for the mock data generation is identical to the previous section with couple of exceptions. We introduce scattered lensing masses (M_{WL}) from true halo mass ($M_{HM,500}$). We introduce a 20% log-normal scatter between the halo mass (M_{HM}) and the weak-lensing mass. All observables are generated in similar manner as in previous Section using the scaling relation models defined in Section 6.3.2.2. We also use the redshift distribution as well as the redshift dependent luminosity thresholds of the eDXL sample. We simulate with small and real measurement uncertainties as done before.

Small measurement uncertainties We use mock samples with 10% measurement uncertainties to test for bias. The mock samples include zero correlation in intrinsic scatters of L_x and Y_{SZ} . We use the identical method used in Appendix A.3.1.2 and fit the scaling relations with $r = 0$. The mean recovered parameters from numerous mock samples and the standard deviations are quoted in Table A.2. We observe significant bias in the slope and intrinsic scatter parameters of the $Y_{SZ} - M_{500}$ scaling relations. On the $L_x - M_{500}$ relation the bias in scaling parameters are lower than their confidence level obtained for single mock data set.

Realistic measurement uncertainties We fit the scaling relations parameters to the mock samples with realistic measurement uncertainties using the same method as in Appendix A.3.1.2. The mock samples considered include zero correlation in intrinsic scatters of L_x and Y_{SZ} . As done in previous section, we fit these mock samples by ignoring the r parameter. The average of recovered modes and the standard deviation of the sample distribution of recovered modes are summarised in Table A.2. We inject $r = 0.6$ for the mock samples along with the 20% intrinsic scatter in lensing mass. We fit these mock data while ignoring r and the scatter in lensing mass. At r equals 0.6, the total bias in the normalisation is significant at a level of 2.7σ with respect to uncertainties expected for a single mock data set. The slope parameter of the $Y_{SZ} - M_{500}$ relation is biased low by $\sim 1\sigma$. We re-fit the mock samples with $r = 0.6$ fixed in the fitting routine. The results are shown in the final row of Table A.2. We observe that the bias in the normalisation drops to 0.5σ which is similar to the bias obtained in the mock samples that were generated with $r = 0$. The bias in the slope value reduces from 0.34 to 0.17.

Table A.2: Scatter between m_{lens} and m_{halo} assumed to be 0.20 (20 %). The redshift dependent luminosity cuts were applied to the set of mock data used here.

r	A_{LM}	B_{LM}	$\sigma_{\ln L}$	A_{SZ}	B_{SZ}	$\sigma_{\ln Y_{\text{SZ}}}$
Input	0.3	1.5	0.6	1.0	1.67	0.4
small measurement uncertainties						
$r = 0$	0.235 ± 0.123	1.476 ± 0.238	0.676 ± 0.151	1.118 ± 0.133	1.356 ± 0.225	0.511 ± 0.079
rel. bias	- 21.6 %	- 1.6 %	12.6 %	11.8 %	- 18.8 %	27.7 %
d	$- 0.53\sigma$	$- 0.10\sigma$	$+ 0.50\sigma$	0.89σ	$- 1.40\sigma$	1.41σ
realistic measurement uncertainties						
$r = 0$	0.33 ± 0.15	1.44 ± 0.22	0.60 ± 0.12	1.09 ± 0.16	1.55 ± 0.32	0.44 ± 0.12
rel. bias	10.0 %	- 4.0 %	- 0.0 %	9.0 %	- 7.0 %	10.0 %
d	0.2σ	$- 0.3\sigma$	$- 0.0\sigma$	0.6σ	$- 0.4\sigma$	0.3σ
$r = 0.6$	0.434 ± 0.213	1.406 ± 0.254	0.501 ± 0.164	1.590 ± 0.217	1.326 ± 0.322	0.342 ± 0.128
rel. bias	44.7%	- 6.3%	- 16.5%	59.0%	- 20.6%	- 14.5%
d	0.63σ	$- 0.37\sigma$	$- 0.60\sigma$	2.72σ	$- 1.07\sigma$	$- 0.45\sigma$
$r = 0.6$ (fixed)	0.34 ± 0.15	1.43 ± 0.22	0.60 ± 0.13	1.10 ± 0.21	1.49 ± 0.21	0.45 ± 0.16
d	0.27σ	$- 0.32\sigma$	$- 0.00\sigma$	$+ 0.50\sigma$	$- 0.86\sigma$	$+ 0.31\sigma$

List of Figures

2.1	MOO J1142+1527, a galaxy cluster at 8.5 billion light years away discovered using the <i>Wide-field Infrared Survey Explorer (WISE)</i> . The image shows infrared observations from the <i>Spitzer Space Telescope</i> combined with near-infrared and visible light seen by the <i>Gemini Observatory</i> . The cluster of red galaxies in the center of the image forms the core part of the cluster. The blue halo shown is the Sunyaev-Zel'dovich effect tracing the thermal gas medium within the galaxy cluster observed with the <i>Combined Array for Research in Millimeter-wave Astronomy (CARMA)</i>	5
2.2	Multi-wavelength composite image of the Abell 520 galaxy cluster. The multi-components are overlaid on the optical image using false colors. X-ray emission from gas medium (Green); Dark matter (Blue); Starlight (Orange).	6
2.3	The dimensionless comoving volume elements for three different cosmologies as a function of redshift.	9
2.4	Time evolution of a massive halo in Millennium II simulations.	11
2.5	Halo mass function for spherical overdense masses from hydrodynamical and dark matter only simulations.	15
2.6	Measured mass function (data points) and model prediction (solid curves).	16
2.7	The distortion caused by the hot ICM in the CMB spectrum. The CMB spectrum is a black body (black solid). To demonstrate the distortion of the spectrum, an unrealistically large Compton parameter ($y = 0.15$) is chosen and this spectrum is represented by the grey dotted-dashed curve. The typical distortion amplitude for galaxy clusters are estimated in the order of 10^{-4} (orange dashed curve). In order to distinguish the distorted spectrum from black body, (right) the differential intensity is shown.	19
2.8	Cluster catalogs from SZ and X-ray selections. The 516 optically confirmed clusters from SPT catalog, 91 clusters from ACT survey, 809 SZ-selected clusters from Planck, 740 X-ray clusters ¹ from ROSAT observations are shown here. The SPT selection is approximately mass selected, whereas the X-ray selection suffers from dimming effect with redshift, the Planck selection suffers from dilution of SZ signal due to the beam.	20
2.9	A sketch of a gravitational lens system. β is the angle between the source S and the optical axis. The angle between the image and the optical axis at the observer is θ . The deflection angle of the light ray at the lens is $\hat{\alpha}$. α is the reduced deflection angle of the image at the observer.	23

- 2.10 Covariance of cluster global properties from ~ 4500 halos in the Millenium Gas Simulations. Off-diagonal terms show the pair-wise correlation between two cluster properties at fixed mass. Two physical processes were considered. A gravity only process (blue) and a pre-heating scenario (red), where the gas was considered to have an entropy floor of 200 keVcm^2 at $z = 4$. The diagonals show the natural logarithmic deviations from the mean mass-scaling relation of each property. 27

2.11 Systematic effects in cosmological constraints from galaxy clusters due to inaccuracies in the mass calibration of cluster observables. Green, blue and violet contours are 68% and 95% confidence regions of the constraints on the parameters obtained by using mass calibration from WtG, CCCP and CMB lensing estimates respectively. The systematic shift in the confidence levels from using the different mass calibrations demonstrate that the cluster abundance constraints on cosmology are currently limited by the uncertainties in the accuracy of mass-observable calibration. 28

2.12 The constraint on σ_8 from different work using galaxy clusters are represented. The first five constraints are from X-ray selected clusters. The constraints from using an optically selected cluster catalog is represented as SDSS. SPT, ACT and Planck denote constraints from SZ selected samples. 29

2.13 *Top:* A mock sample of clusters with X-ray luminosities vs. true mass generated using the Tinker mass function and a $L_x - M$ scaling relation (black solid line) are shown. As predicted by the mass function, more massive objects are rarer than low mass systems. The intrinsic scatter in X-ray luminosity mixes the populations of different mass bins. On applying a threshold cut in luminosity (blue dashed line), this selects the intrinsically bright sources at low mass end. Together, they produce a sample that is biased towards the low mass and intrinsically bright objects. *Bottom:* The sample generated above was used to also generate the Y_{SZ} for each source when (*left*) luminosity and Y_{SZ} are uncorrelated in their scatters at fixed mass, and (*right*) luminosity and Y_{SZ} are strongly correlated in their scatters at fixed mass. 31

3.1 The extended DXL (eDXL) sample selection. The sample is selected in the luminosity-redshift plane from ROSAT catalogues. Above redshift equals 0.27, the sample was selected from REFLEX II catalogue. The low redshift sample is selected from NORAS and REFLEX. In total, 30 galaxy clusters are selected. The grey rectangular outlines enclose the sample selection plane. The details on optical follow-up observations of this sample are given in Section 3.3. As indicated in this figure, three galaxy clusters are missing in the follow-up program. The dotted-dashed curve represent the nominal flux limit (1.8×10^{-12} erg per sec) of the REFLEX II catalogue in the luminosity-redshift plane. This curve corresponds to approximately 80% completeness. The dashed curve is the luminosity curve for higher flux limit (3×10^{-12} erg per sec), which represent the 90% completeness of the REFLEX II catalogue. At low redshift, this curve corresponds to 80% completeness of the REFLEX catalogue. 35

4.1	<i>Left:</i> An input model used in the simulations of mock SZ images. Here, a β profile for the SZE profile was assumed with a core radius equal to the beam size i.e. $1'$ and a temperature decrement of $-0.7 \text{ mK}_{\text{CMB}}$. <i>Middle:</i> The input model was convolved with the point source function which is a linear representation of the filtering process of the reduction pipeline. This gives the filtered SZ image shown here. <i>Right:</i> One of the jack-knived noise realisation of the APEX-SZ observations of the Bullet cluster field was added on top of the filtered SZ image as an example. This is a typical realistic mock image of APEX-SZ observations of a galaxy cluster's SZ intensity.	55
4.2	Same as Figure 4.1 for an input model with core radius of 1.5 arcmin	55
4.3	Same as Figure 4.1 for an input model with core radius of 2.2 arcmin	56
4.4	Schematic description of the Fourier domain fitting.	57
4.5	The noise power spectral density (PSD) in the Bullet cluster field inside three different central areas. The PSD is a measure of average noise power in the region and this is essentially the noise covariance matrix for the χ^2 . A higher PSD implies the weights for the spatial scales are lower in the χ^2	58
4.6	Schematic representation of the radial binning method. This method takes into account the variations in the pixel weights across the APEX-SZ maps.	60
4.7	Bin-to-Bin correlation matrix. <i>Left:</i> White noise. The correlation matrix is essentially a diagonal matrix. <i>Right:</i> Bullet cluster field noise. The noise correlation matrix for the real noise show correlations between neighbouring bins due to the telescope beam. The anti-correlation between the noise in inner and outer bins due to the low-order polynomial baseline applied to the APEX-SZ observations.	61
4.8	Model 3: Distribution of modes for each parameter from multiple mock data. The red lines indicate the input parameter values used for generating mock data with white noise.	63
4.9	Model 4: Distribution of modes for each parameter from multiple mock data. The red lines indicate the input parameter values used for generating mock data with white noise.	64
4.10	Model 3 & 4: Bootstrapped confidence region of the mean value recovered from mock fits for ΔT_0 and r_c . The black solid line indicates the input value of the parameter used for simulating the mock Compton-y images. The red dotted lines represent the 68% confidence region of the distribution. The input values are well-within this confidence region on the average estimated value.	65
4.11	Results for Model A. The distribution of the recovered modes from fitting each of the mock images generated using the Bullet noise fields.	66
4.12	Same as Figure 4.11 but for Model B with core radius of $1.5'$. Although it appears as though the distribution of the recovered core radius is highly skewed and shows some lack of preference for smaller core radius, it is not to be concluded that this is a hint of a biased distribution. Rather this is an indication of incompleteness of the noise realizations limited by the integration time. A similar model was used to test with another noise field (XMM-LSS) and that did not show such level of skewness.	68
4.13	Same as Figure 4.11 but for Model C with core radius of $2.17'$	69
4.14	The bootstrapped distribution of the mean of the sample distributions (shown for each model in figures 4.13, 4.12, 4.11) of core radius and peak temperature decrement of the beta model obtained from the Bullet noise simulated mock data. The black solid line indicates the input values of the parameters, which is within the 68% confidence region (red dotted lines) on the average estimated value.	70

- 4.15 Schematic representation of equal areas of circular and square regions. The smaller radius circle and the smaller length box represent a smaller region typically used in the fitting with the FD method. The larger circle and square represent the region of data that is still available for fitting. The smaller area is referred to by the symbol **S**, and the larger area is represented with the symbol **L** for simplicity. 72
- 4.16 Comparison of the FD and the RB methods using Model 1, and 2 mock data. Results from one mock fit from each model are shown. The contours show the 68% and 95% confidence regions of the parameters from one mock fit. The dashed curves represent the iso-Compton- Y regions (the parameter space which give the same value for Y). The magenta point indicates the input value for the simulated model. 73
- 4.17 *Left:* The parameter covariance for one example Model 3 mock image. The parameter constraints from the radial binning method (blue), and the Fourier domain method (red) are shown. The parameters recovered by both methods are consistent. The input value used for the simulation is shown by the magenta point. The iso-Compton- Y regions in the parameters space is shown as the dotted-dashed lines. *Right:* The marginalised posterior distribution of the Compton- Y from both the methods for the parameter covariance in the left figure are shown here. Not surprisingly, the posterior distribution from both the methods are quite similar. 74
- 4.18 *Left:* Example parameter constraints from mock data of each input model A, B and C. The input values of the parameters, used for the mock data, are indicated by the magenta point. The 68% and 95% confidence regions of the constraints obtained by using the Fourier domain method (red), and radial binning method (blue) are shown. *Right:* The obtained constraints on the Compton- Y for each example shown on the left is shown. The black curve shows the marginalised posterior distribution of the Y obtained from radial binning method. The red curve correspond to the same for the Fourier domain fit. The magenta line indicates the Y value corresponding to the input model for the mock data. The parameter constraints recovered by each of the two methods from fitting mock images with real noise prefer slightly different orientations. The iso-Compton- Y in the parameter space are indicated by dashed lines. The area of the map used by two methods are kept the same. The input values used for the simulation are shown by the magenta point. 75
- 4.19 The recovered parameter covariance from the two fitting methods. The noise covariance matrix of the radial binning method included only diagonal elements, the off-diagonal elements were purposely ignored. 76
- 4.20 The figures A, B, C show the distribution of the best fit values for Compton- Y parameter from the two methods. Integrated Compton parameter Y for the sources with the three different input parameters given in table 4.1. The yellow shaded region is the distribution from the radial binning g method and the histogram filled with blue lines is the distribution recovered from the Fourier domain fitting method. The same set of sources are used for both cases. In all of the cases there seems to be no bias in recovering the Y parameter. The input value of the Y parameter is shown by the black solid line. 77

4.21	The figures A , B and C show the scatter plot of the recovered best fit values of the parameters (core radius and the ΔT_0) of the β model. The input values are indicated by the black cross point. The blue diamond points are the best fit values from the radial binning method and the red points are the best fit values from the Fourier domain method. The ellipses represent the 68% confidence level covariance. These scatter plots demonstrate that there is no qualitative and quantitative difference in the parameter space traced by best-fit values from both the methods.	78
4.22	Parameter covariance of the β model from the two fitting methods and two areas of the map used to fit. <i>Left</i> : Radial binning method (blue) and Fourier domain (red) fitting applied within 552 square arcmin region. <i>Right</i> : Parameter covariance for the same example shown in left figure obtained from the radial binning method applied within 552 square arcmin region (blue) and the Fourier domain fitting applied to 100 square arcmin region (red). The Fourier domain fitting within larger area estimates larger covariance of the parameters than within a smaller area, which is counter to the expectation that the constraints should improve when adding more information. Whereas, the radial binning method when applied to this larger area of the image data constrains the parameter space more precisely than either method applied within a smaller region of the image data.	79
4.23	Distribution of uncertainties on the integrated Compton parameter recovered from fitting to multiple mock realizations of filtered maps. Four different ways of fitting was used. Blue histograms show the uncertainties from radial binning (RB) method applied in large area of map L . Green histograms are obtained for the Fourier domain (FD) fitting in large area. Red histograms are obtained for FD fitting within smaller region S . Black histograms represent the results from RB method applied in small region. On increasing the size of the area of the map used for the fitting, the uncertainties estimated from the two methods show a disagreement. The uncertainties decreases with the area of the map used for the fitting in the radial binning approach. The uncertainties are estimated to be larger in the FD method approach due to the failure of non-stationarity condition in the larger area of the map.	80
5.1	Bin-to-bin correlation matrix for the Bullet cluster field. The bin-width is 1'.	84
5.2	$Y_{\text{sph},500}$ measurements vs. the weak-lensing spherical mass estimates within $R_{500,\text{WL}}$. The non-detections in the integrated Compton parameter are indicated by upper limits at a 2σ level (downward arrows).	85
5.3	Mass estimate from joint fitting analysis vs. the weak-lensing estimate used as the prior. The joint fitting yields generally consistent mass estimate with the input prior values for all clusters with the exception for one cluster, Bullet. Bullet is observed with highest signal to noise in the APEX-SZ maps, and given that this is highly disturbed cluster there is some inconsistency in the input prior mass and the preferred estimate from joint analysis for this cluster.	87
5.4	<i>Left</i> : The angular offset between the centres from ROSAT X-ray and the centres adopted for weak-lensing mass estimate are shown for the 27 eDXL clusters. 80% of the clusters have an offset between optical and X-ray centres that is smaller than half of the APEX-SZ beam. RXCJ1135 shows by far the largest offset of 2.5 arcmin. <i>Right</i> : Integrated Comptonization measured with X-ray centroid vs. optical BCG. The solid line is the 1:1 relation.	88

5.5	$Y_{\text{sph},500}$ measured from two pipelines (BoA and MATLAB (Bender et al. 2016)) are shown here along with their uncertainties and upper limits wherever appropriate. The correlation between the two pipelines using the Kelly method shows that the measurements are statistically consistent for the 41 clusters in the sample. The solid line is the best-fit relation and the dashed line is the one-to-one relation.	91
6.1	<i>Left:</i> ROSAT luminosities $L_x[0.1 - 2.4 \text{ keV}]$ and weak-lensing spherical masses for the eDXL sample with lensing follow-up. The spherical weak-lensing masses are measured within R_{500} . The luminosities are measured within a R_{500} that is independent of the weak-lensing analysis (See section 3.2.3). <i>Right:</i> The integrated Comptonization from APEX-SZ vs. weak-lensing mass estimate for the X-ray selected eDXL sample.	96
6.2	The recovered mean parameters from mock data realisations are shown here for the $Y_{\text{SZ}} - M_{500}$ relation. The red lines mark the input values of the scaling parameters. The error bar on the recovered mean value represents the uncertainty level from single set of mock sample analysis. <i>Left:</i> The black triangular symbols correspond to results from mock data with 10% measurement uncertainties. The blue diamond symbols correspond to results for mock data with realistic uncertainties (See text). <i>Right:</i> The recovered mean parameters for different input scaling relations, where $\sigma_{\ln L_x}$ and $\sigma_{\ln Y_{\text{SZ}}}$ were increased to 0.6 and 0.4 respectively. We simulated the mock data with realistic uncertainties. The bias we see in the normalisation for realistic mock samples indicates that even for cluster sample size of 30 and with realistic measurements, ignoring the correlation show significant bias.	102
6.3	An example of 68% and 95% confidence levels of the normalisation and slope parameters of the $Y_{\text{SZ}} - M_{500}$ scaling relation fit to a mock sample. Red dashed contours: $r = 0$, and the black solid contours: $r = 0.6$. The mock sample was generated using realistic measurement errors and with a correlation ($r = 0.6$) in intrinsic scatters of L_x and Y_{SZ} . The input value of the $Y_{\text{SZ}} - M_{500}$ relation is shown as cross.	103
6.4	Marginalized posterior distributions of the parameters of the scaling models, shown as 1σ and 2σ confidence levels. Colour shaded contours are from the full joint likelihood fit including the correlated intrinsic scatter r with a uniform prior such that $-1 < r < 1$. Red contours indicate the marginalized confidence levels with $r = 0$. The black contours indicate the marginalized confidence levels with $r = 0.5$. The histograms show the marginalised distribution of each recovered scaling parameter when we allow r to vary. The vertical dotted line corresponds to the self-similar expectation of the $Y_{\text{SZ}} - M_{500}$ scaling slope.	107
6.5	The best-fit $Y_{\text{SZ}} - M_{500}$ scaling relation including correlated intrinsic scatter is shown as the solid blue line. The red dashed line shows the fitted scaling relation assuming no correlated scatter. The red and blue line-filled regions denote 68% confidence. Assuming un-correlated intrinsic scatters in luminosity and Comptonization finds a higher normalisation and shallower slope for the $Y_{\text{SZ}} - M_{500}$ relation. The measured weak-lensing masses and integrated Comptonizations plotted here are same as for the eDXL sample shown in Figure 5.2.	108

6.6 Luminosity-mass relation: The luminosities and our scaling relations are measured in the energy band 0.1–2.4 keV. The scaling relations are represented for the luminosity-mass relation for the same energy band. The grey shaded region represents the 68 % confidence level of our scaling relation. The measured values are generally up-scattered from the best-fit relation. This is due to the Eddington & Malmquist biases in the sample which is corrected for in the scaling relation determination through our Bayesian analysis. At the median mass of the sample, our constraints mean a correction factor of 1/3 to a biased estimate from generalizations of ordinary least square methods (e.g., BCES). Previous literature measurements of the luminosity-mass relation for the luminosity in the same energy band of 0.1 – 2.4 keV are plotted for comparison with our constraints. This is discussed in Section 6.6.1.	109
6.7 Same as Figure 6.4 but for $\sigma_{WL HM} = 0.20M_{HM}$. The red contours are marginalized recovered confidence level for $r = 0.0$, and black contours for $r = 0.5$. The histograms show the marginalised distribution of parameters recovered from varying r	110
6.8 <i>Top panel:</i> Luminosity and Comptonization residual data (blue square) after subtracting the best-fit $Y_{SZ} - M_{500}$ and $L_x - M_{500}$ scaling relations for each of the 24 clusters that were detected with APEX-SZ. The contours show the 1σ and 2σ confidence level prediction from random realisations of simulated residuals (grey points) for the best-fit intrinsic covariance model. <i>Lower panel:</i> distribution of residuals of 24 clusters, showing a positive correlation. The Pearson correlation coefficient in residuals is 0.73. The contours represent the 1σ and 2σ confidence levels obtained by combining the distributions shown in the top panel.	112
6.9 <i>Left:</i> Same as Figure 6.8, however, the residuals correspond to best-fit relations for the $r = 0$ covariance model. <i>Right:</i> Distributions of Pearson correlation coefficients in simulated realisations of 24 cluster residuals with un-correlated ($r = 0$) and correlated ($r = 0.5$) intrinsic scatters. The vertical dashed line indicates the measured Pearson correlation coefficient (0.65) in the residual data.	113
6.10 For each cluster we compare the measured mass estimate to the probability distribution function prediction from our modelling.	114
6.11 The best-fit $Y_{SZ} - M_{500}$ scaling relation is shown as the solid black line, with the 68% confidence level indicated by the grey shaded region. The constraints shown here were obtained by ignoring the weak-lensing intrinsic scatter. The scaling relations found in other works that used weak-lensing mass to calibrate total mass within R_{500} are overplotted.	119
6.12 Same as Figure 6.11 but comparing to previous works using mass proxies other than weak-lensing mass. The prediction from simulation work of Sembolini et al. (2013) is also shown for comparison.	120
6.13 Comparison of constraints on the slope of $Y_{SZ,\Delta} - M_\Delta$ scaling relation between the literature and this work. The black solid line is our best-fit slope value and the corresponding 1σ confidence is represented as the grey shaded region. The dashed line represents the self-similar slope value. The 1σ confidence on the slope constraints obtained with $r = 0.0$ is indicated by the line filled region. We inverted the slope constraint for certain works which originally provided results for $1/B_{SZ}$; these are represented as stars. The asterisk on Marrone et al. (2012) represents their result without A383.	121

A.1 An example set of simulated sample with their mock measurements at redshift $z = 0.30$ is shown here. The sample was selected by applying threshold selection (blue horizontal line) in measured luminosities. The simulations of mock data included a correlated intrinsic scatters in the two response variables (L_x , Y_{SZ}) with the correlation coefficient set to $r = 0.0$	139
A.2 An example set of simulated sample with their mock measurements at redshift $z = 0.30$ is shown here. The sample was selected by applying threshold selection (blue horizontal line) in measured luminosities. The simulations of mock data included a correlated intrinsic scatters in the two response variables (L_x , Y_{SZ}) with the correlation coefficient set to $r = 0.80$	140

List of Tables

3.1	Luminosity cuts of the eDXL sample in different redshift bins. The selection is represented graphically in Figure 3.1. The luminosities quoted here are computed in the energy range [0.1 – 2.4] keV. In the final column, we give the number of clusters in each redshift range. In bracket, we mention the actual number of clusters that were completely followed-up in our multi-wavelength observations.	36
3.2	The 27 targets of the eDXL sample. Each cluster position is taken to be the position of the BCG in the optical data. The weak-lensing derived masses ($M_{\text{WL},500}$), and the re-measured ROSAT luminosities (L_x) are reported.	40
3.3	Same as table 3.2, but for other clusters in the full APEX-SZ sample that are not part of the X-ray selected complete sample.	41
4.1	Summary of input parameters of the β model, type of filter function and type of noise used for mock data.	54
4.2	Summary of recovered input values from different mock realizations. The fourth column gives the mean deviation from the input parameter value (p_{input}) and 68% confidence level about the mean deviation of the sample distribution. The final column gives the mean deviation and its confidence level determined by bootstrapping the sample mean.	67
4.3	Results on Y estimations from fitting multiple realizations of mock APEX-SZ maps. .	71
4.4	ΔT_0 parameter estimated from numerous mock realizations of APEX-SZ mock maps. The input value of ΔT_0 was $-0.7 \text{ mK}_{\text{CMB}}$	71
4.5	r_c parameter estimated from numerous mock realizations of APEX-SZ mock maps. .	71
5.1	The 27 targets of the eDXL sample. Each cluster position is taken to be the position of the BCG in the optical data. The integrated Compton-y parameters ($Y_{\text{SZ},500}$), the weak-lensing derived aperture (R_{500}) are reported.	86
5.2	Same as table 5.1, but for other clusters in the full APEX-SZ sample that are not part of the X-ray selected complete sample.	87
5.3	A summary of the components of the analysis used in A. Bender's work	89
5.4	Y_{sph} from Bender et al. ⁰ and this work.	90
6.1	Results of the scaling relations analysis for the eDXL sample using the method described in Section 6.3. The medians and 68% confidence levels of the marginalised distributions are quoted. Centroids for gNFW model fits were fixed to the optical centres (BCG) or X-ray (X) values obtained from the ROSAT survey. The results in the first row (shaded) will be considered as our fiducial result.	106

A.1	The bias in recovered parameters of the scaling relations for the simulated cluster samples with different input correlation coefficient, r , values. The variable, r , refers to the correlation in the intrinsic scattering of L_x and Y_{SZ} at fixed mass. The bias in the SZ scaling parameters grows significant with strong correlation coefficient. The relative bias and significance quoted here were calculated using the equations (A.16) and (A.17)	141
A.2	Scatter between m_{lens} and m_{halo} assumed to be 0.20 (20 %). The redshift dependent luminosity cuts were applied to the set of mock data used here.	143

Copyright

The permission to use the contents of the paper from the journal MNRAS is stated in the link https://academic.oup.com/journals/pages/access_purchase/rights_and_permissions/publication_rights under the section titled “Rights retained by ALL Oxford Journal authors” that reads as follows:

“The right to include the article in full or in part in a thesis or dissertation, provided that this not published commercially;”

Declaration

Hiermit versichere ich, dass die vorgelegte Arbeit – abgesehen von den ausdrücklich bezeichneten Hilfsmitteln – persönlich, selbständig und ohne Benutzung anderer als der angegebenen Hilfs- mittel angefertigt wurde. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle kenntlich gemacht und die vorgelegte Arbeit oder ähnliche Arbeiten sind nicht bereits anderweitig als Dissertation eingereicht worden. Frühere Promotionsversuche wurden nicht unternommen. Für die Erstellung der vorgelegten Arbeit und/oder die Gelegenheit zur Promotion wurde keine fremde Hilfe, insbesondere keine entgelt- liche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberatern/-vermittlern oder anderen Personen) in Anspruch genommen.

Acknowledgements

I would like to thank Prof. Dr. Frank Bertoldi for providing me with the opportunity to pursue my dream of doing research in Astrophysics and Cosmology. I am grateful for his continued support throughout the duration of this thesis. His unwavering support fed into boosting my confidence and persevering through this Ph.D. I thank Dr. Kaustuv Basu for his generous help with the research proposal for my DAAD application. I also thank him for offering me with opportunities to work on APEX-SZ and radio halo projects. I thank and acknowledge his guidance in the early phases of my Ph.D.

I would like to acknowledge financial support from DAAD and IMPRS that enabled me to pursue this Ph.D. I also would like to acknowledge the stipend made available by Prof. Frank Bertoldi and University of Bonn in the last 3.5 years, which made completion of this thesis possible.

I am grateful for all the timely help, support provided by Dr. Martin Sommer, which in many ways were crucial in bringing the project and thesis to an end. I thank him for all the support on the technical aspects of the APEX-SZ data analysis and for introducing me to the BoA pipeline. There is not enough space here to fit in all of the times I could thank him for his detailed feedback on my writing of this thesis, research proposals, applications, paper, etc.

I am grateful for the opportunity to work with many collaborators. The work done in this thesis uses the APEX-SZ observations and this work also benefited from discussions and support from the APEX-SZ collaboration. I thank Dr. Reinhold Schaaf for providing support for the APEX-SZ data analysis software (BoA) and for helping with the installation of AIPS.

I thank Prof. Dr. Thomas Reiprich for always being inclusive and for including me in his group activities, lunches and meetings. This ended up being an important aspect to me throughout this Ph.D. and especially in the last couple of years, as it gave me some anchor that made surviving this Ph.D. possible. I am especially grateful for his patience and the promptness with which he responds to my significant and insignificant questions without deference. I am even more grateful for all the times I could discuss the subject of my thesis and his continued interest from the very time he agreed to be the second advisor. This work has only benefited from his professional and critical feedback.

I appreciate all the professional guidance and help provided by Dr. Florian Pacaud. I thank him for his patience and for always finding time to discuss and qualify my ideas. His professional and critical feedback has been significant in improving the quality of this work. The generosity with which he has mentored and ushered me into fully developing the skills required to achieving success in research is something that I am truly grateful for. His attention to details and handling of these details has set a model for me to follow. I also thank him for his extraordinary help with careful reading of majority of this thesis and providing feedback.

I thank Prof. Adrian T. Lee for taking interest in this work and for being a constant force for pushing

Acknowledgements

things forward. I thank James for introducing me to statistics and for many insightful discussions. I am thankful to Douglas Applegate for helpful discussions on statistics. I would like to acknowledge the many helpful and insightful discussions with Florian Pacaud, James Kennedy, Matthias Klein, Adrian Lee, Martin Sommer.

I thank Christina Stein-Schmitz, for being a friendly face and for many ways in which she makes sure I am not without funding. And for having all the answers about the bureaucracy. I thank the IMPRS office for providing support regarding admin works and for providing opportunities for soft-skills development and travel funding.

I thank the current and past members of the SZ group: Jens, Miriam, Ana, Matthias, Sandra, for their interaction and some good times. I thank James, Jens, Kevin, Ana, for giving feedback on parts of the thesis and versions of the paper.

I am particularly grateful for the wonderful lectures that I got to attend in the beginning of this Ph.D., that helped in navigating through this forest of jargony astronomy.

I am fortunate to have found a good friend in Eugenia, who has been there through all these years ever since I came to Bonn. I have been really lucky to get nice office-mates. I thank both past and present office-mates for really good memories and a vibrant atmosphere: Matthias, Sandra, Sameera, Felix, Basilio.

My closest and dearest friends not in Bonn, Debo, Shravani, Ajinkya, Sanhita, Chitra and Ranjini for being a positive force and for never stop believing in me. Himadri, for all the Christmas market hopping and fun times and for being a good company.

Thanks to Chimmu, Manju and Rithika, I got to spend some weekends in a home away from home. I especially cherish all the times I got to remember my inner kid while spending time with my little niece Rithika.

I am deeply grateful and thankful for all the years of support and encouragement meted out to me by my loving parents and sister. Without their support, encouragement and patience, it would not have been possible to accomplish any of this. I also thank wholeheartedly my extended family members (Mohana perima, peripa and Arun for arranging my first flight ticket to Germany). All of their timely support has made possible for me to give my fullest to concluding the work presented in this thesis.