Cosmological Constraints from the South Pole Telescope Galaxy Cluster Survey

Tijmen de Haan

Department of Physics
McGill University
Montréal, Québec
Canada

September 9, 2014

A Thesis submitted to McGill University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy

Contents

List of Figures

List of Tables

Abstract

The South Pole Telescope (SPT) is a 10 m telescope located at the geographic south pole. From February of 2007 to November of 2011, we used the SPT to perform a five year survey of 2500 square degrees of the southern sky in millimeter waves. In this thesis, we describe work on the SPT project, such as digital feedback techniques for bolometer readout that are enabling next-generation instruments, including an implementation that is currently in use on the SPT. We discuss the 2500 square degree dataset, how to filter the data, generate maps, and extract a catalog of galaxy clusters. Using this list of hundreds of SPT-selected galaxy clusters we constrain cosmological parameters, noting that the scale of underlying cluster masses is the dominant systematic error. We present a novel Bayesian method for jointly fitting cosmological parameters as well as an arbitrary number of observable-mass scaling relations, in a computationally efficient way. Using this method, we compute constraints on cosmological models.

Résumé

Le Télescope du pôle Sud (SPT) est un télescope de 10 mètres de diamètre situé au pôle Sud géographique. De février 2007 à novembre 2011, nous avons utilisé le SPT afin de mener une enquête de cinq ans sur les quelque 2500 degrés carré carrés du ciel austral en ondes millimétriques. Dans cette thèse nous décrivons le travail accompli dans ce projet du SPT tel que les techniques de rétroaction numérique pour la lecture de bolomètres; ces techniques nous ayant permis de mettre à niveau des instrumentations comme celles déployées et présentement utilisées sur le SPT. Nous examinons l'ensemble des données obtenues, leur filtration, la production des cartes, puis l'extraction d'un catalogue d'amas de galaxies. En utilisant cette liste comprenant des centaines de galaxies sélectionnées par le SPT, nous contraignons les paramètres cosmologiques, ce qui nous amène à noter que l'échelle de masse d'amas de galaxies sous-jacente est l'erreur systématique dominante. Nous présentons une nouvelle méthode bayésienne pour contraindre les paramètres cosmologiques conjointement avec un nombre arbitraire de relations entre les observables et la masse, d'une manière efficace. En utilisant cette méthode, nous calculons les contraintes sur les modèles cosmologiques.

Contributions of Authors

Though this thesis was written by me, I am only one of around fifty to a hundred scientists working on the South Pole Telescope project. Even the subjects specifically covered in this dissertation contain invaluable contributions by many of my colleagues. In fact, some of the minor points presented here in order to maintain logical flow are not my work. Therefore, I will clarify some of the places where work was performed by others. In Chapter 2, while I played a major role in designing, testing, and deploying DAN and DEVB, the firmware implementation was almost entirely done by Graeme Smecher. An earlier version of this chapter was also submitted as conference proceedings as de Haan et al. (2012). Chapter 3 describes the SPT data analysis pipeline for galaxy cluster extraction. My contributions here are limited. I performed various pipeline tests, and implemented the atmospheric removal algorithm.

Conversely, I have worked on many projects throughout my Ph.D. and I cannot cover all aspects in this work. Some of the projects that go unmentioned include the following. In 2008, we found a single cluster candidate which could not be confirmed by ancillary observations in different wavebands. I worked on quantifying the probability of this non-detection (Brodwin et al., 2010). We also detected an extremely large decrement in the SPT maps in 2009. This cluster was later confirmed as SPT-CL J2106-5844, found to lie at very high redshift (z=1.13). I led the analysis effort of quantifying the "rareness" of such an extreme cluster in the surveyed volume (Foley et al., 2011; Stalder et al., 2013). I also deployed to the south pole to work on the instrument, where aside from menial tasks such as greasing the telescope, I made significant contributions to installing readout electronics, including the McGill digital frequency multiplexing platform and cryogenic thermometry readout.

Acknowledgments

First and foremost, I want to thank my supervisor Matt Dobbs for his ideas, knowledge and guidance. They were invaluable to this work and my development as a scientist in general. A huge thanks also goes out to my unofficial co-supervisor Gil Holder for many useful discussions and a seemingly exhaustive knowledge of cosmology and statistics.

I want to thank all the grad students, postdocs, and engineers I've had the pleasure of working with. Thanks also to the members of the SPT collaboration, especially Ryan Keisler and Brad Benson for their excellent leadership during south pole deployments.

Finally, a big thank you to my parents Arthur and Denise, and my sister Emma for their love and support over the years.

Cosmological Constraints from South Pole Telescope Data

1

Introduction

In this chapter, we will briefly introduce some of the basic concepts of our current knowledge of cosmology. The concepts and variables defined here form the underpinning of our (remarkable) understanding of cosmology today. We will introduce the concept of galaxy clusters, including their historical significance for the field of cosmology, and explain how the South Pole Telescope (SPT) can contribute to precision cosmology by counting their abundance on the sky. As a cosmological probe of the growth of structure over cosmic time, this galaxy cluster abundance measurement will be analyzed in the context of cosmological parameters, some of which will be measured, some confirmed, and some more precisely constrained than before in the literature.

1.1 Introductory Cosmology

Physics on large scales is well-described by the theory of general relativity (GR). Published in 1915 by Albert Einstein, GR is a geometric description of gravity, the dominating force on cosmological scales. Here we will briefly derive the theoretical framework that is required for the cosmological analysis using the galaxy cluster data obtained with the SPT. We will place particular importance on the assumptions that are made.

1.1.1 Derivation of the Spatially Flat FRW Metric

The fundamental assumption underlying modern cosmology is known as the Cosmological Principle. It states that the universe is isotropic and homogeneous on large scales. The universe is relatively smooth, and the tiny fluctuations we see can be described statistically in a way that doesn't have preferred directions or preferred locations. The reason that we do see structure around us today can be explained by those tiny seeds of structure growing over time due to gravity, and on relatively small length scales (such as that of a galaxy cluster or smaller) due to other fundamental forces. Restricting ourselves to gravitional effects only for the moment, GR can be compactly summarized with the Einstein equation

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \ . \tag{1.1}$$

On the left hand side we have the Einstein tensor $G_{\mu\nu}$, a rank 2 tensor that can be derived from the metric $g_{\mu\nu}$ which describes the curvature of spacetime. On the right hand side, we have the energy-momentum tensor. The Einstein equation implies the fact that locally, the curvature of spacetime is sourced by components of the stress-energy tensor: energy, momentum, pressure, and shear.

For the remainder of this introduction to physical cosmology, we will have to make several simplifying assumptions about both the spacetime metric, as well as the stress-energy tensor. We would like to stress that these assumptions are paramount to deriving the rest of our understanding of cosmology. They are a good fit to available data, and while many alternatives have been falsified through observations, they are in no way fundamental.

First, we will assume the simplifying functional form for the metric

$$\begin{aligned} g_{00} &= -c^2 \\ g_{ii} &= a(t)^2 \end{aligned} \tag{1.2}$$

Type of energy density	equation of state $\mathcal{W}$
radiation	$-1/3$
matter	0
cosmological constant	-1
constant dark energy	w
arbitrary dark energy	$w(a)$

Table 1.1 Commonly used components of the stress-energy tensor $T_{\mu\nu}$ and their equation of state.

with the other elements zero. This is simply the Minkowski metric with a time-dependent "scale factor" $a(t)$ in front of the spatial components. This is the spatially flat case of what is known as the Friedmann-Robertson-Walker (FRW) metric.

Secondly, we will parameterize the stress-energy tensor using terms of the form

$$\begin{aligned} T^{00} &= \rho \\ T^{ii} &= p = \mathcal{W}\rho c^2 \end{aligned} \tag{1.3}$$

again with the other elements zero. Here ρ describes the energy density of the component in question, p describes its pressure, which is directly related to density through the equation of state parameter $\mathcal{W}$. Table 1.1 shows some of the relevant components of the stress-energy tensor we will use later on.

In order to evaluate the Einstein equation, we expand the Einstein tensor as follows:

$$G_{\mu\nu} = (\delta^{\gamma}_{\mu}\delta^{\zeta}_{\nu} - \frac{1}{2}g_{\mu\nu}g^{\gamma\zeta})(\partial_{\epsilon}\Gamma^{\epsilon}_{\gamma\zeta} - \partial_{\zeta}\Gamma^{\epsilon}_{\gamma\epsilon} + \Gamma^{\epsilon}_{\epsilon\sigma}\Gamma^{\sigma}_{\gamma\zeta} - \Gamma^{\epsilon}_{\zeta\sigma}\Gamma^{\sigma}_{\epsilon\gamma}) \ . \tag{1.4}$$

The Christoffel symbols ($\Gamma_{\alpha\beta\gamma} = \frac{1}{2}\left(\partial_{\beta}g_{\alpha\gamma} + \partial_{\gamma}g_{\alpha\beta} - \partial_{\alpha}g_{\beta\gamma}\right)$) are

$$\begin{aligned} \Gamma^{0}_{ii} &= \tfrac{a\dot{a}}{c^2} \\ \Gamma^{i}_{0i} &= \tfrac{\dot{a}}{a} \end{aligned} \tag{1.5}$$

which gives the Einstein tensor. Writing out the Einstein equation in matrix form:

$$\begin{pmatrix} 3\left(\frac{\dot{a}}{a}\right)^2 & 0 & 0 & 0 \\ 0 & -\frac{1}{c^2}\left(2\ddot{a}a+\dot{a}^2\right) & 0 & 0 \\ 0 & 0 & -\frac{1}{c^2}\left(2\ddot{a}a+\dot{a}^2\right) & 0 \\ 0 & 0 & 0 & -\frac{1}{c^2}\left(2\ddot{a}a+\dot{a}^2\right) \end{pmatrix} = \frac{8\pi G}{c^4}\begin{pmatrix} c^4\sum_i \rho_i & 0 & 0 & 0 \\ 0 & c^2\sum_i \mathcal{W}_i\rho_i & 0 & 0 \\ 0 & 0 & c^2\sum_i \mathcal{W}_i\rho_i & 0 \\ 0 & 0 & 0 & c^2\sum_i \mathcal{W}_i\rho_i \end{pmatrix} . \tag{1.6}$$

These are called the Friedmann equations. They describe the evolution of the scale factor given a set of energy density contributions. Rewriting, the 0^{th} element states that

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\sum_i \rho_i \ , \tag{1.7}$$

and the trace states that

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\sum_i \rho_i\left(1+3\mathcal{W}_i\right) \ . \tag{1.8}$$

These equations, describing the background evolution of the scale factor, are more often seen using different parameters which will be described now.

1.1.2 Reformulating into Observational Jargon

Much of the purpose of this thesis is to make experimental contributions to

1. *testing the assumptions stated above.*
2. *determining the free parameters in the Friedmann equations.*
3. *measuring deviations from the Friedmann equations on smaller scales, where the cosmological principle is insufficient.*

In order to more easily compare the Friedmann equations to observable quantities, we will introduce several variables. We refer the reader to Appendix B for a summary of all commonly used variables in this dissertation.

First, the Hubble parameter H is defined as $H = \dot{a}/a$ and its value today (by convention, $a = 1$ at present) is denoted with H_0. Its numerical value has been measured to be on the order of 0.1 Gyr^{-1} or 100 km/s/Mpc. Secondly, we will reparameterize the density parameters ρ_i in terms of the critical density

$$\rho_c = \frac{3H^2}{8\pi G} \,, \tag{1.9}$$

by defining $\Omega_i = \rho_i/\rho_c$. Finally, we introduce redshift $z = 1/a - 1$.

Substituting this new parameterization reduces the Friedmann equations to

$$\boxed{\sum_i \Omega_i = 1} \,, \tag{1.10}$$

and

$$\boxed{\frac{H^2}{H_0^2} = \sum_i \Omega_{i,0}(1+z)^{-(1+3w_i)}} \,. \tag{1.11}$$

Assuming that the only relevant contributions to the stress-energy tensor are from matter, radiation, and a cosmological constant ($w = -1$) denoted by Λ, this equation becomes

$$H(z) = H_0\sqrt{\Omega_\gamma(1+z)^4 + \Omega_M(1+z)^3 + \Omega_\Lambda} \,. \tag{1.12}$$

This equation succinctly describes the expansion rate as a function of redshift.

1.1.3 Deviations from the FRW metric

As alluded to earlier, the FRW metric is meant to describe physics on the largest scales. On smaller scales, there is clearly structure in the matter density as a function of spatial position. This structure has largely grown gravitationally, starting as minute fluctuations in a highly homogenous, hot, early universe. Perhaps the most compelling

Symbol	Parameter description
A_s	Normalization of spectrum of primordial fluctuations
n_s	Power law slope of primordial fluctuations
Ω_c	Energy density of cold dark matter today
Ω_b	Energy density of baryonic matter today
θ_s	Characteristic angular scale of primary CMB fluctuations

Table 1.2 The parameters of the ΛCDM cosmological model.

evidence for this comes from the primary CMB fluctuations, where the maximum excursions are $\mathcal{O}(10^{-4})$ on top of a highly uniform background density.

The way structure forms on few-Mpc scales is remarkably well-understood. Modern cosmological measurements corroberate that this structure can be predicted from purely gravitational simulations using only five parameters, shown in Table 1.2.

This cosmological model with vaccuum energy Λ, cold dark matter, and no spatial curvature is called ΛCDM.

1.2 Measuring Cosmological Parameters with Cluster Counts

Galaxy clusters are the largest gravitationally collapsed structures in the universe. They are the tracers of extreme peaks in the matter density field on few Mpc scales and their abundance is a strong function of the aforementioned cosmological parameters. They consist of $\sim 88\%$ dark matter. The remainder of the mass is largely made up of a diffuse, $\sim 10^7$ K plasma, while stars only make up a few percent of the total mass.

The first mention of clusters of galaxies in the literature is in 1784, by Charles Messier. He found that galaxies cluster on the sky, though they were then called "nebulae", since their extragalactic origin was not known. Clusters have since been rich laboratories of physics on Mpc scales. Perhaps most notably, Fritz Zwicky used the virial theorem on the Coma cluster to infer the existence of dark matter in 1933, opening a new area of physics. Observations of the Bullet cluster (1E 0657-558) performed in the late 1990s show that most of the mass in a cluster-cluster merger is collisionless, providing almost indisputable evidence for dark matter.

Today, since dark matter is by far the dominant component in the dynamics of clusters, the non-linear gravitational process of cluster formation can be understood to great accuracy using numerical simulations. These so-called "N-body" simulations only need to include basic gravitational physics of Lagrangian test particles in an FRW metric.

The key observables predicted by such theory are the mass and redshift distribution of clusters. Therefore, in order to place constraints on cosmological parameters, one must accurately determine the mass and redshift of each of the detected clusters. Finding the cluster redshift is typically done using optical techniques, either through galaxy photometry, or the location of spectral lines using spectroscopy. X-ray spectroscopy, namely the identification of the 7 keV iron line, is another useful method for determining cluster redshift.

Mass determination will be the topic of much of this dissertation. Finding an accurate systematic mass scale for a given cluster sample has proven to be the dominant systematic error on cosmological parameter determination. The statistical techniques required to turn abundance measurements into cosmological constraints are non-trivial, primarily due to the fact that the theoretical abundance falls steeply with underlying cluster mass. In real data, this steep population function is inevitably sampled using a noisy proxy, leading to Eddington bias (see Chapter 5). A correct treatment of abundance-based constraints, as well as incorporating realistic uncertainties on the (critical) knowledge of the overall mass scale, will be presented in Chapter 6.

Since clusters are overdensities in the continuum matter density field, one must choose a definition of halo mass. Connecting measured mass estimates to theoretical predictions can be difficult when these definitions are not identical. In this work, we use M_{500}, the mass enclosed in a sphere that contains an average matter density five hundred times greater than the critical density of the universe.

1.2.1 The Sunyaev-Zel'Dovich Effect

The thermal Sunyaev-Zel'dovich (SZ) effect is the inverse Compton scattering of CMB photons off the hot electrons in the intracluster medium (ICM). Due to the fact that this is a scattering process, increasing the mean photon energy while preserving photon number, the SZ effect manifests as a spatial flux decrement below the null frequency of $\sim$ 220 GHz. This is shown in Figure 1.1.

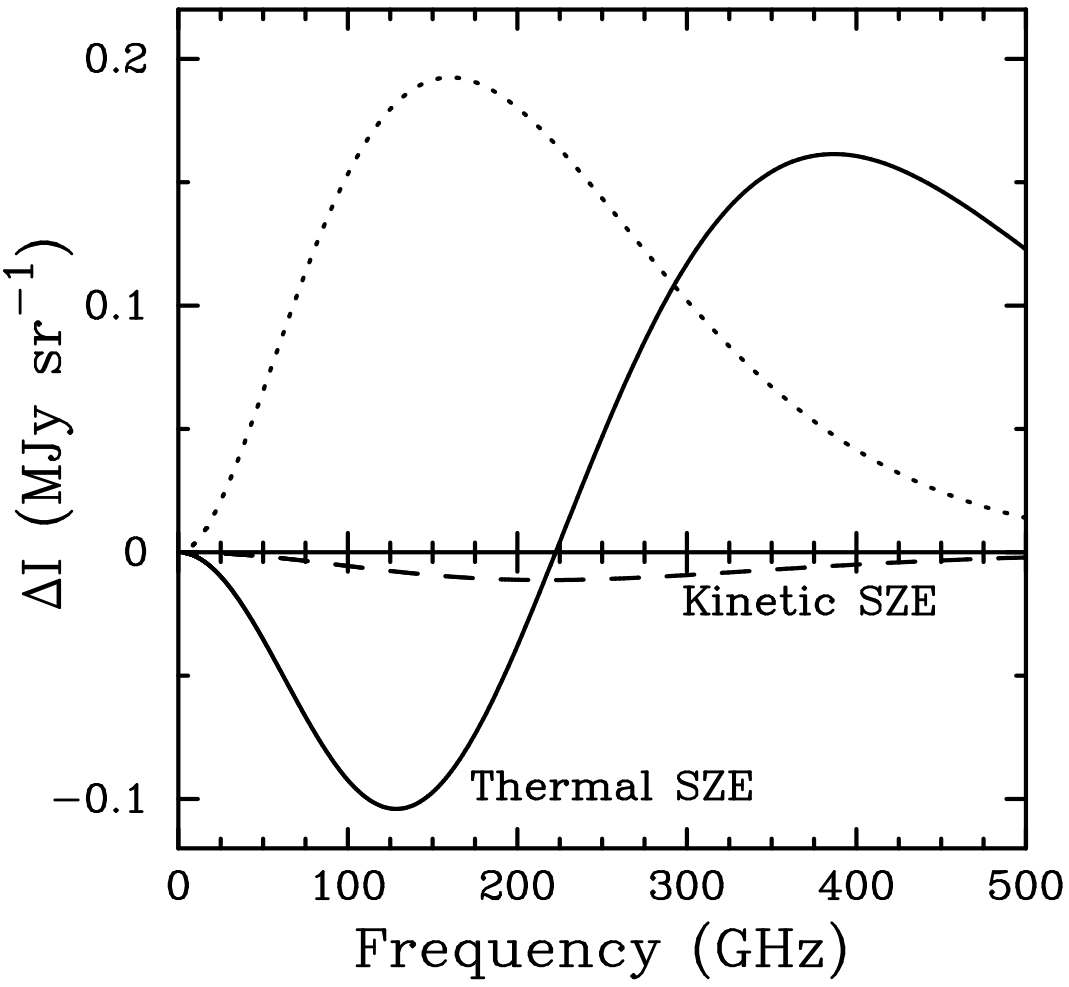

Figure 1.1 The spectral shape of the thermal SZ effect. Below roughly 220 GHz, the thermal SZ effect causes a characteristic flux decrement primary. The primary CMB and the kinetic SZ effect are as shown dotted and dashed lines, respectively. Figure adapted from Carlstrom et al. (2002).

The amplitude of the effect is proportional to the density of hot electrons present in the ICM, and proportional to their temperature

$$\frac{\Delta T}{T_{\rm CMB}} \propto \int dl n_e T_e \ , \tag{1.13}$$

where the integral runs along the line of sight. The thermal SZ effect is independent of redshift, allowing an SZ survey to select clusters out to high redshift, only limited by the time it takes for clusters to form.

Beyond the thermal SZ effect, there are related and complementary concepts such as relativistic corrections to the thermal SZ effect, the non-thermal SZ effect due to a possible ultra-relativistic non-thermal electron population, and the kinetic Sunyaev-Zel'dovich effect (kSZ) due to the proper motion of hot electrons relative to the CMB rest frame. For a detailed review of both the tSZ effect and these related concepts, see Carlstrom et al. (2002).

1.3 The South Pole Telescope

The SPT is a 10-meter diameter telescope located at the geographic South Pole. With a mirror surface accuracy of better than 20 μm RMS, it is capable of performing in the sub-mm waveband, though the first three generations of cameras observe in the millimetre-wave band. The 10-m diameter and millimeter waveband result in a ~ 1 arcminute beam, which was chosen for the primary science goal: the discovery and abundance measurement of hundreds of massive galaxy clusters (Ruhl et al., 2004), which have a characteristic scale of one arcminute at redshifts beyond $z = 0.25$. For this purpose, the three-band (see Figure 1.3) SPT-SZ camera was installed and saw first light in 2007, surveying 2500 square degrees of southern sky during five years. Roughly simultaneously, similar surveys were performed with the ground-based ACT experiment and the *Planck* satellite.

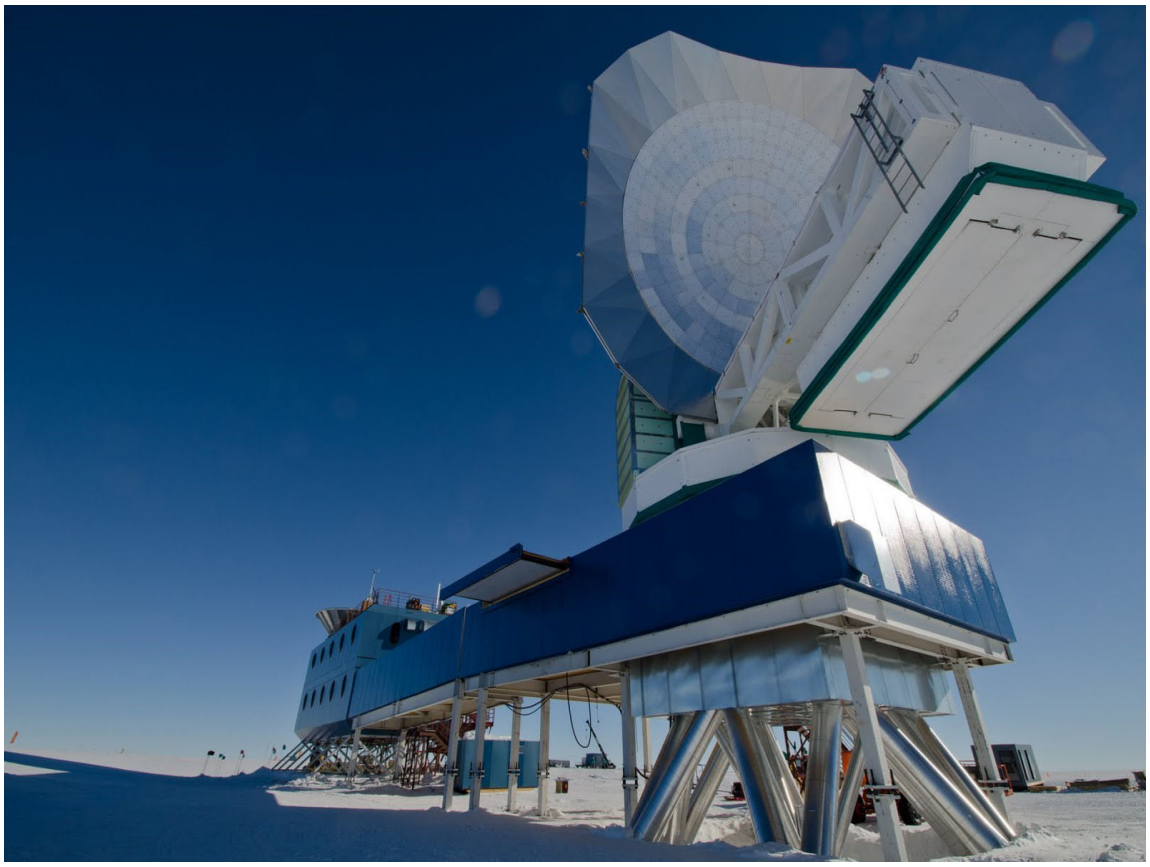

Figure 1.2 The 10-m South Pole Telescope in the austral summer of 2012-2013. Photograph by Ryan Keisler.

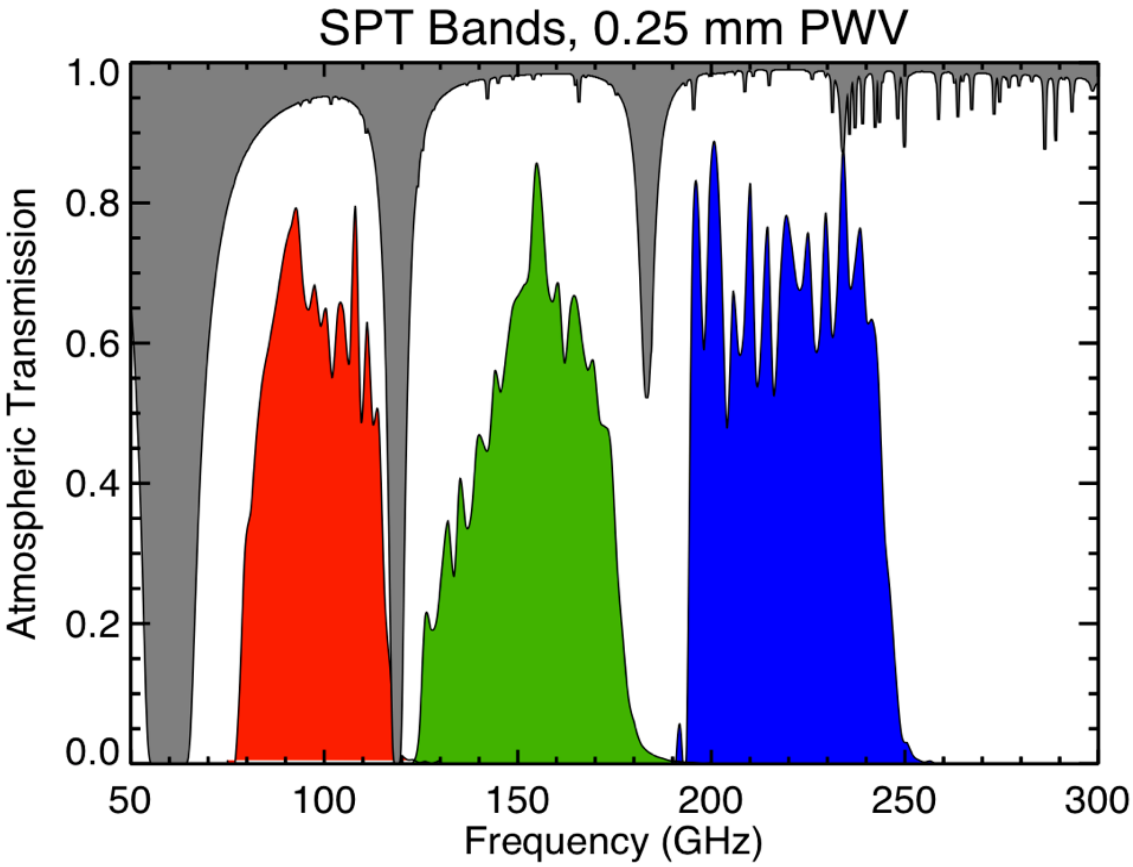

Figure 1.3 The three spectral bandpasses used in the SPT-SZ experiment. The red, green and blue bandpasses are denoted as the 90 GHz, 150 GHz and 220 GHz channels in this dissertation. Overplotted is the expected atmospheric transmission at the south pole. Note that the major telluric lines are avoided, allowing for a low-loading, low-opacity measurement of the background astrophysical sky. Figure by Brad Benson.

1.3.1 SPTpol and SPT-3G

In November of 2011, we installed the next-generation SPT camera: SPTpol. As its name suggests, the camera is polarization-sensitive. The focal plane is populated with seven 150 GHz detector modules and 180 dual-polarization 90 GHz detectors. Its primary science goal is to measure the odd-parity patterns (B-modes) in the polarization of the CMB, which are sourced by both gravitational lensing due to large scale structure, and primordial gravitational waves. The galaxy clusters discovered through the SZ effect also provide a wealth of cosmologically and astrophysically interesting information, though analysis of this sample is beyond the scope of this work. We will, however, discuss the advances in detector readout technology, which are implemented on SPTpol, in §2. These technological advances will be critical for the development of SPT-3G, the third generation camera planned to be installed on the SPT in 2016.

The remainder of this dissertation will describe the SPT experiment in greater detail, including contributions to readout electronics, low-level data analysis, and an optimal cluster finding algorithm. A new Bayesian method for computationally efficiently computing the cosmological likelihood of the SPT sample will be presented, followed by the cosmological constraints inferred from this measurement.

2

Frequency Multiplexed Readout

The SPT is one of a large number of mm-wave astrophysics experiments (Schwan et al., 2011; Reichborn-Kjennerud et al., 2010; The Polarbear Collaboration et al., 2010) that employ frequency multiplexed, voltage biased transition edge sensor (TES) bolometers to convert the incident radiation power into a current, which is measured by superconducting quantum interference devices (SQUIDs). TES bolometers are exquisitely sensitive power measurement devices. To achieve such high sensitivities, they are kept at cryogenic temperatures, limiting the number of wires that can be used for their readout. For that reason, multiplexing is necessary in order to achieve large pixel-count cameras.

The SPT-SZ instrument, with which the cluster data presented in this work was obtained uses an analog implementation of frequency multiplexed TES bolometer readout. This system is presented in Dobbs et al. (2011). The current SPTpol camera uses a digital implementation known as the McGill digital frequency multiplexing (DfMux) platform (Dobbs et al., 2008). This modern version is more compact, lower in cost and power consumption, improves the low-frequency noise performance of the experiment, and allows for a multiplexing factor greater than the analog system (which ran at 7x). In addition, it easily allows for extensions to the readout scheme, which is the topic of this chapter. A simplified schematic of this readout scheme is shown in Figure 2.1.

Due to the steep total power-resistance relation of the TES, a voltage bias results in strong electrothermal feedback (ETF) (Irwin, 1995). In essence, the sum of optical

and electrical power deposited on the bolometer is kept constant at the level dictated by the ETF loop gain. Series impedances to the TES spoil ETF and decrease stability, as we further explore in §2.2.2. The main sources of series impedance are the input impedance of the SQUID and the inductance of the cryogenic wiring.

In addition, the SQUID has a highly non-linear response and has limited dynamic range. Providing negative feedback to the SQUID linearizes it and suppresses its effective input impedance. This negative feedback was previously provided broadband using a SQUID flux locked loop shunt feedback (Dobbs et al., 2011, 2012) (hereafter referred to as shunt feedback). It is shown as the dashed part of Figure 2.1. However, the shunt feedback has several issues. First, it needs to provide strong negative feedback over the entire bandwidth, while maintaining stability. Managing phase shifts and rolling off the open loop gain to maintain the stability of this loop is a challenge. Due to details not relevant to this work (see e.g. Lueker (2011)), this requirement currently limits the usable bandwidth for frequency multiplexing to $\sim$ 1.3 MHz and restricts wiring lengths. In addition, the presence of a superconducting leg in the LCR comb (i.e. a latched TES) can cause the shunt feedback to go unstable (Lueker, 2011).

Here we describe an alternative feedback scheme known as Digital Active Nulling (DAN), which does not suffer from these limitations, allowing a higher multiplexing factor, reduction of device parameter requirements, and improved stability. DAN is therefore a key technology for enabling multi-kilopixel bolometer arrays, for future ground-based, balloon-borne, and satellite applications. SPTpol is using this system for part of the focal plane as well as for significant speed gains during the setup and tuning of the SQUIDs and bolometer array. The previous mode of operation that did not use DAN will be referred to as "classic" operation.

DAN is similar to baseband feedback (den Hartog et al., 2011, 2012; Takei et al., 2009) (BBFB), except for the electrical injection of the nulling signal, whereas BBFB is applied directly to the SQUID through a separate feedback coil. If the coupling

constant (defined as the ratio of mutual inductance to the geometric mean of self inductances) of the input coil to the feedback coil were unity, this would be equivalent to DAN as far as the feedback loop is concerned. However, BBFB applications have so far had a small mutual inductance of the feedback coil to the input coil, in which case the increased linearity and dynamic range are preserved, but the SQUID input impedance suppression is equal to the coupling constant.

Furthermore, in §2.2 we present a novel method of suppressing the effect of impedance in series with the TES known as Digitally Enhanced Voltage Bias (DEVB). Here, we once again take advantage of the small bandwidth of the TES and digitally correct for the voltage drop across the stray impedance; effectively providing strong voltage bias at each TES bias frequency separately.

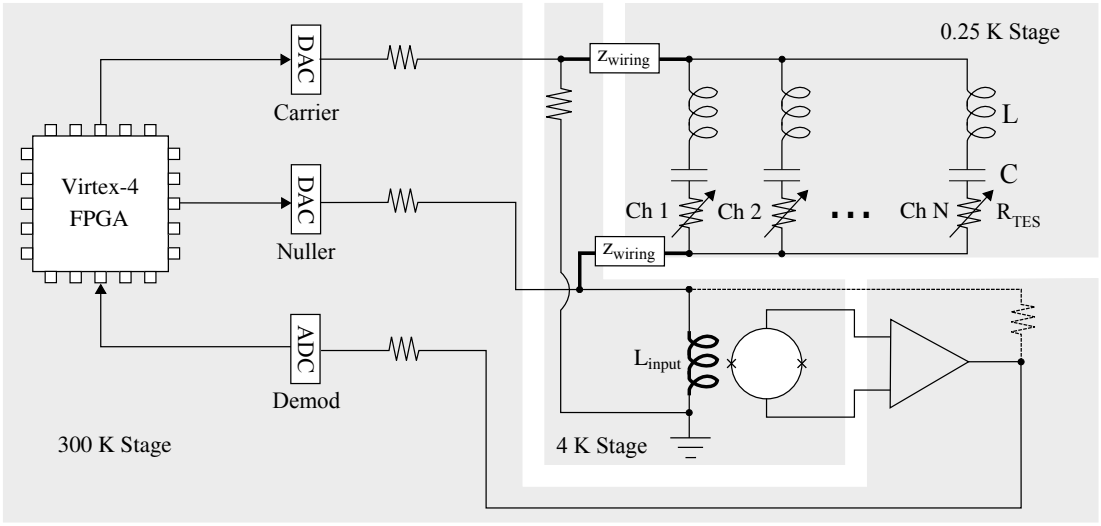

Figure 2.1 Simplified schematic showing the voltage biased, frequency multiplexed bolometer readout system. The digitized carrier waveform provides strong voltage bias to each TES at its respective resonant frequency, as set by the series LC filter. The SQUID provides a current measurement, which is amplified, digitized and demodulated. In classic operation, the digitized nuller signals provided static sinusoidal feedback, in addition to broadband feedback from the shunt feedback. The latter, shown dashed, is omitted in DAN operation, where the nuller signal actively zeroes the SQUID current across the bolometer bandwidth. The thick lines highlight the main sources of series impedance to the bolometer.

Series impedances spoil the voltage bias. In the extreme limit where the series impedance dominates the total impedance, the TES is effectively current biased and is unstable to perturbations. An increase in current causes an increase in electrical power and hence in resistance, leading to an instability. As shown in Appendix A, for an effective series complex impedance z_s and TES resistance R_{TES}, the stability criterion is given by

$$\frac{d\log V_{\mathrm{TES}}}{d\log R_{\mathrm{TES}}} \leq \frac{|z_s|}{|z_s + R_{\mathrm{TES}}|}\,. \tag{2.1}$$

Once the logarithmic slope of the transition exceeds the fractional effective series impedance, the TES will latch into a superconducting state, such that it is no longer useful as a detector. In §2.2.2, we show that DEVB allows for stable operation in this regime.

2.1 Digital Active Nulling

We have implemented DAN on the DfMux platform, with the flowchart shown in Figure 2.2. Each circuit board provides 8 digital-to-analog converters, 4 analog-to-digital converters, and a Xilinx Virtex-4 Field-Programmable Gate Array (FPGA) for signal processing tasks. Each digital-to-analog converter synthesizes a "comb" waveform at 25 MSPS, consisting of 16 sinusoidal terms with independent amplitude, phase, and frequency controls. The DfMux also provides 64 complex demodulator channels, and streams data from each of them over Ethernet at a programmable sampling rate which, for this work, we set to 191 Hz. The software stack consists of an embedded Linux system, running C code and controlled remotely with a user-friendly python API (Smecher et al., 2012; Story et al., 2012b).

DAN is implemented as a separate discrete time integral control loop at each bias frequency. For one such bias frequency, the feedback loop is shown in Figure 2.2. We choose integral control for its desirable property of having an effective loop gain proportional to $1/\delta f$, where δf denotes the frequency separation of the signal from the bias frequency. This implies that it has infinite gain at $\delta f = 0$, where the bulk of

the signal (the bias voltage) resides. For stability, we require that the open loop gain falls below unity at 180° of phase shift, such that no poles of the closed loop transfer function have a positive real component. The phase shift is dominated by digital time delays, which, in our implementation, range from 5.5 μs - 10.6 μs depending on the bias channel. Thus, we can easily achieve stability by letting the open loop gain fall below unity at $\sim$ 10 kHz, which implies that at the highest bandwidth that may be required for science, say 100 Hz, we still have a loop gain of 100 and hence 99% effective nulling.

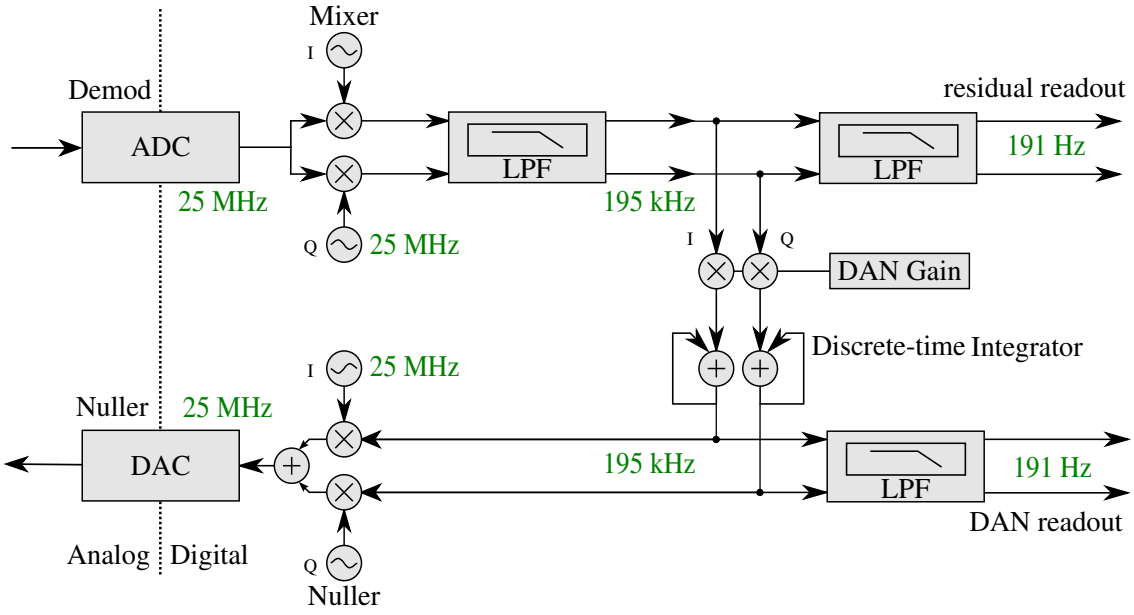

Figure 2.2 Schematic diagram showing the firmware implementation of the DAN algorithm signal path. Data rates are shown in green. Both readout paths are shown. The residual readout path can be used to characterize the properties of the loop, though as long as the loop is operating correctly, the DAN readout alone respresents the bolometer current.

2.1.1 Experimental Results

We initially characterized our implementation of DAN by bypassing the cryogenic components of the readout system and using a warm resistive network that has a similar transfer function to that of the cryogenic components. A measurement of the

residual signal at the current sensor as a function of frequency (hereafter referred to as a network analysis) of DAN in this characterization setup reveals the closed loop transfer function shown in Figure 2.3.

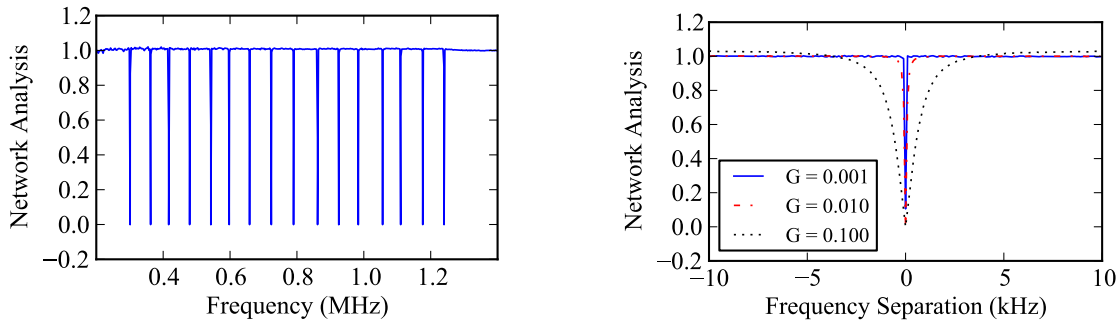

Figure 2.3 Left: The broadband closed loop response of DAN. At 16 arbitrarily chosen bias frequencies where DAN is operating, the transfer function of an injected signal dips to zero i.e. the signal is nulled. Right: Zoom on the network analysis near one bias frequency for a variety of gain settings. The effective bandwidth of the nulling loop is shown to increase with G, a commanded gain parameter proportional to the open loop gain of the DAN feedback loop.

We then operated DAN with cryogenic bolometers, including the SQUID readout. In contrast to classic operation, we found that one advantage of the DAN mode of operation is that the time required to map out the TES transition is reduced to the trivial exercise of stepping in bias amplitude while recording the DAN output. This mapping is performed after each cryogenic cycle on the SPTpol experiment. Using DAN decreases this tuning time (which previously took tens of minutes) by an order of magnitude, while simultaneously measuring the voltage-resistance characteristic at a much higher resolution. This speedup, simplification and improvement of the bolometer tuning algorithm has been implemented and is currently one of the uses of DAN occurring daily on SPTpol, and was used for the first science flight of the EBEX balloon-borne experiment (Reichborn-Kjennerud et al., 2010).

In the cryogenic test setup, we verified that the loop is operating correctly by streaming the residual demodulator signal and seeing a highly suppressed white noise level that increases with frequency in the manner expected from the integral control

loop. We also verified that the DAN signal is the same as the demodulator signal when the gain parameter is set to zero. In particular, we provide a small sinusoidal test signal, which applies time varying power to the TES. The resulting demodulated current measurement(Lueker et al., 2009) is found to be identical between the cases where DAN is enabled and where it is disabled. In addition, we compare the white noise level between these two modes of operation and find that they agree. A histogram of the white noise levels for a cryogenic SQUID without TES bolometers attached is shown in Figure 2.4.

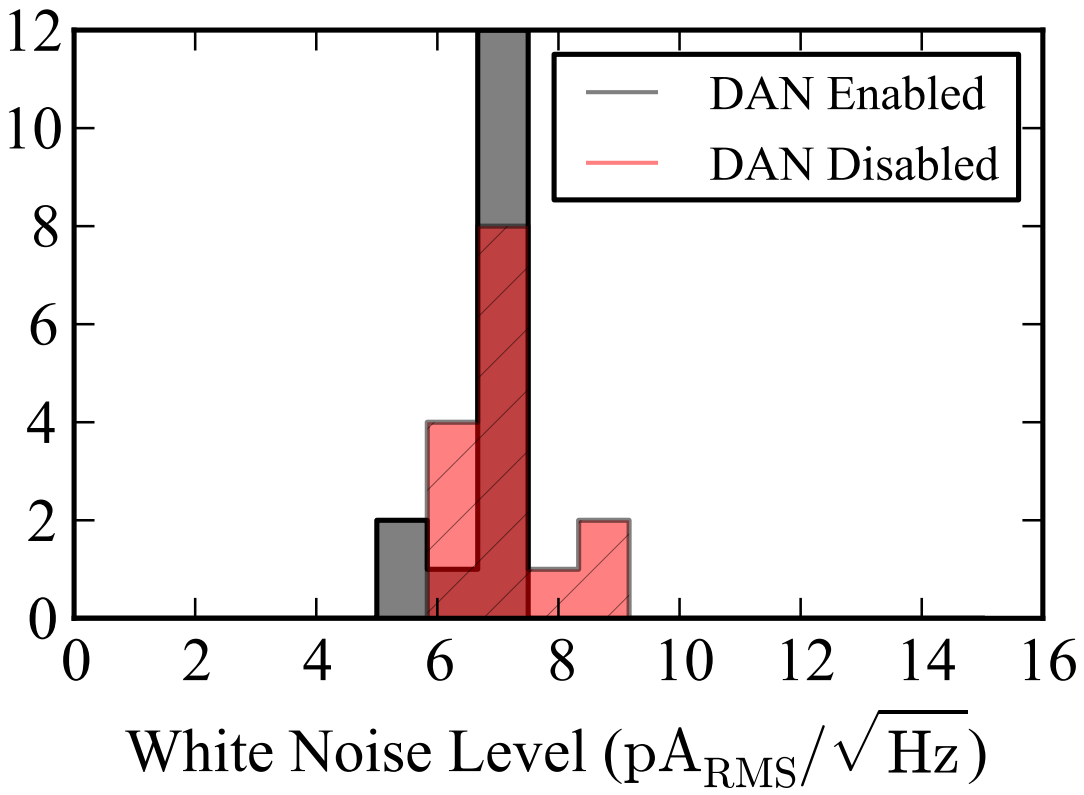

Figure 2.4 Histogram comparing the white noise level for 15 DAN channels operating on a SQUID, with DAN enabled and DAN disabled. The two distributions appear statistically consistent, confirming our expectation that the DAN operation does not alter the noise properties of the readout system appreciably.

2.1.2 Improvements due to DAN

Operating the system in DAN mode has several advantages. First, the strongly reduced effective SQUID input impedance results in a stiffer voltage bias, improving TES stability.

DAN nulls the SQUID current near bias frequencies, improving SQUID linearity and stability due to the large reduction of the total signal amplitude at the SQUID input. Similarly, improved nulling reduces the dynamic range requirement on the ADC, avoiding overloads in the case of unstable bolometers or changes in the bolometer operating point.

Removal of the shunt feedback improves system stability, allowing longer wire lengths between the 4K SQUID and room temperature electronics, and increasing the allowed bandwidth. Additionally, the shunt feedback instability to a single superconducting leg in the comb is eliminated. DAN does not interact across bias frequencies, so any instabilities are localized to one TES, instead of the whole comb. We have experimentally verified this by allowing various numbers of individual bolometers to go superconducting and found no measurable effect on the noise level of the remaining bolometers on the same comb. This is in stark contrast to classic operation.

The increase in usable bandwidth, removal of sensitivity to single misbehaving bolometers, and reduced usage of SQUID dynamic range, all enable a higher multiplexing factor. This higher multiplexing factor and the longer allowed wiring length are key for reducing heat load on cold stages such that it falls within the capabilities of available space cryogenics, as well as enabling lower cost, higher pixel count ground based experiments.

Both increasing wiring lengths and moving the bolometer channels to higher frequency causes an increase in the wiring inductance (bolded in Figure 2.1), which is a source of series impedance to the TES. We note that the presence of this inductance also shifts the optimal bias frequency and that, at this new bias frequency, the resulting effective impedance in series with the TES is smaller than $j\omega L_{\text{wiring}}$.

The amplitude and phase of the effective complex series impedance are non-trivially related to the details of the LCR network, though an analytic calculation is possible.

The series impedance due to wiring can be mitigated by moving this wiring inside the DAN feedback loop, suppressing it in a manner similar to the SQUID input impedance suppression. This involves increasing the number of wires from each SQUID to a cryogenic stage near the detectors from two to six. An alternate scheme to mitigate this wiring inductance is DEVB, which we discuss in §2.2.

2.2 Digitally Enhanced Voltage Bias

While the SQUID input impedance is strongly suppressed by DAN, residual series impedance (e.g. from wiring inductance, capacitor ESR, or magnetic coupling to lossy material) can limit TES stability. This series complex impedance can be measured to good accuracy by tuning the bolometers into their transition, cooling them below their superconducting temperature, and stepping down the bias voltage until the TES latches into its superconducting state. At this point, the current is proportional to the applied voltage (as we have experimentally verified), giving a measure of the residual complex impedance, which is typically ∼10-30% of the bolometer normal resistance. We will show how this impedance can be used to dynamically adjust the voltage bias such as to keep the TES voltage constant.

TES bolometers respond to incident power with a time constant $\tau_0 = C/G$ where C is the heat capacity of the TES and G is the thermal conductance to the heat sink. Under voltage bias, this time constant is sped up by the loop gain of electrothermal feedback $\mathcal{L}$ (e.g. Richards (1994)) as

$$\tau_{\rm eff} = \frac{\tau_0}{1+\mathcal{L}}\,. \tag{2.2}$$

DEVB actively controls the voltage across the TES on time scales faster than $\tau_{\rm eff}$, providing a constant voltage bias across the TES while strongly suppressing the effective series complex impedance.

We denote this series impedance as $z_s = R_s + jX_s$. Given values of R_s and X_s, a constant voltage bias to the TES denoted V_{TES} can be provided by measuring the current through the bolometer and providing the voltage

$$V_{\mathrm{OUT}} = \sqrt{(V_{\mathrm{TES}} + R_s I)^2 + X_s^2 I^2}\,, \tag{2.3}$$

where I denotes the magnitude of the measured current and V_{TES} is the programmable voltage set-point. Since, as with DAN, there are digital delays between measuring I and providing V_{OUT}, we perform this computation at a fixed loop gain below unity in order to introduce an effective DEVB time constant and ensure stability of the feedback loop. The DEVB time constant is discussed further in §2.2.1.

In addition to providing improved TES stability, the enhanced voltage bias across the LCR comb suppresses cross-talk due to that same series impedance. In the case of an inductive series impedance, this term is proportional to the ratio of the stray inductance to the inductance of the LC resonator(Dobbs et al., 2011). With DEVB, this effect is suppressed by the DEVB loop gain.

2.2.1 Preliminary Implementation

A hardware prototype of DEVB has been implemented based on a customized version of the DfMux firmware(Smecher et al., 2010), outlined in Figure 2.5. This firmware supports 64 bias and nuller channels, as well as 68 single phase demodulator channels. The DEVB feedback loop is implemented on one of those bias channels, using two of the demodulator channels, aligned 90° out-of-phase to provide I and Q current measurements at the bias frequency. The resulting signals are fed to a DEVB module, which calculates an output amplitude for the synthesizer block by explicitly computing Equation 2.3, using the quadrature sum of the I and Q measurements as the current amplitude. The synthesizer modulates the amplitude of the bias sinusoid using this output. The input to, and the output from, the DEVB module are further decimated to 191 Hz and streamed across the network for setup and analysis.

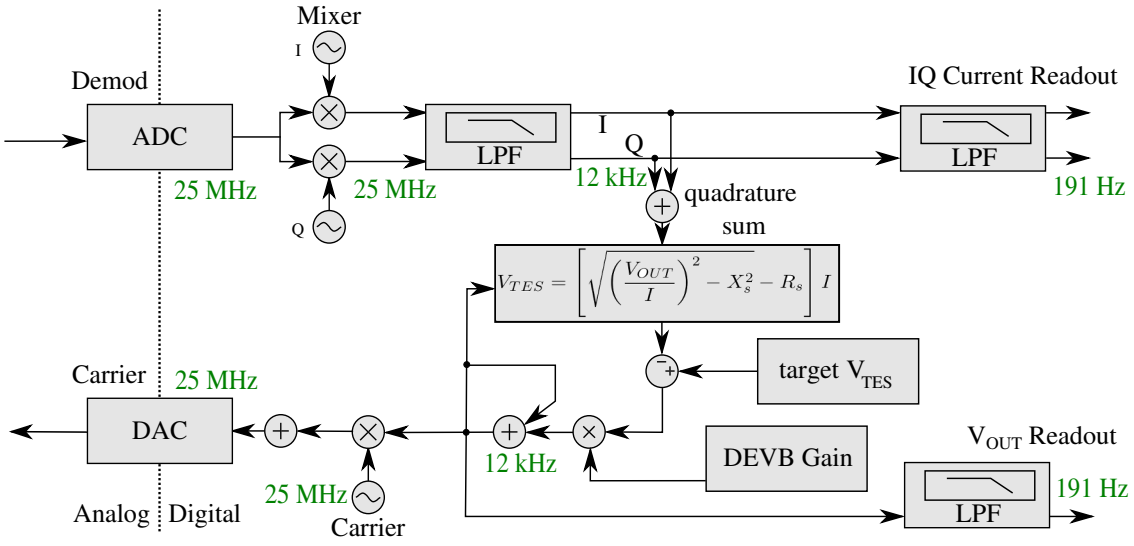

Figure 2.5 Firmware diagram showing the DEVB signal path. Data rates are shown in green. The complex current measurement and output voltage amplitude are read out simultaneously.

Figure 2.5 shows the computation within the DEVB module. At the input to this module, fixed-point signals from the demodulator are converted into floating-point numbers for computation. The algorithm itself is separated into two parts: calculation of a target amplitude, and a control element generating the feedback signal by stepping the current amplitude a fraction towards the target amplitude. We denote this fraction the DEVB gain. The control signal generated by these blocks is finally converted back into fixed-point numbers suitable for the synthesizer block.

We run the control loop at 12 kHz and set the DEVB gain to 1/3, yielding an approximate DEVB time constant of 4 kHz; much faster than the $\sim$ 100 Hz time constant of a typical TES.

2.2.2 Measurements

In order to demonstrate the enhanced performance of TES detectors under DEVB, we test the stability of a detector in the presence of a large stray impedance as given in Equation 2.1. Specifically, we apply a large carrier voltage to a single detector

in an LC resonance, designed to allow for frequency multiplexing. By setting the frequency of this voltage significantly away from the LC resonance we create an effective additional impedance in series with the bolometer. In this case we set the series impedance to 1.4 j Ω, while the normal TES resistance is $\sim$ 0.9 Ω. We then reduce the applied voltage, dropping the TES into its transition. The TES is lowered through its transition until it latches and becomes superconducting. The resulting total impedance as a function of applied voltage and the resistance of the TES alone as a function of the voltage across the TES alone are shown in Figure 2.6. This test, performed with DEVB enabled and DEVB disabled, shows that the use of DEVB lowers the point at which the TES latches, demonstrating that DEVB allows the TES to drop much lower into its transition in the presence of large series impedance. Enabling this mode of operation relaxes design constraints on many elements of the readout system.

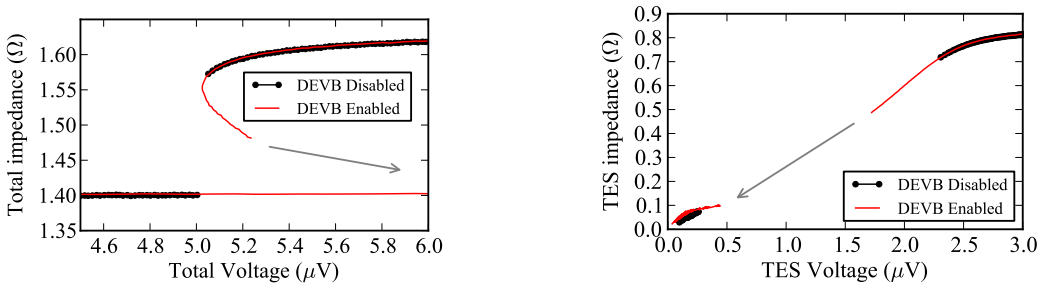

Figure 2.6 TES transition shown as the provided voltage is slowly decreased with DEVB enabled and disabled. The TES latches into a superconducting state when it reaches 60% of its normal resistance with DEVB enabled, compared to latching at 90% of its normal resistance with DEVB disabled. The left panel shows the raw data, whereas the right panel shows the data once the stray series impedance is subtracted off. The arrows are used to show the latching event, which occurs faster than the data rate.

2.2.3 Future Prospects for DEVB

While we have demonstrated the concept of DEVB, extending this feedback mechanism to an array of bolometers requires several improvements to the firmware, as well

as the readout scheme and feedback mechanism.

Equation 2.3, as well as the computation of the magnitude of the complex demodulated current, contains a square root operation, which is challenging to implement efficiently on an FPGA. An alternate DEVB scheme would involve controlling both the I and Q components of the bias voltage V_{OUT} rather than controlling only its amplitude. The set point for V_{OUT} would then become

$$\begin{aligned} V_{\mathrm{OUT}}^{I} &= V_{\mathrm{TES}}^{I} + I^{I} R_s - I^{Q} X_s \\ V_{\mathrm{OUT}}^{Q} &= V_{\mathrm{TES}}^{Q} + I^{Q} R_s + I^{I} X_s \,. \end{aligned} \tag{2.4}$$

This complex DEVB scheme, when implemented in fixed-point logic, would be easily scalable to the full 64 bias and complex demodulation channels using the current Xilinx Virtex-4 FPGA.

The current implementation of DEVB does not support DAN. In future versions of DEVB, we would implement both the DEVB and DAN modules simultaneously. We would choose to run DAN at a significantly faster effective time constant than the DEVB control loop, such that the SQUID current is nulled, regardless of the changes in V_{OUT} made by the DEVB module. This approach would allow us to employ the benefits of DAN, such as the higher usable bandwidth and reduced SQUID dynamic range requirement, while retaining the suppressed series impedance from wiring inductance without needing to increase the heat load from the 4K stage to the millikelvin stage by increasing the wire count.

Finally, it is possible to implement DEVB with frequency-dependent coefficients such that a frequency-dependent series impedance could be suppressed. This would allow narrower LC filters to be used, reducing channel spacing and hence increasing the multiplexing factor for a given usable bandwidth and constraint on electrical cross-talk.

We have introduced DAN, a method of keeping the current through the sensing portion of the readout system very close to zero at frequencies near the TES bias frequencies. This digital feedback mechanism increases stability and linearity of the SQUIDs, while suppressing their effective input impedance. We have demonstrated a fully implemented version of DAN including stability results and nominal noise performance, and described its use on the SPTpol experiment.

While this technique is similar to BBFB, the suppression of series impedance is unique to DAN. Removing the shunt feedback from the readout system increases the usable bandwidth for frequency multiplexing, enabling higher multiplexing factors and reducing wiring length constraints. Hence, DAN is a key technology for applying frequency multiplexed readout to satellite applications, which generally require long wiring lengths.

However, the use of higher bias frequencies increases the series impedance due to wiring. This effect can be mitigated by increasing the number of wires per multiplexing module from the 4 K SQUID to (close to) the millikelvin stage from two to six.

We also described DEVB, a digital feedback technique which suppresses series impedances to the TES by measuring them and providing the actively controlled bias voltage required to hold the TES at a constant potential. This technique has been demonstrated to increase TES stability, despite large ($z_s \gtrsim R_{\mathrm{TES}}$) complex series impedances, without increasing wire count.

3

SPT Data Analysis

3.1 Observations

The cosmological constraints presented in this dissertation are based on observations performed with the SPT from 2007-2011. With this five-year survey, we mapped the region $20h < \mathrm{RA} < 7h$, $-65° \leq \mathrm{dec} \leq -40°$, totalling 2500 square degrees of the southern sky. The depths to which this area was mapped are shown in Table 3.1. The full definition of map depth can be found in Schaffer et al. (2011).

Here, we consider only the 95 GHz and 150 GHz data. The 220 GHz data is significantly less deep and the observations were contaminated by atmospheric fluctuations at large angular scales where they could be useful for removing CMB fluctuations. The addition of this frequency band does not significantly improve the efficiency of cluster detections. For robustness, we therefore choose to leave it out of the analysis presented here.

The 2500 square-degree survey was performed in sections that are rectangular in right ascension and declination. These subregions are referred to as *fields*. The fields are denoted by their approximate centers e.g. `ra23h30dec-55`. Each of the 19 fields

Frequency (GHz)	Noise Level (μK-arcmin)
95	40
150	18
220	70

Table 3.1 The noise level as defined in Schaffer et al. (2011) for the 2500 square degree SPT-SZ survey in each of the three frequency bands.

was scanned in the following manner. The telescope is swept in azimuth at a constant velocity ($\sim 0.25^\circ/s$) from one end of the field to the other, and back. This is followed by a step in telescope elevation, and repeated until the field is covered vertically. This scan strategy allows the entire SPT focal plane to cover the field and is referred to as an *observation*. Observations are repeated for several months until the approximate depth of 18 μK at 150 GHz is reached.

The scan strategy differs for several fields. For some fields, the azimuth scans cover only one half of the field in right ascension at any one time. When half the field has been scanned, the other half is immediately started, with the duration of each half being exactly half the right-ascension extent of the field measured in hours. Due to the fact that the SPT is at a latitude of -90°, this results in the two halves of the field tracing out an identical scan pattern with respect to the ground. The observations using this approach are called "lead-trail" observations. For one other field (`ra21hdec-60`), the telescope azimuth drive was under investigation, so approximately 75% of the data was taken by moving the telescope in elevation from the highest declination to the lowest declination, only changing telescope azimuth in between observations.

3.2 Data processing and mapmaking

The time-ordered data from each detector was filtered in several steps, enumerated below.

1. The bolometers have intrinsic optical time constants that differ between detectors. In order to homogenize the frequency response between detectors, we divide the Fourier-space representation of the timestream of a particular detector by its best-fit single-pole filter. In order to avoid aliasing when binning into maps, we also apply a Fourier-space filter that is a low-pass filter of the form

$$H(\omega) = \begin{cases} 1 & \omega < \omega_0 \\ e^{-(\omega/\omega_0)^6} & \omega \geq \omega_0 \end{cases}, \tag{3.1}$$

where ω_0 is chosen to correspond to 25 Hz for all detectors.

2. The very lowest frequency components of the bolometer timestreams are contaminated by instrumental factors such as cryogenic stage temperature drifts, readout noise, and atmospheric fluctuation power. Since these sources of signal overlap with a region of the sky power spectrum dominated by primary CMB fluctuations, rather than cluster signal, we can remove these effects by applying a high-pass filter to the data. When applying a high-pass filter, bright point sources must be masked to minimize their effect on the rest of the map. The algorithm we developed for doing so uses linear least-squares to fit and subtracts a series of basis functions as follows. Briefly, the data is organized as a matrix $\mathbf{X}$ containing the bolometer timestreams, i.e. it is $(N_{tod} \times N_{bolometers})$. This is then multiplied by unity, minus the linear least squares operator as follows

$$\mathbf{X} = \mathbf{X} - \mathbf{V}\left(\mathbf{V}^\top \mathbf{V}\right)^{-1} \mathbf{V}^\top \mathbf{X} \,, \tag{3.2}$$

where $\mathbf{V}$ is the Vandermonde matrix in the case of polynomial filtering. In our case, we keep the first two polynomials (mean and slope), but replace the other rows of the Vandermonde matrix by sine and cosine functions, approximating a Fourier-domain hard high pass filter. While less computationally efficient than an FFT-based method, this explicit linear least-squares method conveniently allows for masking of the brightest mm-wave sources, which otherwise would cause filtering artifacts in the map domain. We first go through a preliminary iteration of this data analysis procedure without masking, in order to determine the locations of these brightest sources. We then use the measured positions in the second and final iteration of the map-making procedure as the mask locations.

3. Similar to the temporal high-pass filter, we apply a filter at each time sample but across all bolometers. This step is performed identically to the high-pass filter

applied to the time ordered data for each bolometer, except that $\mathbf{X}$ is replaced by its transpose and we use a Vandermonde matrix where each row contains a constant for bolometers in a given detector module, and zero otherwise. This results in a (masked) subtraction of the average signal across each detector module.

4. The time-ordered data is calibrated according to the procedure outlined in §3.2.1.

5. The filtered and calibrated time-ordered data is combined with the reconstructed pointing and binned into a map by simple inverse-variance-weighted binning and averaging.

We performed this process on simulated timestreams, and compared the resulting maps to an approximate Fourier-domain filter which will be described in §3.3. The resulting maps were tested by performing a galaxy cluster search, and the resulting cluster significances differed by $\sim 1\%$. This implies that the filtering procedure described above can be well-approximated by a Fourier-domain filter. Specifically, one with an isotropic 0.5° high-pass filter and a 1° high-pass filter in the scan direction.

Based on comparisons of radio source positions derived from SPT maps and positions of those sources in the AT20G catalog (Murphy et al., 2010), we find a small overall astrometry correction on the order of $\sim 5''$ on the sky. Since this is small compared to the angular extent of the galaxy clusters, we do not apply this correction in this work.

3.2.1 Calibration

The relative and absolute calibration of the detector response (i.e. the size of the signal measured for a given fluctuation on the sky) is computed using a tiered approach.

There are several issues to keep in mind when designing, implementing or evaluating a calibration strategy. The calibration of a bolometer may drift with time, and is

expected to change with weather and especially after a cryogenic cycle. The different bolometers also have a different calibration relative to each other.

A modulated thermal source is installed behind the secondary mirror, with a shutter in place. This source illuminates the focal plane with a 6 Hz square wave. This system would be sufficient to calibrate the focal plane if it weren't for the following issues. First, its illumination of the focal plane is non-uniform. Secondly, the atmospheric opacity can change over time, altering the relation between power received by the bolometer and the true power on the sky. Thirdly, the source is known to slowly dim over time, eventually burning out and needing to be replaced. Lastly, the overall absolute intensity of the source is not known to sufficient accuracy.

In order to address the non-uniform illumination, opacity and dimming over time, we use another calibration: frequent observations of the galactic HII source RCW38. The emission from RCW38 is sufficiently bright in all three spectral bands to yield a signal-to-noise ratio (SNR) of several for most bolometers in a $\sim$ 15 minute scan, while not being too bright to cause non-linear response of the detectors. The RCW38 response is averaged over an entire observing season to yield a relative calibration for each detector, which is used as a correction for the non-uniform illumination of the calibrator. The atmospheric opacity correction is computed by averaging the daily RCW38 response of each focal plane module. Finally, note that the opacity correction also automatically corrects for slow changes in the temperature of the modulated thermal source.

The absolute flux of RCW38 is known at the $\sim$ 10% level (Coble et al., 2003). Analysis of *Planck* satellite data (Planck Collaboration et al., 2013c) has resulted in an absolute calibration of better than 1% through measuring the Doppler modulation of the 2.7 K CMB monopole. This modulation is caused by both the orbit of the satellite around L2, and the orbit of the earth around the sun, both of which are known to much better than 1%. Since the SPT data has excellent overlap with the Planck data over an intermediate range of angular scales ($\ell \in [670, 1170]$), comparing

transfer function-deconvolved bandpowers gives an overall calibration that is more than sufficient for the analysis presented here.

3.3 Cluster Extraction

Seeing galaxy clusters in the raw SPT maps by eye is fairly difficult. In addition, characterizing the selection function would be extremely difficult. Therefore, we choose to implement a numerical algorithm to extract the cluster sample. Furthermore, based on the work of Haehnelt and Tegmark (1996); Herranz et al. (2002a,b); Melin et al. (2006), if the cluster profile, transfer function and noise power spectrum are known, the optimal filter can be constructed. For details of the definition of the term "optimal", see Melin et al. (2006).

For the cluster profile, we choose to use a projected spherical isothermal β-model:

$$\Delta T = \Delta T_0 (1 + \theta^2/\theta_c^2)^{-\frac{3}{2}\beta+\frac{1}{2}},$$

where the amplitude ΔT_0 is chosen to normalize the profile and the core radius θ_c is a free parameter. We choose $\beta = 1$. After extensively checking the cluster extraction procedure, we find a negligible dependence on the choice of functional form for the assumed cluster profile. For instance, using an Arnaud profile (Arnaud et al., 2010), a Nagai profile (Nagai et al., 2007), or a Gaussian profile (with the width a free parameter) results in a very similar cluster catalog, which has no significant impact on the results presented here.

Heuristically, this can be seen in Figure 3.1. The total noise power spectrum is steeply falling at large angular scales due to contributions from primary CMB fluctuations. At small angular scales, noise power becomes increasingly important. In the range of angular scales where the SZ signal is well-measured with the SPT, the cluster profile is typically quite flat, yielding little dependence on the exact details of the assumed profile.

As mentioned in 3.2, the transfer function was initially computed from time-

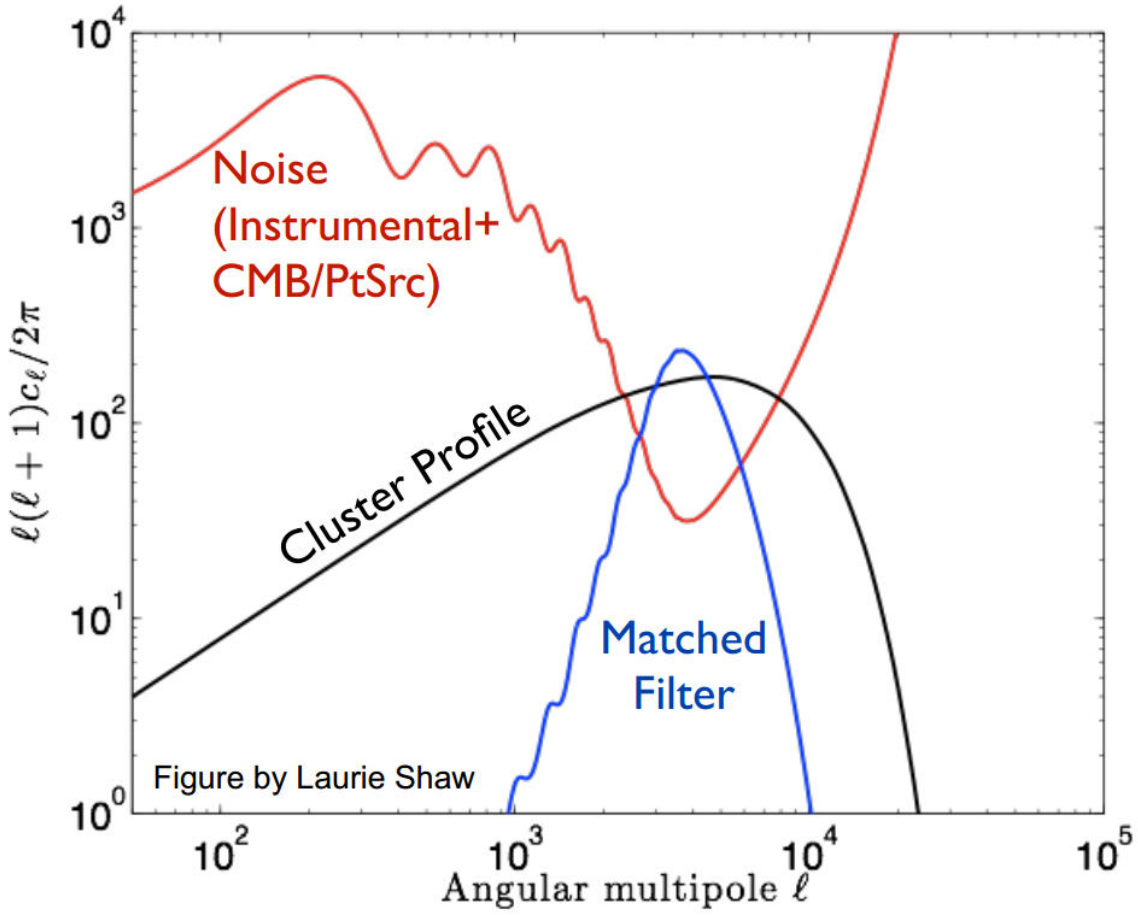

Figure 3.1 An approximate representation of the noise and signal terms resulting the SPT matched filter, shown in the 1-D spatial Fourier domain.

domain simulations. However, to make the choice of transfer function robust and computationally simple, we choose the analytic transfer function described there. For a particular subset of the SPT data (the `ra5h30dec-55` field), we verified the effect of switching to the Fourier-equivalent of time-domain simulations and find no appreciable change in the resulting cluster catalog.

The noise PSD is estimated from "jackknives": the repeated co-adding of the single observation maps while randomly multiplying half the maps by minus one. This causes any power which is stationary on the sky and hence coherent between any subsets of the data, to cancel. Noise power is incoherent on long timescales and acts such that the sign flip does not alter its relevant stochastic properties. The jackknife maps are transformed into the spatial Fourier domain, squared and averaged. This results in a noise PSD that accurately describes the SPT noise in the approximation it is gaussian. This approximation has been verified to be sufficiently accurate in the simulations described in 4.5.

With these data products in hand, the SPT maps were filtered with a spatial matched filter of the form

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

where ψ is the matched filter, B is the transfer function, S is the assumed source template, and the relevant noise power has been broken into astrophysical (N_{astro}) and noise (N_{noise}) components, all in the spatial Fourier domain.

The astrophysical sources of noise (for this measurement) N_{astro} include contributions from power from primary and (a gaussian approximation to) lensed CMB fluctuations, an SZ background, and point sources. These components are the same as those used for the cluster simulations, which are described in detail in Chapter 4.

The marginalization over the cluster profile core radius θ_c was performed numerically. We choose twelve core radii evenly spaced between $0.25'$ to $3.0'$. The resulting

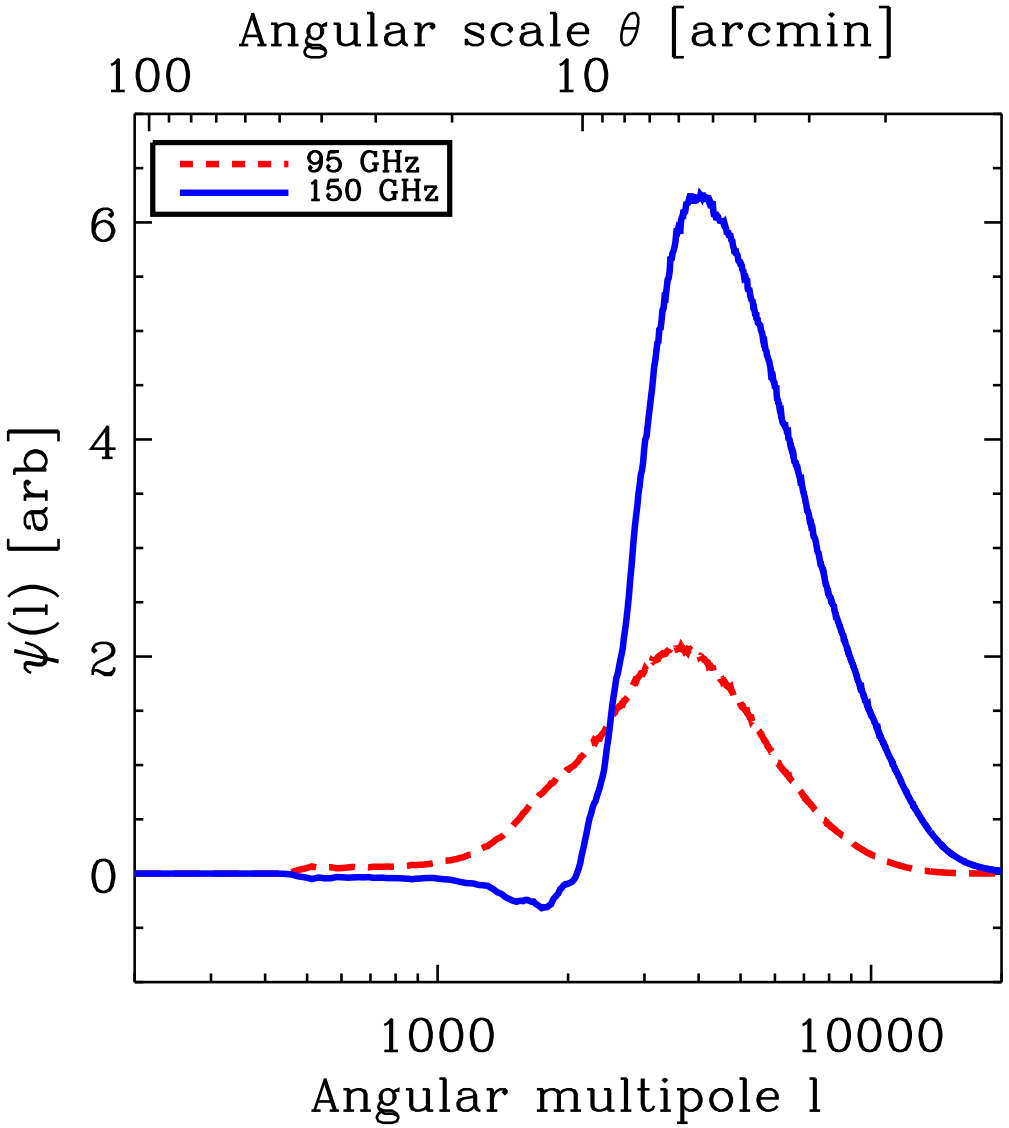

Figure 3.2 The SPT matched filter for $\theta_c = 0.25'$ is shown in the 1-D spatial Fourier domain. The filter approaches zero for low k_x modes, since those are contaminated by sources of noise that vary slowly during the scan, such as the atmospheric power and cryogenic stage drifts. The ripple present in the 150 GHz contribution to the filter is due to the contributions of primary CMB fluctuations to the total noise PSD. Figure from Bleem et al. (2014).

cluster lists are cross-matched and, for each candidate, we choose the value of θ_c that maximizes the cluster detection significance. The 150 GHz matched filter for $\theta_c = 1'$ is shown in Figure 3.2.

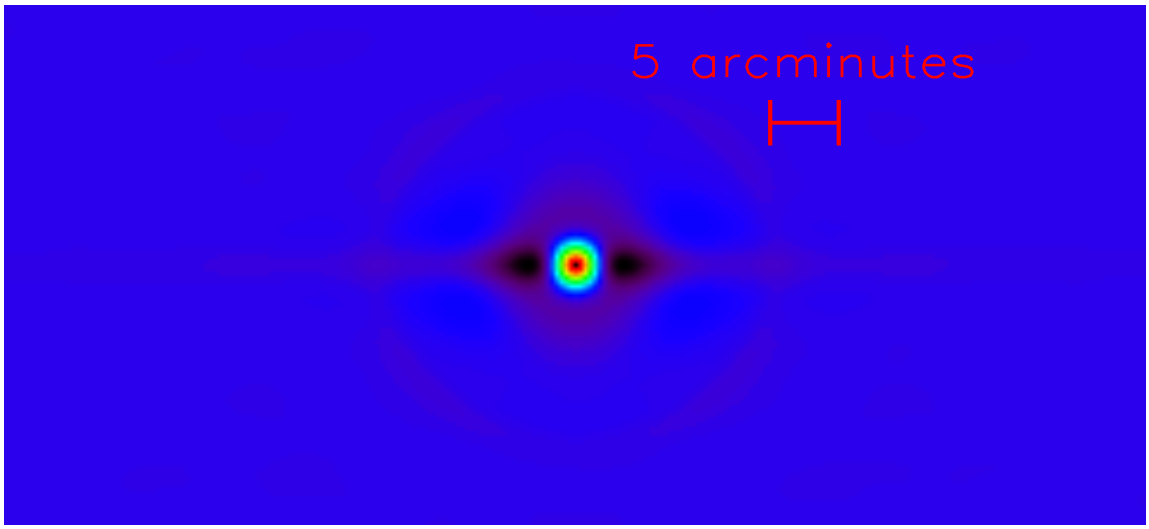

Figure 3.3 The SPT matched filter shown in the map domain. Convolving the map with this filter maximizes the signal-to-noise of a cluster with the assumed profile in the presence of the measured noise PSD N and transfer function B. Note the filter goes negative to the left and right of the main lobe in order to suppress large angular scale noise in the scan direction.

The candidates are found in the filtered map using a (negative) peak detection algorithm similar to SExtractor (Bertin and Arnouts, 1996). The signal-to-noise or detection significance, which we denote with ξ, is determined as the amplitude of the filtered map at the candidate location, divided by its standard deviation in the nearest 90 arcminute strip in declination. For the majority of this work, we only consider candidates with $\xi > 5$. In §4.5, this choice will be shown to lead to a low false detection rate of $\sim 5\%$, while still yielding a large cluster sample.

Because the matched filter supresses power on large scales, the real-space filter will qualitatively show "ringing". An example of the SPT real-space matched filter is shown in Figure 3.3. Bright point sources, once convolved with this matched filter, show significant decrements due to ringing. To avoid spurious detections of cluster

candidates from the wings of these bright point sources, all positive sources above a given flux (roughly 6 mJy, or $5\,\sigma$ in a version of the map filtered to optimize point-source signal-to-noise) were masked to a radius of $4'$ before the matched filter was applied. Roughly 150 sources were masked in each field, of which 90-95% are radio sources. Furthermore, we have also found some spurious detections near very high significance SZ clusters. For this reason, we also exclude candidates within a 10 arcminutes of $\xi > 20$ candidates. The masked area corresponds to $\sim 5\%$ of the total survey area, so any possible decrease in the measured abundance due to point source-cluster correlation (either through association or lensing effects) is negligible for this work.

We have presented the low-level data analysis used to reduce the raw SPT data into maps. To good approximation, these maps are related to the true mm-wave sky with a 2D Fourier-domain transfer function, and a 2D Fourier-domain noise power spectral density. Using this, we have used a matched filter approach in order to effectively extract cluster candidates.

4

Simulations of the Millimeter-Wave Sky

An accurate sky model of the mm-wave sky is required for understanding the selection function of the SPT galaxy cluster survey. Observed with SPT at 95, 150, and 220 GHz, the SPT-SZ maps can be modeled as a sum of several stochastic components, including

- thermal Sunyaev-Zel'dovich effect from hot gas in galaxy clusters and groups.
- primary CMB fluctuations.
- Poisson-distributed emissive point sources, with various spectral indices.
- clustered emissive point sources.
- realizations of the instrument noise using random signal-free combinations of SPT data. This includes any sources that are incoherent in time and uncorrelated with the CMB sky, such as atmospheric fluctuations.

The goal of this chapter is to reproduce realizations of the SPT sky that have statistical properties as close as possible to the real SPT-SZ observations. The model for non-SZ sources of mm-wave anisotropy were presented in §3.3. Here we will describe how the SZ sky is simulated and how these simulations are used as inputs for the cluster likelihood code presented in §6.1. In addition, we describe the use of these mm-wave sky simulations to obtain the field-to-field variation in the SPT-SZ survey of the effective cluster selection function, as well as the false detection rate.

4.1 CMB and Point source realizations

The simulated CMB skies are generated as Gaussian random fields, using the power spectrum from the best-fit WMAP7 ΛCDM model. The point source realizations are also generated as Gaussian random fields. The simulated radio population follows the results of De Zotti et al. (2005); Vieira et al. (2010) and Reichardt et al. (2012). We assume 100% correlation between the bands, a spectral index of $\alpha = -0.53$ and, at 150 GHz, an amplitude of $D_\ell = \ell(\ell+1)C_\ell = 1.28\mu K^2$ at $\ell = 3000$. The amplitudes and spectral indices of the DSFG contributions are also constrained by recent SPT measurements (Reichardt et al., 2012). At 150 GHz and $\ell = 3000$ the Poisson contribution has amplitude $D_\ell = 7.54\mu K^2$ and the clustered contribution $D_\ell = 6.25\mu K^2$; we use α=3.6 for both contributions.

4.2 SZ Simulations

The SZ sky is sourced by hot ($\lesssim 10^8$ K) plasma. Using the SPT-SZ survey, the SZ effect has been detected in individual massive clusters, detected statistically in the power spectrum (Lueker et al., 2010; Shirokoff et al., 2011; Reichardt et al., 2012), as well as a detection of tSZ signal in the bispectrum (Crawford et al., 2013), which is the Fourier-domain equivalent of the three-point correlation function.

SZ skies can be simulated using several techniques. Here we show several of these techniques and motivate using the most sophisticated and up-to-date SZ model as an input to the cosmological likelihood analysis that is central to this dissertation.

4.2.1 Shaw Semi-Analytic Gas Model

Shaw et al. (2008) use a semi-analytic technique to paste gas models onto halos identified in purely gravitational N-body simulations (Ostriker et al., 2005; Bode et al., 2007). Loosely, the algorithm solves the equation of hydrostatic equilibrium for the halo snapshot in the N-body simulation, adds the appropriate gas, computes the comptonization, and integrates along the line of sight. The resulting 2-D image

is pasted into 100 square degree maps. Once all halos are treated this way, the result is a sky map, an example of which is shown in Figure 4.2.

4.2.2 Arnaud Profile Simulations using DES Mocks

Another rudimentary way to simulate an SZ sky is to assume an analytic gas model for each halo. We chose the model from Arnaud et al. (2010), which has been calibrated by local X-ray observations. We then used the 5000 square degree DES Aardvark 0.2 simulation, pasting on the relevant profile for each halo, given its mass and redshift. A typical resulting simulated SZ map is shown in Figure 4.1.

4.2.3 Hydrodynamical Simulations

The OverWhelmingly Large Simulation project (OWLS) is a large smoothed particle hydrodynamical simulation. It contains prescriptions for non-gravitational physics such as star formation and AGN feedback (McCarthy et al., 2013; Le Brun et al., 2013; Schaye et al., 2010). It is therefore expected to be a more realistic SZ sky, albeit over a smaller area.

4.3 SZ-Mass Scaling Relation from Simulations

In §3.3, we introduced the cluster significance ξ, which is simply the peak height of the filtered synthesized SZ map, divided by the RMS of this map. This is the SZ observable used in this work. Because the observable allows for a very well understood selection function, and cosmological constraints are dominated by the unknown normalization of the scaling relation, we find no advantage to going to a different SZ observable such as the re-extracted Y_{SZ}.

Here, we introduce another SZ parameter related to the cluster detection process. This second parameter is the unbiased significance ζ. It is defined as the value of ξ that would be found in the absence of instrumental noise and astrophysical contaminants. Due to the fact that ξ is determined by maximizing the significance after searching in two-dimensional position space and source template size, the average ξ found across

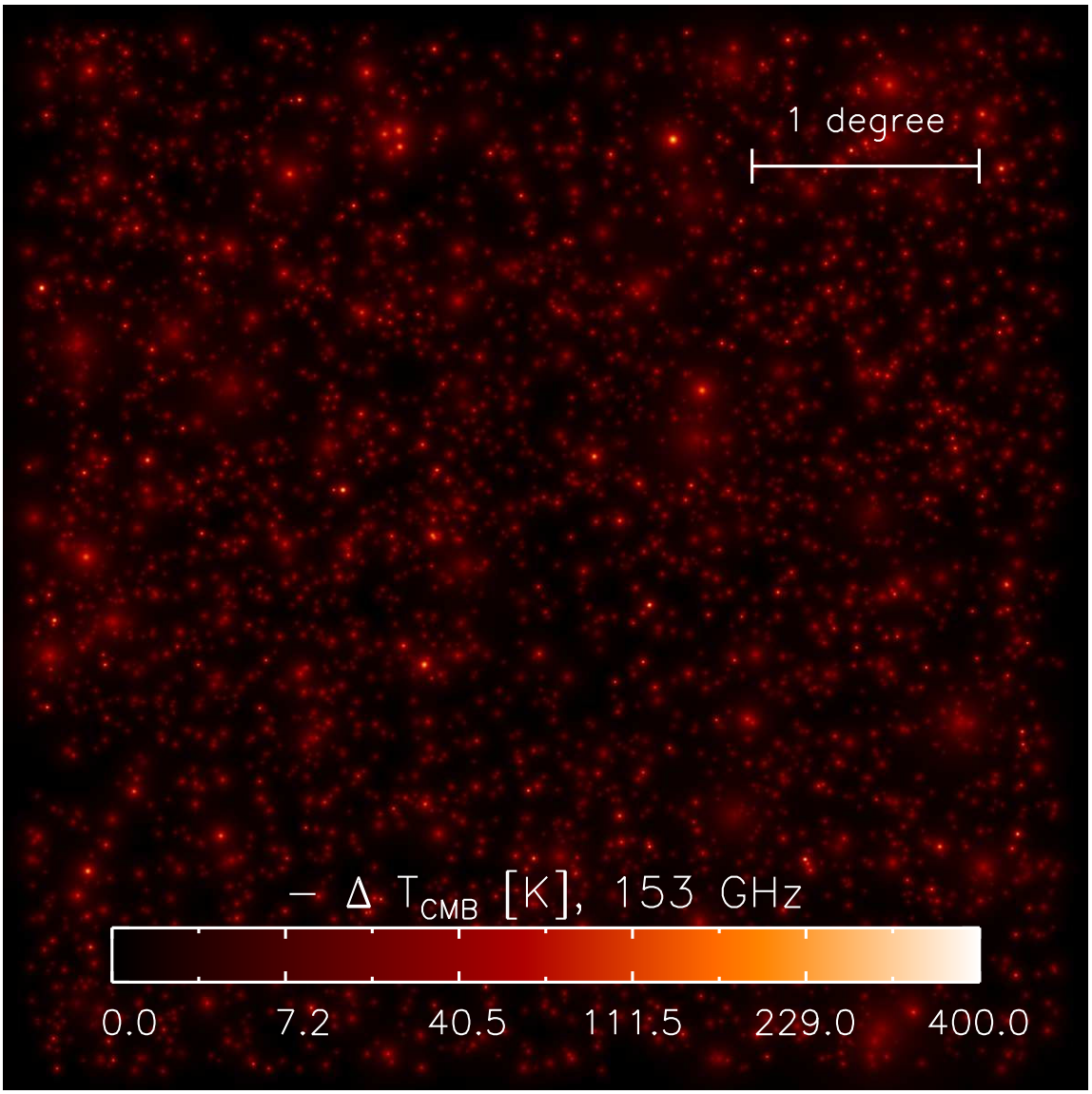

Figure 4.1 This region of simulated SZ sky was created by pasting Arnaud profiles onto halos identified in a large N-body simulation. Note the azimuthal symmetry of each individual halo.

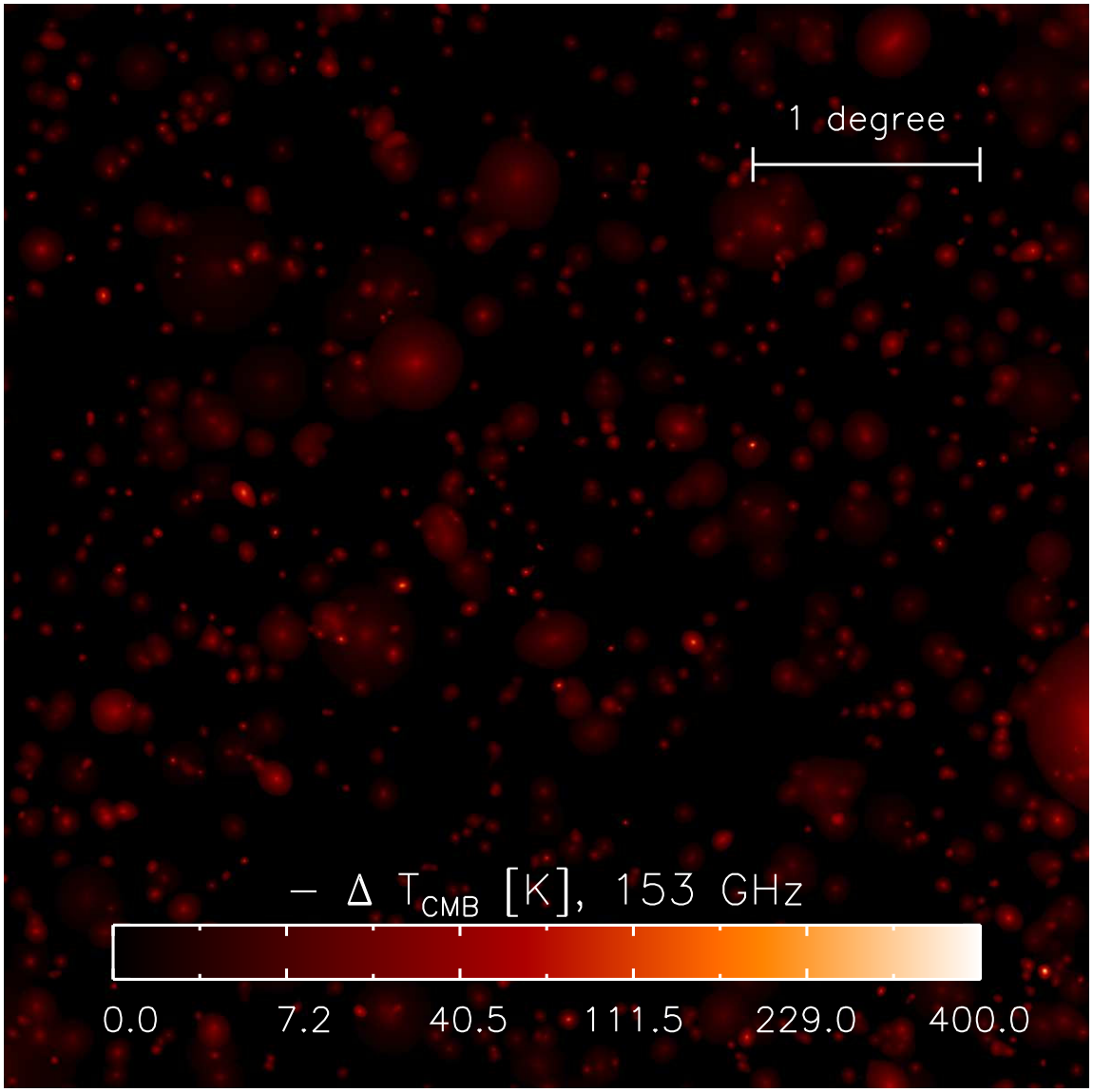

Figure 4.2 In this SZ simulation, the Bode gas model is computed for each halo identified in the N-body simulation, and pasted onto this map. Note that the individual halo gas models are truncated.

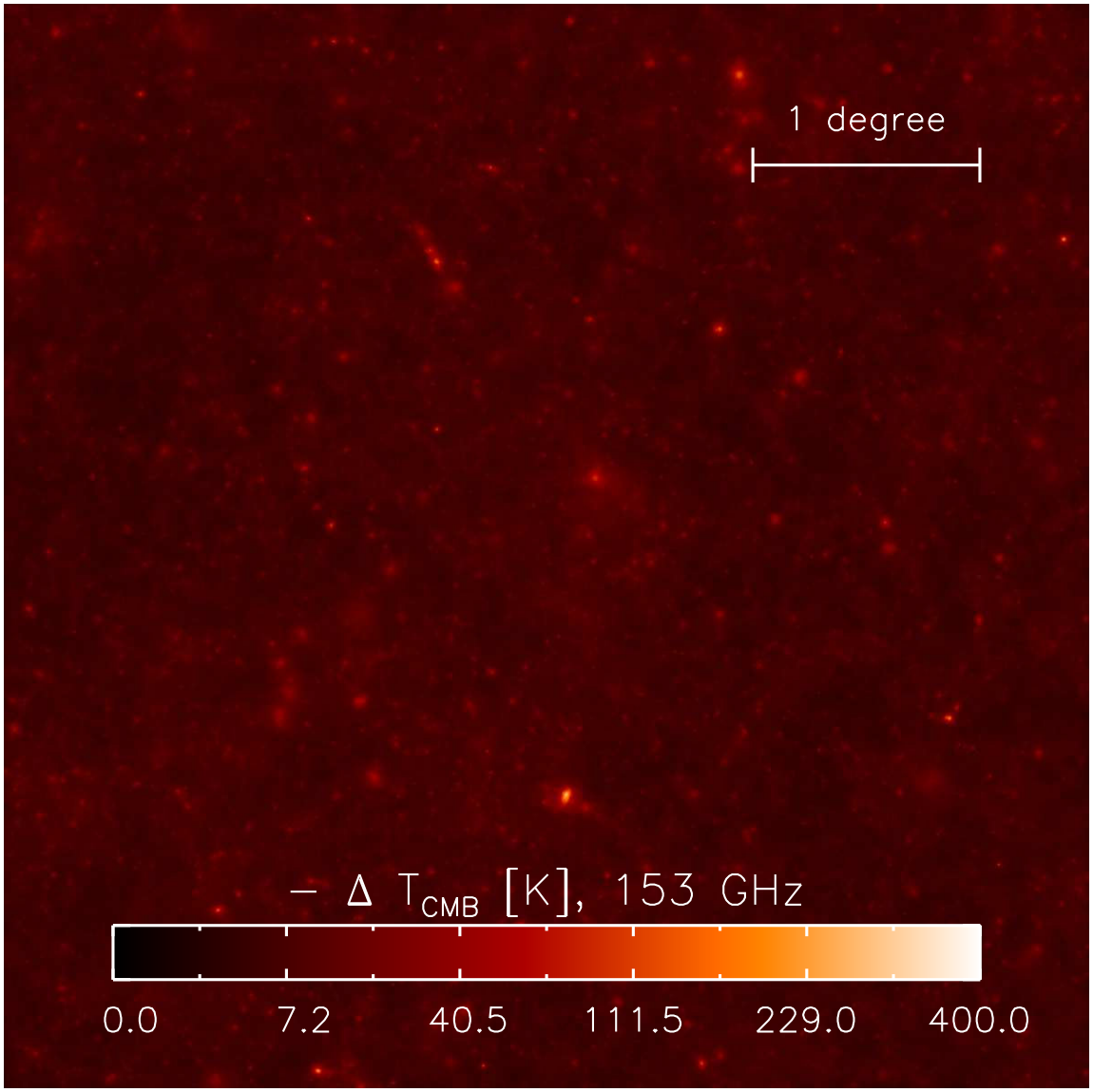

Figure 4.3 This simulation was performed using smoothed particle hydrodynamics. This realization of the simulated SZ sky does not rely on identifying individual halos.

many noise realizations is given by:

$$\langle\xi\rangle^2 = \zeta^2 + 3. \tag{4.1}$$

Loosely, ξ^2 can be considered as the χ^2 away from the null hypothesis. Due to the fact that there are three independent nuisance parameters that are being marginalized over in the case of $\langle\xi\rangle$ (and none in the case of ζ^2) we expected three degrees of freedom to bias the effective χ^2 by 3 on average. This has been verified in simulations (see Figure 4.4). Since this relation cannot hold down to very low ζ, we only model this maximization bias for $\zeta > 2$. We find that changing the location of this cutoff to $\zeta > 1.5$ or $\zeta > 2.5$ has negligible impact on the results presented in this work. Due to the definition of ξ, and the fact that the astrophysical contaminants and instrument noise are Gaussian to a high degree, it is related to $\langle\xi\rangle^2$ by a Gaussian of unit width. We then parameterize the ζ - M scaling relation as

$$\zeta = A_{SZ}\left(\frac{M}{3\times10^{14}M_{\odot}h^{-1}}\right)^{B_{SZ}}\left(\frac{E(z)}{E(0.6)}\right)^{C_{SZ}}, \tag{4.2}$$

with an additional log-normal intrinsic scatter parameter $\sigma_{\ln\zeta}$.

In order to obtain a ζ-mass relation from the simulated SZ sky, we run a modified version of the matched filter algorithm that takes both a simulated map containing all the components mentioned in §4, as well as a pure SZ map. The SZ map is used to identify candidates, while the full map is used to normalize the detection.

The position and ζ for each detection is then used to cross-match with the simulation catalog in order to find the true position, redshift and mass for each cluster. We use a simple cross-matching algorithm that searches for the nearest cluster above a minimum mass threshold. In order to screen out false associations we use two statistics: the positional difference between the matched filter detection and the distance away from a typical zeta-mass relation. We reject outliers with a threshold in positional difference chosen from manual inspection. For the scaling-relation based

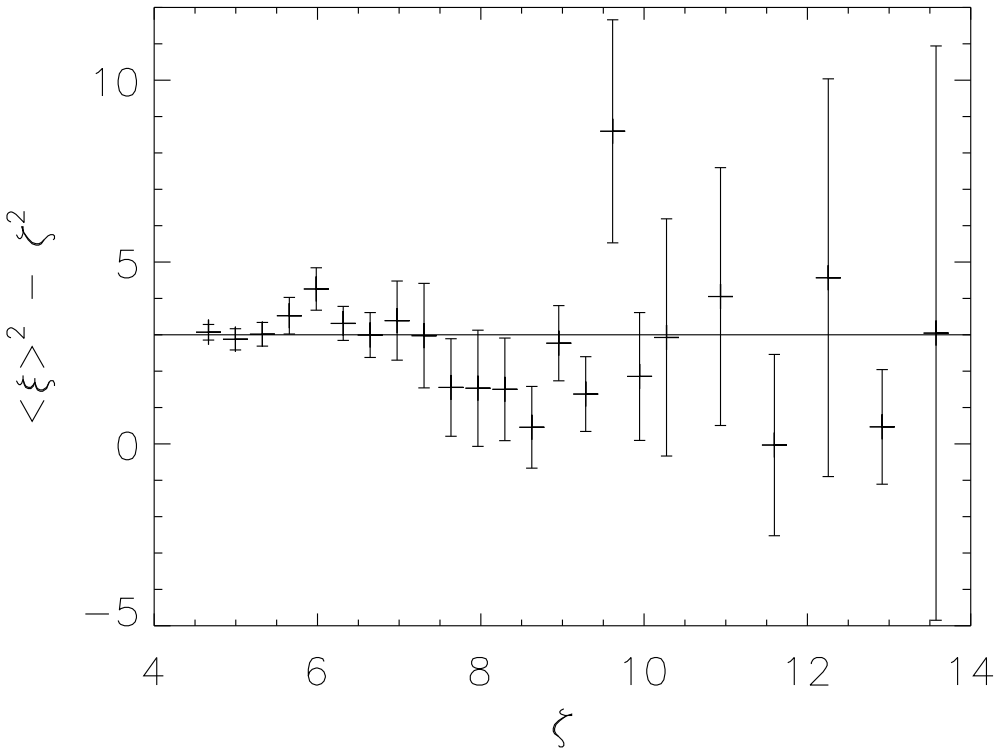

Figure 4.4 The result of 25 realizations of simulations of the relation between $\langle\xi\rangle$ and ζ. The quadratic difference between measured ζ and $\langle\xi\rangle$ is binned and plotted against a wide range of ζ. These data are consistent with $\langle\xi\rangle^2 = \zeta^2 + 3$, with $\chi^2 = 27$ for 23 degrees of freedom. Figure taken from Vanderlinde et al. (2010).

threshold, we compute the quantity $\frac{\zeta}{(M_{500})^{1.3}}$, which should be fairly constant for all clusters, we then fit a Gaussian to the histogram of those values and reject outliers that are outside $\pm 3\ \sigma$.

We then cut the cluster catalog for $M_{500} > 1 \times 10^{14} M_{\odot}/h$ and fit a scaling relation with the functional form given in Equation 4.2. The fit is performed by minimizing the mean absolute deviation from the scaling relation. We have checked that a linear least-squares method gives negligibly different results. When using the robust statistic, we find that the details of the method used for rejecting false associations have little to no impact on the final scaling relation results.

The results from the SZ simulations described in §4.2 are shown in Table 4.1. Note that in the Arnaud model, since any two halos of a given mass and redshift are assigned the same gas profile, the halo-to-halo scatter is expected to be zero. The fact that we measure $\sigma_{\ln \zeta} = 0.05$ is due to the deviations from the power law evolution parameterization assumed with both mass and redshift in 4.2.

For the cosmological analysis carried out in this work, we choose nominal simulation to be from the OWLS simulation, with AGN feedback model 8.0, because it reproduces the global X-ray, tSZ, optical, and black hole scaling relations (Le Brun et al., 2013), as well as the observed pressure profiles of the local group and cluster population. In order to encapsulate residual uncertainty in how faithfully the SZ properties are represented in this simulation, we adopt conservative priors of 30%, 20%, 100%, and 100% on A_{SZ}, B_{SZ}, C_{SZ}, and $\sigma_{\ln \zeta}$, respectively. As we will show in §7.2, these priors, while not entirely uninformative, do not dominate the constraints on the scaling relation. Rather, the inclusion of follow-up observations in the X-ray provide the bulk of the constraining power.

4.4 Field Scaling Factors

The SPT-SZ fields have somewhat different noise levels. This adds a systematic offset to the ζ-mass relation for each field. Ignoring this offset would increase the

Scaling Relation Parameter	Shaw09 WMAP5 highfb2	**OWLS AGN8.0**	OWLS AGN8.5	OWLS AGN8.7	OWLS NO-COOL_UVB	OWLS REF	Arnaud
A_{SZ}	6.24	**5.38**	4.36	3.28	6.05	4.69	6.20
B_{SZ}	1.33	**1.34**	1.39	1.42	1.22	1.27	1.37
C_{SZ}	0.83	**0.49**	0.72	0.81	-0.07	0.33	0.34
$\sigma_{\ln\zeta}$	0.24	**0.13**	0.12	0.14	0.16	0.14	0.05

Table 4.1 The SPT-SZ scaling relation parameters extracted from SZ simulations. The nominal simulation is shown in bold.

inferred intrinsic scatter in the relation. Therefore, we choose to simulate the zeta-mass relation for each field separately. We find that the change in B_{SZ}, C_{SZ}, and $\sigma_{\ln\zeta}$ is small i.e. adds negligible scatter to the overall ζ-mass relation when ignored. We therefore choose only to rescale the overall normalization (A_{SZ}) per field. For the i^{th} field, we rescale as $A^i_{SZ,eff} = FSF^i A_{SZ}$. Normalized to the 2008 observations of the `ra5h30dec-55` field, the field scaling factors (FSF) are shown in Table 4.2. We have checked the simulations by adding known cluster profiles to the real maps, applying the cluster-finding algorithm and checking the recovered unbiased significance. This semi-analytic test agrees well with the results of the simulations.

4.5 False Detection Rate Simulations

For the $\xi > 5$ cut used in this work, the number of false detections due to noise fluctuations is significantly greater than one in the 2500 square degree SPT-SZ survey. If the number of expected false detections is of the same order as the Poisson uncertainty on the total number of detected clusters, it needs to be taken into account in analysis. In order to simulate the false detection rate (FDR), we repeat the FSF simulations with the SZ component omitted. Any detections are then false by definition.

The result of the FDR simulations is shown in Figure 4.5. The total number of expected false detections is 18.5 for the full 2500 square degree survey at $\xi > 5$. We

Field Name	Field Scaling Factor
ra5h30dec-55	1.33
ra23h30dec-55	1.39
ra21hdec-60	1.29
ra3h30dec-60	1.25
ra21hdec-50	1.11
ra4h10dec-50	1.27
ra0h50dec-50	1.14
ra2h30dec-50	1.19
ra1hdec-60	1.18
ra5h30dec-45	1.08
ra6h30dec-55	1.16
ra3h30dec-42.5	1.20
ra23hdec-62.5	1.18
ra21hdec-42.5	1.15
ra1hdec-42.5	1.19
ra22h30dec-55	1.13
ra23hdec-45	1.19
ra6h30dec-45	1.16
ra6hdec-62.5	1.18

Table 4.2 Field scaling factors for the 19 SPT-SZ fields. These are used to rescale the ζ-mass relation for each field.

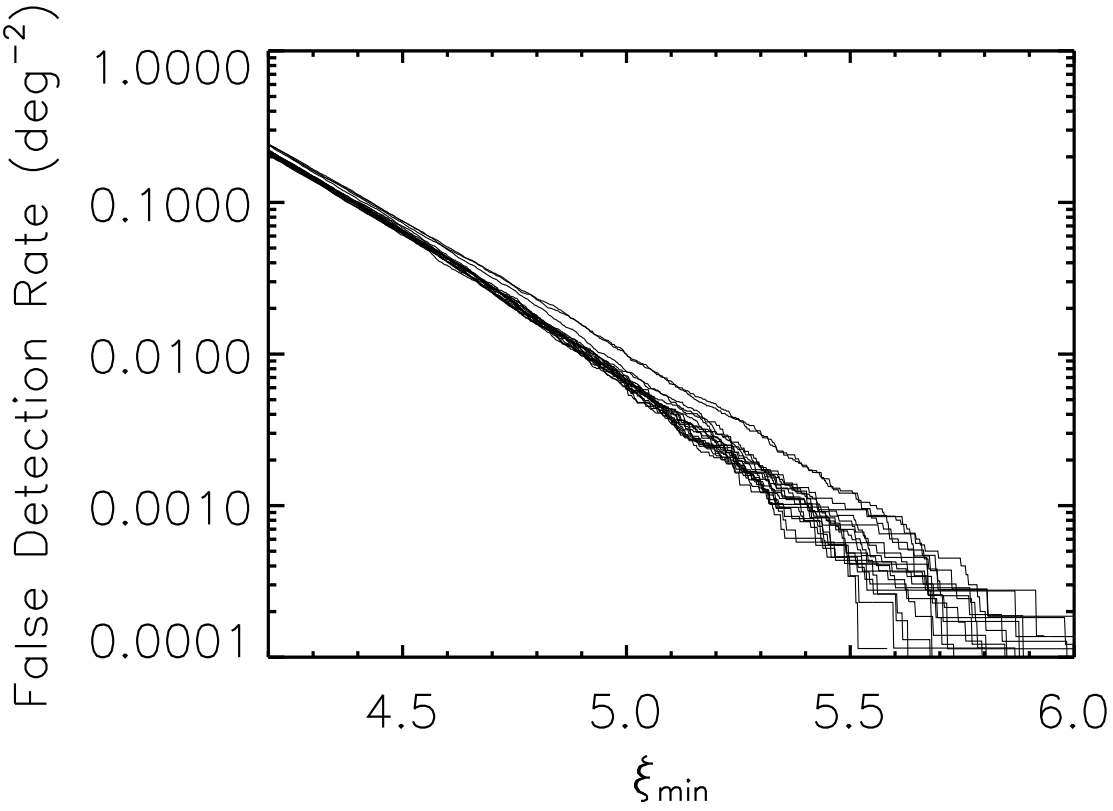

Figure 4.5 The number of false detections found per square degree of simulated sky for each of the 19 SPT fields.

Field Name	α	β
ra5h30dec-55	16.8	4.6
ra23h30dec-55	17.6	4.0
ra21hdec-60	25.6	4.1
ra3h30dec-60	20.5	4.7
ra21hdec-50	25.3	4.1
ra4h10dec-50	16.8	5.5
ra0h50dec-50	20.8	5.1
ra2h30dec-50	15.0	4.8
ra1hdec-60	17.2	5.4
ra5h30dec-45	15.9	4.8
ra6h30dec-55	17.8	4.6
ra3h30dec-42.5	16.8	4.3
ra23hdec-62.5	14.9	4.9
ra21hdec-42.5	17.1	4.5
ra1hdec-42.5	18.4	5.6
ra22h30dec-55	16.5	5.2
ra23hdec-45	17.0	5.2
ra6h30dec-45	14.8	4.2
ra6hdec-62.5	16.5	4.7

Table 4.3 False detection rate parameters for the 19 SPT-SZ fields.

fit a model of the form

$$N(>\xi) = \alpha \exp\left(-\beta(\xi - 5)\right) \ , \tag{4.3}$$

where $N(>\xi)$ is defined as the number of false detections expected over the full survey area if the full survey had the same statistical properties as the field in question. For each field, the resulting values of α and β are shown in Table 4.3.

In this chapter, we have presented simulation work required as inputs to the cosmological results central to this dissertation. We have shown a variety of cluster simulations ranging from a simple spherically symmetric functional form per cluster to full hydrodynamical N-body simulations. These simulations provide a consistent picture of the cluster scaling relation and motivate priors on the scaling relation parameters. In addition, the simulations are combined with a full sky model to estimate the false detection rate (FDR) and scaling relation re-normalization factors between SPT fields (field scaling factor, or FSF).

5

The Cluster Selection Function

In this chapter, we will outline a crucial point for cluster cosmology; many aspects will be widely applicable to any statistical analysis relying on objects selected from a non-uniform population distribution. When samples are selected from a steeply falling underlying population function using a noisy measurement, the resulting statistical properties of the sample can be counter-intuitive. For instance, a cluster of a slightly lower mass is more likely to scatter up into an SZ-selected sample than a higher mass cluster scattering down out of the sample, simply because there are more low mass clusters in the underlying population. In observational astronomy, this effect on the measured population is known as *Eddington bias*.

Here, we will present a discussion of the cluster mass function, selection function and how to perform posterior mass estimates in the context of the SPT-SZ survey.

5.1 Mass Function

As explained in 1.2, the cluster mass function describes the abundance of clusters as a function of mass and redshift, or $dN/dMdz$. This is the quantity that is predicted from N-body simulations, and allows cluster abundance measurements such as the one presented in this dissertation to place constraints on the underlying theoretical model.

We use the fitting functions from the work of Tinker et al. (2008) to calculate the cluster mass function based on the cosmological parameters and associated matter power spectra estimated using the standard software package CAMB (Lewis et al.,

M_{500}, z	$\frac{\partial \ln dN}{\partial \ln \sigma_8}$
$z = 0.0, M_{500} = 10^{14} M_\odot$	1.9
$z = 0.5, M_{500} = 10^{14} M_\odot$	2.8
$z = 1.0, M_{500} = 10^{14} M_\odot$	4.5
$z = 0.0, M_{500} = 10^{15} M_\odot$	7.4
$z = 0.5, M_{500} = 10^{15} M_\odot$	10.2
$z = 1.0, M_{500} = 10^{15} M_\odot$	15.4
$\int dM dz \, \mathrm{sf}_{\mathrm{SPT}}(M, z)$	6.1

Table 5.1 The power-law scaling of cluster abundance with changes in σ_8. The last row shows the abundance integrated over the SPT selection function in mass and redshift i.e. the total number of clusters measured in an SPT-like survey is proportional to $\sigma_8^{6.1}$. Note that, at fixed number count, measuring cluster abundance at high redshift is a more powerful probe of σ_8 than a low-redshift measurement.

2000). The Tinker mass function is derived from large purely gravitational N-body simulations that are expected to be accurate to better than 5% (Jenkins et al., 2001; Warren et al., 2006; Bhattacharya et al., 2011), which is a subdominant source of uncertainty for the work presented here.

Figure 5.1 shows a typical mass function, which can be seen to be very steep. In addition to being steep with mass, the cluster abundance also scales steeply with cosmological parameters. Table 5.1 shows the scaling of the cluster abundance with σ_8, the normalization of the matter power spectrum. Clearly, this steepness implies that measurements of the cluster mass function are a statistically powerful way of probing cosmological parameters.

5.2 Selection Function

The SPT selection function used in this work is $z > 0.25$, and cluster significance $\xi > 5$. The choice of selection on ξ as a variable is natural, since this is simply choosing a cutoff in the list of SPT cluster candidates sorted by detection significance, as found using the method described in §3.3. The numerical choice of cutoff was mostly based on the false detection rate described in §4.5. For a ~ 400 cluster sample (approximately the size of the SPT sample with $\xi > 5$), we expect on the order

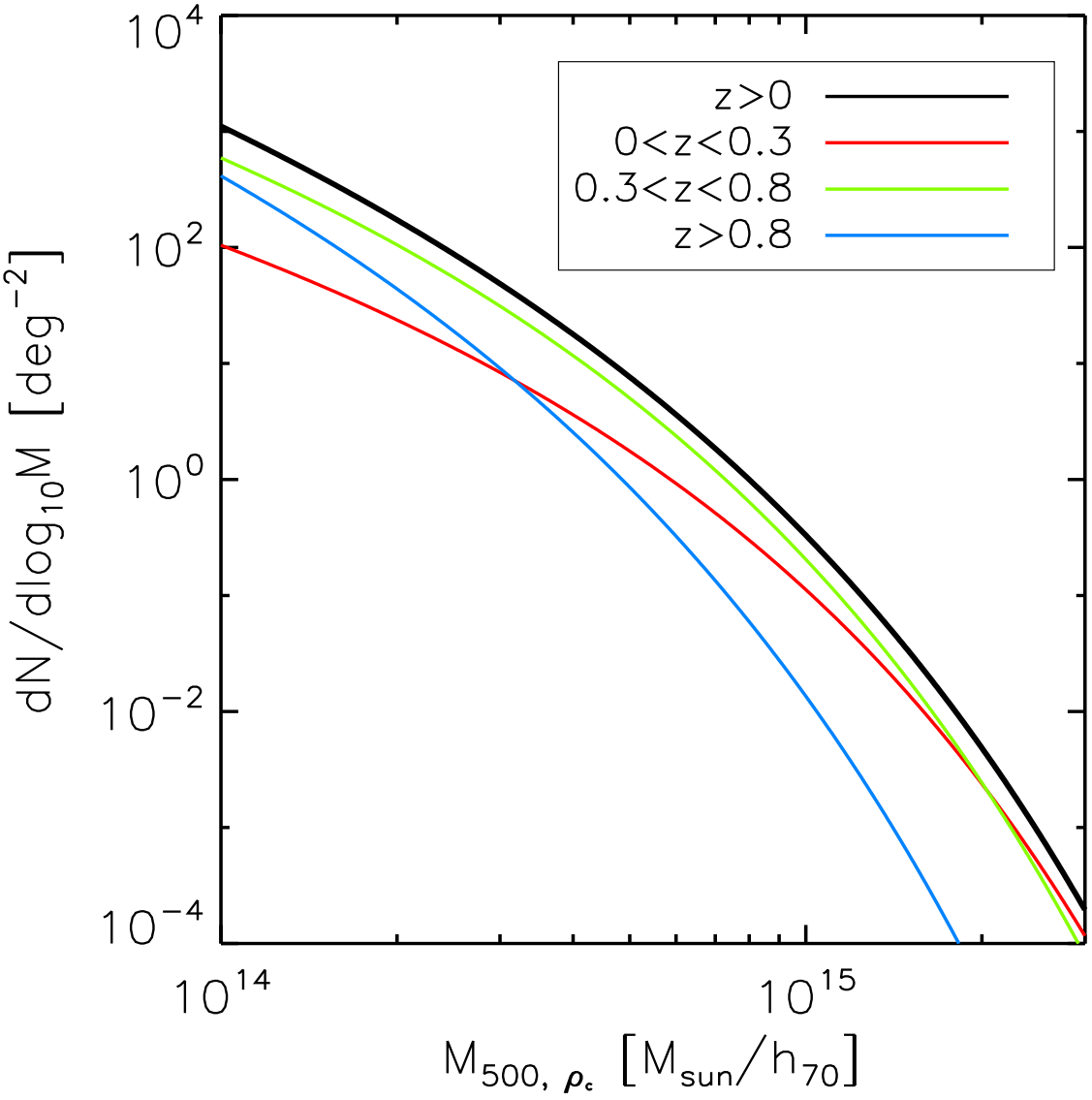

Figure 5.1 The Tinker mass function for some fiducial choice of cosmological parameters is shown in black. The colored curves show the contributions from different redshift ranges. Note the cluster abundance for the highest masses is sourced primarily by the lowest redshifts, due to growth of structure. The lower mass range is dominated by higher redshifts, due to the increased surveyed volume.

of 20 false detections, equal to the Poisson error of $\sqrt{400} = 20$. This means that, as long as the false detection rate is relatively well-understood, it is a subdominant source of error. In addition, uncertainty in the $\langle\xi\rangle$-ζ relation (Equation 4.1) at low ξ was another driver for us to choose the conservative $\xi > 5$ cutoff. If the dominant systematic uncertainty in the cosmological analysis (the SZ-mass scaling relation) can be reduced to be similar to the statistical (Poisson) uncertainty, we would consider using a more aggressive ξ cut, substantially increasing the number of total SPT-detected clusters and hence reducing the Poisson uncertainty.

The redshift cutoff was chosen due to the breakdown of power-law behaviour in the ζ - M simulations we presented in §4.2. Heuristically, the power law behaviour of mass in the redshift variable $E(z)$ has a soft index because the SZ effect is independent of redshift and the ℓ-space profiles of the clusters are relatively insensitive to the angular diameter distance to the cluster over the range $0.25 \lesssim z \lesssim 2$. At lower redshift, the angular diameter distance rapidly decreases and the ℓ-space profile shifts to lower ℓ. In combination with the extremely red power spectrum of the total noise power spectral density $\mathbf{N}$ due to the primary CMB fluctuations, this causes a break in the redshift dependence of the M-ζ relation. We therefore choose to perform this analysis for $z > 0.25$ clusters. The intrinsic mass function $dN/dMdz$ is not very steep in z around $z = 0.25$, so any errors in redshift determination (clusters with $z_{\rm true} > 0.25$ and $z_{\rm measured} < 0.25$ or vice versa) have negligible impact on the parameter constraints presented in this work.

5.2.1 Computation of the Total Cluster Number

Here we outline several algorithms for computing the total theoretically expected number of clusters using a given point in cosmological and scaling relation parameter space for the SPT cluster cosmology model. This quantity will be denoted by $N_{\rm clust}$. Not only is this computation necessary in the likelihood expression that will be developed in §6.1, the different methods provide a pedagogical introduction to concepts

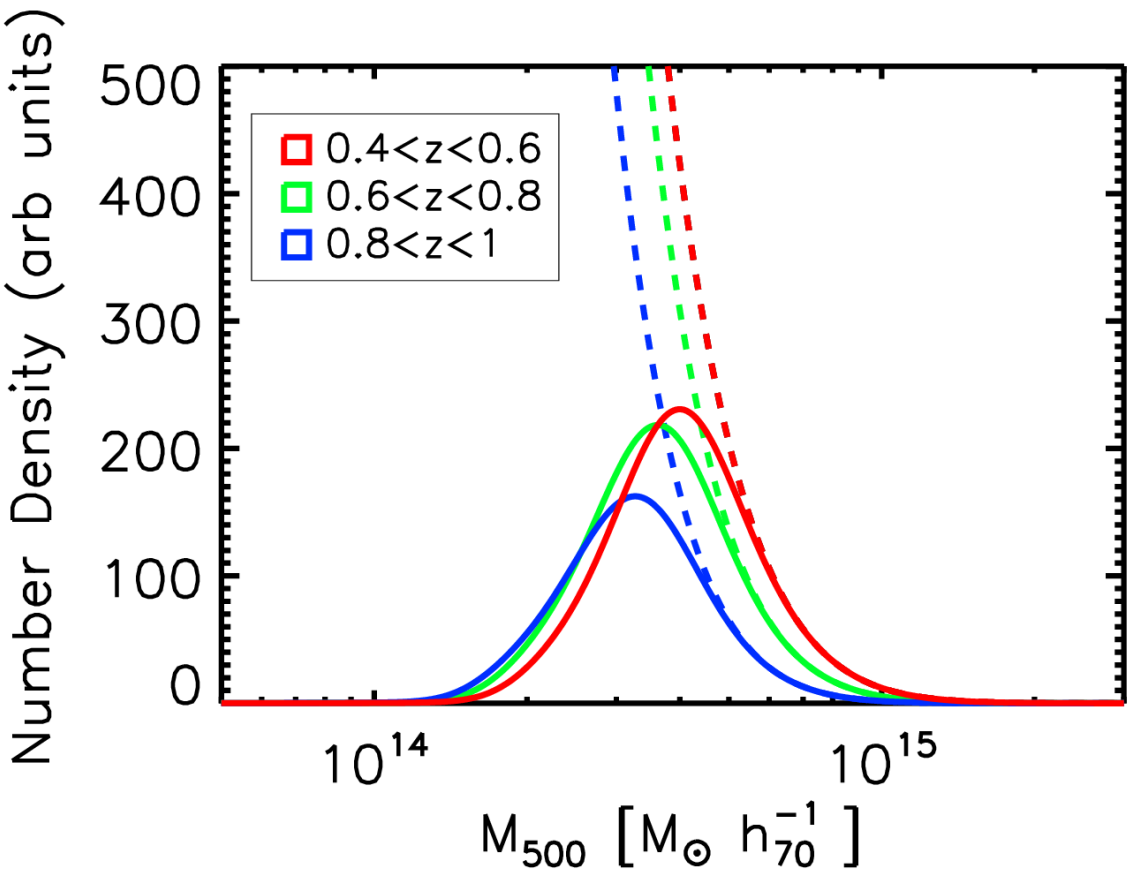

Figure 5.2 The Tinker mass function for the same choice of cosmological parameters as in Figure 5.1 is shown on a linear abundance scale. Here the mass function, integrated over several choices of redshift range, is shown as dashed curves. The solid curves show the same mass function multiplied by the SPT selection function. Note that the majority of SPT-selected clusters come from the region where the SPT sample is not 100% complete, highlighting the importance of rigorous statistical treatment of the selection function.

that are paramount to unbiased and self-consistent estimates of similar quantities like the mass of an individual cluster (§5.3).

In §5.2.1.1, we will present a computationally intensive, but highly robust Monte-Carlo method for computing the total number of expected clusters. We then present a different grid-based method in §5.2.1.2, which was used in earlier versions of this cluster likelihood analysis (Vanderlinde et al., 2010; Benson et al., 2013; Reichardt et al., 2013). The grid-based method resulted in a $\sim 10\%$ bias in the estimated number of clusters, mildly affecting implications about cosmological parameters. We will therefore introduce a highly accurate, computationally inexpensive, reverse propagation-based algorithm in §5.2.1.3. This algorithm will be used in the likelihood algorithm introduced in §6.

5.2.1.1 Monte-Carlo Method

This method draws mass, redshift pairs ($\{M, z\}$) from the mass function according to Poisson statistics. From each of these $\{M, z\}$ pairs, we generate a set of observables according to the scaling relation described in 4.3. We now have a set of $\{\xi, z\}$ pairs. We then choose all pairs with ξ and z above their respective thresholds. Since this procedure relies purely on forward modeling using a Monte-Carlo method, without any implicit applications of Bayes' theorem, it is the simplest algorithm presented here.

However, this method has an intrinsic $\pm\sqrt{N_{\rm clust}}$ error for every time it is run. One way of computing the underlying value of $N_{\rm clust}$ is to run this M number of times such that that error is reduced by a factor of $1/\sqrt{M}$. However, making this numerical error subdominant to typical statistical uncertainties for an SPT-SZ-like survey takes approximately one minute on a single CPU, which is too slow to use for the cluster likelihood that will be described in §6.1. Due to its simplicity, the Monte-Carlo method does turn out to be a valuable benchmark for comparing the numerical accuracy of alternative algorithms.

5.2.1.2 Observable-Space Grid Method

In the observable-space grid method, we generate a two-dimensional grid in (ξ, z) and integrate it numerically. The algorithm for doing so is as follows.

- We construct a two-dimensional grid of the number of clusters as a function of redshift and mass by multiplying the Tinker et al. (2008) mass function by the comoving volume element. The gridding is set to be very fine in both mass and redshift, with $\Delta z = 0.01$ and the mass binning set so that $\Delta\zeta < 0.0025$ (see below). The grids are constructed to extend beyond the sensitivity range of SPT, $0.01 < z < 3.00$ and $0.1 < \zeta < 50$. Extending the upper limits was found not to impact cosmological results, as predicted number counts have dropped to negligible levels above those thresholds.

- The parameterized scaling relation (§4.3) is used to convert the mass for each bin to unbiased significance ζ for assumed values of A_{SZ}, B_{SZ}, and C_{SZ}.

- We convolve this grid of number counts (in ζ-z space) with a Gaussian in $\ln\zeta$ with width set by the assumed intrinsic log-normal scatter $\sigma_{\ln\zeta}$ in the scaling relation.

- The unbiased significance ζ of each bin is converted to an ensemble-averaged significance $\langle\xi\rangle$ using Equation 4.1.

- We then convolve this grid with a unit-width Gaussian in ξ to account for noise and astrophysical contaminants, with the resulting grid in the native SPT catalog space, ξ-z.

- Using the lower limits in ξ and z as described in 5.2 as lower bounds, we perform a numerical integral using a spline-interpolation technique together with the trapezoid rule. The choice of numerical integration technique is found to be adequate for our choice of gridding density ($\sigma_{N_{\rm clust}} \lesssim 1$).

This method relies on forward-propagating the mass function into observable space, hence not using implicit applications of Bayes' theorem. However, the frequent use of change of variables does require several Jacobians to be computed carefully or, in other words, the changing bin sizes to be taken into account. Furthermore, the numerical error due to finite gridding converges slowly, so increasing numerical precision causes rapid growth in the required number of computational operations. Our implementation of this method was found not to agree with the expectation from the Monte Carlo method at the 10% level. It is currently unclear whether this is due to an implementation error or a fundamental problem with this algorithm.

5.2.1.3 Reverse-Propagation of the Selection Function

Like the observable-space grid method, this method also starts with a (M, z) grid. For each value of z, we compute the effective selection function in M. To do so, we use the Bayesian prescription for handling missing data; we explicitly start with the relevant indicator function $I(\xi) = \Theta(\xi - \xi_{\rm min})$ where Θ denotes the Heaviside step function. This is modified by the $\xi - \langle\xi\rangle$ relation, which is a Gaussian of unit width, denoted here with $P(\langle\xi\rangle|\xi)$. The selection function (denoted sf) becomes:

$$\mathrm{sf}(\langle\xi\rangle) = \int d\xi I(\xi) P(\langle\xi\rangle|\xi) = \frac{1}{2} \times \left(1 + \mathrm{erf}\left(\frac{\langle\xi\rangle - \xi_{\rm min}}{\sqrt{2}}\right)\right) \ . \tag{5.1}$$

Using the ζ-$\langle\xi\rangle$ relation from Equation 4.1 and the scaling relation from 4.2, we can convert this to the relevant selection function in mass. The total number of clusters is then computed using

$$N_{\rm clust} = \sum_i \int_0^\infty \mathrm{sf}(M_{500}) \left.\frac{dN}{dM_{500}dz}\right|_{z=z_i} dM_{500} \times (\Delta z) \ . \tag{5.2}$$

This method can be performed rapidly (0.2 seconds in our implementation) and yields results consistent with the Monte Carlo method to much higher accuracy than required in this work.

5.2.1.4 Discussion

The three methods presented above are canonical examples of how to handle the cluster selection function. The Monte Carlo method demonstrates a simple example from the class of algorithms using a brute-force Monte Carlo method. The model for the cluster population is written down as a set of the stochastic processes which are directly simulated in such a method, without having to write down the explicit probabilistic distributions which describe them. The use of *forward-propagation* means applications of Bayes' Theorem (and hence details of choosing priors or determining normalization constants) are avoided. Given their simplicity, such methods should always be implemented as a golden standard that should be used to cross-check more complicated methods, since selection effects in cluster population analysis can be counter-intuitive and subtle.

Our implementation of the observable-space grid method is found to disagree with the brute-force method at the $\sim 10\%$ level. While this is acceptable for the previously published works Benson et al. (2013) and Reichardt et al. (2013), this work has Poisson uncertainty comparable to that systematic. We therefore choose not to use this method for the results in Chapter 7.

Reverse-propagation of the selection function is an example of modelling a data-space operation into theory space. In this case, the data-space operation is the indicator function requiring that a cluster property exceed a minimum observable value to be included in the cluster sample. This operation can be computed in mass-space, by applying a series of changes of variable, and integrating over the relevant probabilistic distributions relating variables.

5.3 Posterior Mass Estimation

Given an ensemble of clusters at fixed mass and redshift, the scaling relations describe the distribution of clusters as a function of the observables. Considering a fixed redshift, and only the SZ (selection) observable ξ, this distribution is denoted with

$P(\xi|M)$ (the probability of observing some cluster significance, given a mass). In this section, we describe how to obtain $P(M|\xi)$, or the posterior mass estimate associated with a cluster.

First, we consider an unbiased posterior distribution estimating the true underlying cluster mass for a single given cluster. Here, we apply Bayes' theorem. The maximum entropy Bayesian prior is to weight each member of the underlying population equally i.e. to set the prior equal to the population function. This gives

$$P(M|\xi, z) \propto P(M|z)P(\xi|M, z) \ , \tag{5.3}$$

where $P(M|z) = dN/dM|_z$ describes the mass function for that slice in redshift, $P(\xi|M, z)$ is simply the scaling relation described in §4.3, including the ζ-M relation, the log-normal intrinsic scatter, the effects of maximization boosting and the unit Gaussian scatter due to measurement noise and astrophysical contaminants.

When including multiple simultaneous mass proxies that may be correlated with the selection observable, this becomes

$$P(M|\xi, \xi_1, \xi_2, ..., z) \propto P(M|z)P(\xi, \xi_1, \xi_2, ..., |M, z) \ , \tag{5.4}$$

where the likelihood factors as $P(\xi, \xi_1, \xi_2, ..., |M, z) = P(\xi|M, z)P(\xi_1|M, z)P(\xi_2|M, z)...$ in the case of uncorrelated observables.

A commonly assumed simplification (Mortonson et al., 2011; Harrison and Hotchkiss, 2013) is to linearly expand the mass function to (locally) be a power law, while taking the likelihood $P(\xi|M, z)$ to be log-normal with a variance that is small compared to the range over which γ changes. At fixed redshift, the mass function is written as

$$\frac{dN}{dM} \propto M^{\gamma} \tag{5.5}$$

where γ is the relevant power law index of the mass function (of order a few for clusters

considered in this work). Let $M_{\rm obs}$ denote the value of M obtained by substituting ξ into the scaling relation and let $\sigma^2_{\ln M}$ define the variance of the log-normal distribution used to approximate the width of the M-$M_{\rm obs}$ relation. The posterior estimate then becomes

$$\begin{aligned} P(M|\xi,z) &\propto M^{\gamma}\tfrac{1}{M_{\rm obs}}\exp\left(-\tfrac{1}{2}\tfrac{(\ln M-\ln M_{\rm obs})^2}{\sigma^2_{\ln M}}\right) \\ &\propto \exp\left(-\tfrac{1}{2}\tfrac{\ln(M)^2-2\ln M\ln M_{\rm obs}-2\sigma^2_{\ln M}\gamma\ln M}{\sigma^2_{\ln M}}\right) \\ &\propto \exp\left(-\tfrac{1}{2}\tfrac{\left(\ln(M_{\rm obs})^2-(\ln M(1+\frac{\gamma\sigma^2_{\ln M}}{\ln M_{\rm obs}}))^2\right)}{\sigma^2_{\ln M}}\right) . \end{aligned} \tag{5.6}$$

Taking the first two moments gives a simple prescription for estimating the mean and variance of this type of posterior mass estimate as

$$\begin{aligned} \ln M &= \ln M_{\rm obs}+\gamma\sigma^2_{\ln M_{\rm obs}} \\ \sigma^2_{\ln M} &= \sigma^2_{\ln M_{\rm obs}} . \end{aligned} \tag{5.7}$$

While this estimate is relevant for estimating the posterior mass of a single cluster, another type of mass estimate is relevant for population studies, such as estimating the "rareness" of a given cluster. This is a mass estimate performed at constant number density, such that

$$\left.\frac{dN}{dMdz}\right|_{M,z} = \left.\frac{dN}{dMdz}\right|_{M_{\rm obs},z} . \tag{5.8}$$

Writing these expressions explicitly and solving for the first two moments in a fashion identical to that performed above yields a prescription for estimating this type of posterior mass estimate:

$$\begin{aligned} \ln M &= \ln M_{\rm obs}+\tfrac{1}{2}\gamma\sigma^2 \\ \sigma^2_{\ln M} &= \sigma^2_{\ln M_{\rm obs}} . \end{aligned} \tag{5.9}$$

Note that the correction factor for the mass estimate relevant to population studies (constant number density) is only half that compared to the mass estimates where

the *probabilities* are set equal.

In this chapter, we have presented the theoretical abundance of clusters as a function of mass and redshift (the mass function), and shown how this translates to observable abundances. The process is non-trivial due to the selection of objects from a steeply falling population using a noisy proxy. We present numerical techniques for computing the total number of expected clusters and find that this number scales as $\sigma_8^{6.1}$ for the SPT survey. The closely related inverse problem of computing unbiased cluster mass estimates is also solved to linear order. As previously mentioned in Mortonson et al. (2011), we derive the surprising fact that the difference between the mass one measures and the underlying true mass differs by a factor of two depending on whether probability or number density is held constant.

6

SPT Cosmological Cluster Likelihood

This chapter focuses on evaluating the cosmological likelihood of the SPT-SZ cluster sample. Starting from basic Poisson statistics, we derive a naive method and show why it is computationally intractable. We then introduce a newly developed method using Bayesian statistics and Monte-Carlo sampling to estimate the cosmological likelihood, self-consistently marginalizing over an arbitrary number of mass tracers.

6.1 Likelihood Evaluation

In this dissertation we develop a different numerical technique that allows for a large number of simultaneous follow-up mass observables.

We start with the usual binned Poisson statistic. In order to keep the likelihood evaluation computationally tractable, we forward-model the theoretical cluster mass function onto observable space $\vec{x}$, and use the Poisson statistic. For a given observable-space bin x_i, the probability of observing n events, with expectation value $y(x_i)$ is

$$P_i = \frac{e^{-y(x_i)} y(x_i)^n}{n!}. \tag{6.1}$$

We choose to take the limit of small bins where y becomes arbitrarily small at all x_i, and n is zero except at the locations of x_i where clusters have been observed. If we

let x_j denote the bin that contains the $j^{\rm th}$ cluster, we obtain

$$\ln \mathcal{L} = \ln \prod_i P_i = \sum_i \ln P_i = -\sum_i y(x_i) + \sum_j \ln y(x_j). \tag{6.2}$$

In this case, we will treat the observable space as being $\vec{x} = [z, \xi, Y_{\rm X}, M_{\rm WL}]$, though this method fully (and efficiently) generalizes to a larger number of observables. Writing the model expectation value $y(x_i) = N(z_i, \xi_i, Y_{{\rm X},i}, M_{{\rm WL},i})$. Going to one-dimensional indices gives

$$\ln \mathcal{L} = -\sum_{i_1,i_2,i_3,i_4} N(z_{i_1}, \xi_{i_2}, Y_{{\rm X},i_3}, M_{{\rm WL},i_4}) + \sum_j \ln N(z_j, \xi_j, Y_{{\rm X},j}, M_{{\rm WL},j}), \tag{6.3}$$

where the i_1, i_2, i_3, i_4 sums again run over all possible values and j only runs over the bins where clusters were detected. Going to the continuous limit would result in a divergent likelihood expression. This can be understood by the fact that the model is increasingly unlikely to produce our particular realization of the cluster catalog as we go to finer binning. The divergence can be removed by simply adding $-\ln \Delta z \Delta\xi \Delta Y_{\rm X} \Delta M_{\rm WL}$ to Equation 6.3. This quantity depends only on bin size, so that $\Delta \ln \mathcal{L}$ for different values of cosmological or scaling relation parameters remains meaningful. We then obtain the expression

$$\ln \mathcal{L} = -\int dz d\xi dY_{\rm X} dM_{\rm WL} \frac{dN(z, \xi, Y_{\rm X}, M_{\rm WL})}{dz d\xi dY_{\rm X} dM_{\rm WL}} + \sum_j \ln \frac{dN(z_j, \xi_j, Y_{{\rm X},j}, M_{{\rm WL},j})}{dz d\xi dY_{\rm X} dM_{\rm WL}}. \tag{6.4}$$

Vanderlinde et al. (2010) and Benson et al. (2013) evaluated this expression on a two-dimensional and three-dimensional uniformly spaced grid, respectively. Each dimension was gridded into several hundred points, such that the computational cost rises exponentially with the number of mass observables. The computational cost was trivial for V10, challenging for B11, and likely untractable in the case of four observables.

Instead of computing Equation 6.4 on a uniformly spaced grid, we use the following

numerical techniques.

Firstly, we note that the follow-up mass observables immediately integrate out of the first term, since the selection function is purely in ξ and z. Therefore, we only need to perform a two-dimensional integral

$$-\int dz d\xi dY_{\rm X} dM_{\rm WL} \frac{dN(z,\xi,Y_{\rm X},M_{\rm WL})}{dz d\xi dY_{\rm X} dM_{\rm WL}} = -\int_{z_{\rm cut}}^{\infty} dz \int_{\xi_{\rm cut}}^{\infty} d\xi \frac{dN}{dz d\xi}. \tag{6.5}$$

We evaluate this expression with

$$\int_{z_{\rm cut}}^{\infty} dz \int_{\xi_{\rm cut}}^{\infty} d\xi \frac{dN}{dz d\xi} = \int_{z_{\rm cut}}^{\infty} dz \int_{0}^{\infty} dM \frac{dN}{dz dM} P(\xi > \xi_{\rm cut}|M), \tag{6.6}$$

where $P(\xi > \xi_{\rm cut}|M)$ is simply the significance cut modelled through the scaling relation as described in §4.3.

The second term of Equation 6.4 is more challenging to evaluate. Going back to more compact notation with $\vec{x}$ being a general set of mass observables (previously $[\xi, Y_{\rm X}, M_{\rm WL}]$), we start by writing the expectation density as an integral over the mass function,

$$\frac{dN(\vec{x}_j)}{d\vec{x}} = \int d\ln M' \prod_i \int dx_i' P(x_{i,j}|x_i') P(\vec{x}'|z, \ln M') \frac{dN}{dz d\ln M'} \ . \tag{6.7}$$

The probability density functions $P(x_{i,j}|x_i')$ describe the measurement error of the $i^{\rm th}$ observable. The $P(\vec{x}'|z, \ln M')$ describe the joint scaling relations, implemented with multi-dimensional log-normal intrinsic scatter. In this case of three mass observables (ξ, $Y_{\rm X}$, $M_{\rm WL}$), this is implemented with 6 parameters: three parameters describing the variance and three correlation coefficients, each of which is marginalized over in this work. The x_i integrals must be evaluated in the vicinity of every cluster, for every value of mass, at every step of the global Markov Chain Monte Carlo (MCMC). While this appears computationally challenging, we will now present an efficient numeric method.

6.1.1 Analytic Simplification and Monte-Carlo Integration

The computational bottleneck for evaluating the cluster likelihood in the presence of several mass observables lies in evaluating Equation 6.7. Considering a slice in redshift, we have to perform an $(N_{obs}+1)$-dimensional integral. To do so, we perform Monte-Carlo integration, randomly sampling the mass function $\frac{dN}{d\ln M'}$ with points drawn from the probability distribution $\prod_i P(x_{i,j}|x'_i)P(\vec{x}'|\ln M')$. Let $\ln\vec{m}'$ denote the integration variable cluster mass observables (i.e. $\vec{x}'$ substituted into the observable-mass scaling relations). We can then write

$$P(\ln\vec{m}'|\ln M') = \frac{1}{\sqrt{(2\pi)^{N_{obs}}|\mathcal{C}|}}\exp\left(-\frac{1}{2}((\ln\vec{m}'-\ln M')^\top\mathcal{C}^{-1}(\ln\vec{m}'-\ln M')\right), \tag{6.8}$$

where $\mathcal{C}$ is the mass observable covariance matrix containing the intrinsic scatter and correlation coefficient parameters introduced in §7.1 such that $\mathcal{C}_{kl} = \langle(m_i - M)(m_j - M)\rangle_{kl}$. Now, we wish to obtain samples of the integration variable $\ln M'$, given the bin at which we are attempting to evaluate $dN(\vec{x}_j)/d\vec{x}$. We thus apply Bayes' theorem

$$P(\ln M'|\vec{x}_j) = \frac{\int\prod_i P(x_{i,j}|x'_i)P(\vec{x}'|\ln M')d\vec{x}'}{\int\prod_i P(x_{i,j}|x'_i)P(\vec{x}'|\ln M')d\ln M'}\prod_i P(x_{i,j}|x'_i)P(\vec{x}'|\ln M'). \tag{6.9}$$

It can be shown that the prefactor is unity when the probabilities involved are either Gaussian or log-normal. However, this condition does not hold due to the treatment of maximization bias in §4.3. Instead, the two integrals become

$$\begin{aligned}
\int\prod_i P(x_{i,j}|x'_i)P(\vec{x}'|\ln M')d\vec{x}' &= \int\exp\left(-\tfrac{1}{2}\left(\langle\xi\rangle-\zeta\right)^2\right)\frac{\exp\left(-\frac{1}{2}\frac{\left(\ln M'-\ln m_\zeta(\zeta)\right)^2}{\left(F_{SZ}/B_{SZ}\right)^2}\right)}{\zeta}d\zeta \\
&= \frac{1}{\sqrt{2\pi\left(F_{SZ}/B_{SZ}\right)^2}}\sqrt{2\pi} \\
\int\prod_i P(x_{i,j}|x'_i)P(\vec{x}'|\ln M')d\ln\vec{M}' &= \int\exp\left(-\tfrac{1}{2}\left(\langle\xi\rangle-\zeta\right)^2\right)\frac{\exp\left(-\frac{1}{2}\frac{\left(\ln M'-\ln m_\zeta(\zeta)\right)^2}{\left(F_{SZ}/B_{SZ}\right)^2}\right)}{\zeta}d\ln M' \\
&= \frac{1}{\sqrt{2\pi\left(F_{SZ}/B_{SZ}\right)^2}}\int\exp\left(-\tfrac{1}{2}\left(\langle\xi\rangle-\zeta\right)^2\right)\tfrac{1}{\zeta}d\zeta,
\end{aligned} \tag{6.10}$$

such that the prefactor becomes the computationally simple one-dimensional integral

$$\frac{\int \prod_i P(x_{i,j}|x'_i) P(\vec{x}'|\ln M') d\vec{x}'}{\int \prod_i P(x_{i,j}|x'_i) P(\vec{x}'|\ln M') d\ln M'} = \frac{1}{\sqrt{2\pi}} \int \exp\left(-\frac{1}{2}\left(\langle\xi\rangle - \zeta\right)^2\right) \frac{1}{\zeta} d\zeta, \tag{6.11}$$

which we evaluate using the trapezoid rule. In order to efficiently draw samples from Equation 6.9 we first, for each i, draw samples from the measurement error $P(x_{i,j}|x'_i)$, which is assumed to be independent for each i. We then substitute these values into the scaling relations to obtain an ensemble of $\ln \vec{m}'$. The remaining task is then to draw random deviates $\ln M'$ that follow the probability distribution explicitly shown in Equation 6.8, given each value of $\ln \vec{m}'$. To do so, we first define

$$\begin{aligned} \Psi &= (\ln \vec{m}' - \ln M')^\top \mathcal{C}^{-1} (\ln \vec{m}' - \ln M') \\ &= \Sigma_{ij} (\ln m'_i - \ln M')(\ln m'_j - \ln M')(\mathcal{C}^{-1})_{ij}. \end{aligned} \tag{6.12}$$

We proceed to complete the square

$$\begin{aligned} \Psi &= \Sigma_{ij} \ln m'_i \ln m'_j (\mathcal{C}^{-1})_{ij} - 2 \ln M' \Sigma_i \ln m'_i \Sigma_j (\mathcal{C}^{-1})_{ij} + \ln M'^2 \Sigma_{ij} (\mathcal{C}^{-1})_{ij} \\ &= T_0 - 2 T_1 \ln M' + T_2 \ln M'^2 \\ &= \frac{(\ln M' - T_1/T_2)^2}{\left(1/\sqrt{T_2}\right)} + T_0 - \frac{T_1^2}{T_2}, \end{aligned} \tag{6.13}$$

where $T_0 = \Sigma_{ij} \ln m'_i \ln m'_j (\mathcal{C}^{-1})_{ij}$, $T_1 = \Sigma_i \ln m'_i \Sigma_j (\mathcal{C}^{-1})_{ij}$ and $T_2 = \Sigma_{ij} (\mathcal{C}^{-1})_{ij}$. which is quadratic in $\ln M'$, such that Equation 6.8 is a log-normal distribution in M' with a known mean, width and normalization. We compute T_0, T_1 and T_2 explicitly and sample from the resulting distribution.

Having obtained samples of $\ln M'$, we average the mass function $dN/d\ln M'$ and combine with Equation 6.11 to obtain an unbiased estimator for Equation 6.7. Our implementation of this estimator has been demonstrated to be unbiased through extensive simulations. These simulations involve drawing samples from the mass function and forward-modeling them into a catalog of mock SZ-selected clusters, with all appropriate observables.

The fit to the simulated catalogs recovers the input parameters with well-behaved residuals. The error on the mean is found to decrease as the inverse square root of the number of deviates drawn. In practice, for three mass observables and several hundred clusters, we draw 5000 deviates per cluster, resulting in RMS noise on the likelihood surface of $\Delta\chi^2 \lesssim 0.1$, with an approximate execution time of one second on a single CPU thread for the 2500 square-degree SPT-SZ sample used in this work.

This likelihood module is implemented in IDL and integrated into CosmoMC (Lewis et al., 2000) using Callable IDL.

In this chapter, we have presented a novel statistical technique for simultaneously constraining cosmological parameters and scaling relation parameters. The method is computationally inexpensive, even in the presence of a large number of mass observables.

7

Results

In this chapter, we use the cluster abundance measurement from the 2500 square degree SPT-SZ survey to constrain cosmological parameters. The uncertainties are dominated by systematics: in particular the SZ-mass scaling relation presented in §4.3. Here we assume the Gaussian priors described there, corresponding to conservative widths centered on the results from the **AGN8.0** OWLS simulation (Le Brun et al., 2013; McCarthy et al., 2013). In addition, we incorporate *Chandra* X-ray information to improve knowledge of the scaling relation.

7.1 Yx-Mass Scaling Relation Parameterization

The follow-up mass observable used in this work builds on the work of Vikhlinin et al. (2009), Benson et al. (2013), and Reichardt et al. (2013). We use $Y_{\rm X}$ as a proxy for total cluster mass, as described in McDonald et al. (2013). We write the scaling relation as

$$\frac{M_{500}}{10^{14}M_{\odot}/h} = \left(A_{\rm X}h^{3/2}\right)\left(\frac{Y_{\rm X}}{3\times10^{14}M_{\odot}\,{\rm keV}}\right)^{B_{\rm X}}E(z)^{C_{\rm X}}, \tag{7.1}$$

where the parameters $A_{\rm X}$, $B_{\rm X}$, $C_{\rm X}$ describe the normalization, mass evolution and redshift evolution of the relation, respectively. We introduce the parameter $\sigma_{\ln Y_{\rm X}}$ which models a log-normal intrinsic scatter in the $Y_{\rm X}$-M relation, as well as a correlation coefficient $\rho_{\zeta,Y_{\rm X}}$.

In addition to the SZ-mass scaling relation, this means we marginalize over a total

Parameter	Parameter description
A_s	Normalization of spectrum of primordial fluctuations
n_s	Power law slope of primordial fluctuations
Ω_c	Energy density of cold dark matter today
Ω_b	Energy density of baryonic matter today
θ_s	Characteristic angular scale of primary CMB fluctuations
τ	Optical depth to reionization
A_{SZ}	Normalization of the SZ - mass scaling relation
B_{SZ}	Mass evolution of the SZ - mass scaling relation
C_{SZ}	Redshift evolution of the SZ - mass scaling relation
$\sigma_{\ln\zeta}$	Intrinsic scatter in the SZ - mass scaling relation
A_X	Normalization of the X-ray - mass scaling relation
B_X	Mass evolution of the X-ray - mass scaling relation
C_X	Redshift evolution of the X-ray - mass scaling relation
$\sigma_{\ln Y_X}$	Intrinsic scatter in the X-ray - mass scaling relation
ρ_{ζ,Y_X}	Correlation coefficient for the intrinsic scatter in the scaling relations
Extension Parameters:	
Σm_ν	Sum of the neutrino masses
N_{eff}	Effective number of relativistic species
w	Dark energy equation of state parameter

Table 7.1 The parameters that are varied in the joint cosmological and scaling relation fit. The extension parameters are only varied where mentioned.

of nine scaling relation parameters in order to extract cosmological parameters. These parameters, along with the cosmological parameters we use, are shown in Table 7.1.

7.2 Cosmological Constraints

The cluster catalog provides constraints on cosmological parameters through several mechanisms. The number of clusters found, averaged over the surveyed volume, strongly depends on the amplitude of the matter power spectrum. Given the median redshift of $z \sim 0.5$, the parameter combination that is constrained is $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71 \text{ km/s/Mpc}}\right)^{-0.1}$. Note that this well-constrained parameter combination has a slightly weaker dependence on Ω_M than low redshift cluster probes. For instance, Vikhlinin et al. (2009) find the optimally constrained parameter combination to be $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.41}$.

In addition to total cluster counts, the scaling relations described in §4.3 and §7.1 depend differently on cosmological parameters. Specifically, while the mass estimate inferred from the ζ-M relation depends weakly on redshift, the mass estimate inferred from the $Y_{\rm X}$- M scaling relation depends strongly on the angular diameter distance to the cluster, and hence the redshift. The self-consistency conditions implicit to the likelihood expression described in §6.1 therefore provide some constraint on the angular diameter distance to each cluster for which *Chandra* data is available. These constraints will be published in Benson et al. (*in prep*).

Furthermore, the power law slope and shape of the mass function changes with mass scale. In principle, the observed counts therefore provide independent information on the normalization of the ζ-M relation. This is not an important effect in the constraints presented in this chapter, mainly due to the fact that we marginalize over a large range in scaling relation parameters, as listed in §4.3.

Finally, the redshift dependence of the cluster abundance contains information on the growth function, as well as a dependence on the cosmic volume surveyed. The $N_{\rm clust}$ clusters presented in this work provide a large enough statistical sample to signficantly constrain cosmological parameters by measuring the evolution of cluster abundance. Note however, that this effect is driven by the systematic uncertainty on the scaling relation, as parameterized with power-law redshift evolution of the observable-mass scaling relations. This is further discussed in the context of dark energy in §7.2.4.

7.2.1 ΛCDM

Cosmological probes other than cluster counts have been extremely successful in narrowing down the cosmological parameter space. As mentioned in Chapter 1.1, one commonly used parameterization that fits the majority of experimental data is the ΛCDM model. This model has six degrees of freedom: a comoving cosmological baryon density $\Omega_b h^2$, cold dark matter density $\Omega_c h^2$, angular scale to the sound hori-

zon θ_s and optical depth to reionization τ. The density fluctuations are parameterized using a power-law scalar primordial power spectrum with amplitude A_s and power law index $(n_s - 1)$.

The six-parameter ΛCDM model is not well constrained by cluster counts alone. The cluster likelihood is flat in the optical depth due to reionization τ as well as the primordial scalar spectrum power law index n_s, once an appropriate pivot point is chosen. Therefore, when discussing cluster constraints without the inclusion of CMB data, we fix these parameters to the PlanckXVI best-fit values. We choose to marginalize over $\Omega_b h^2 = 0.022 \pm 0.002$, a BBN prior on the baryon density from measurements of the abundances of deuterium (Kirkman et al., 2003). We find that the cluster likelihood is very flat over this range, meaning that the cluster constraints do not strongly depend on the details of this prior.

The remaining three degrees of freedom in the ΛCDM model that are relevant to the cluster likelihood are H_0, Ω_M and σ_8, though in exploring the likelihood surface we actually vary the base parameters $\Omega_b h^2$, $\Omega_c h^2$, $\ln 10^{10} A_s$ and $100\theta_{MC}$.

Using the likelihood algorithm presented in §6 and the ΛCDM cosmological model, we measure

$$\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71 \text{ km/s/Mpc}}\right)^{-0.1} = 0.74 \pm 0.02 \tag{7.2}$$

Compared to the *Planck* constraints (Planck Collaboration et al., 2013a) (with the low-ell polarization powerspectrum as measured by WMAP, denoted as WP), the cluster-based measurement of σ_8 is quite low. In order to visualize this directly in data space, we plot the number of SPT-selected clusters expected as a function of redshift, marginalizing over the *Planck*-allowed ΛCDM cosmology. We use the nominal scaling relation parameters from the AGN8.0 model discussed in 4.2.3. This is shown in Figure 7.1. The measured abundance in combination with the AGN8.0 scaling relation is clearly discrepant with the predicted abundance from *Planck* under a ΛCDM cosmology.

While this visualization in data space is enlightening, it is necessary to marginalize

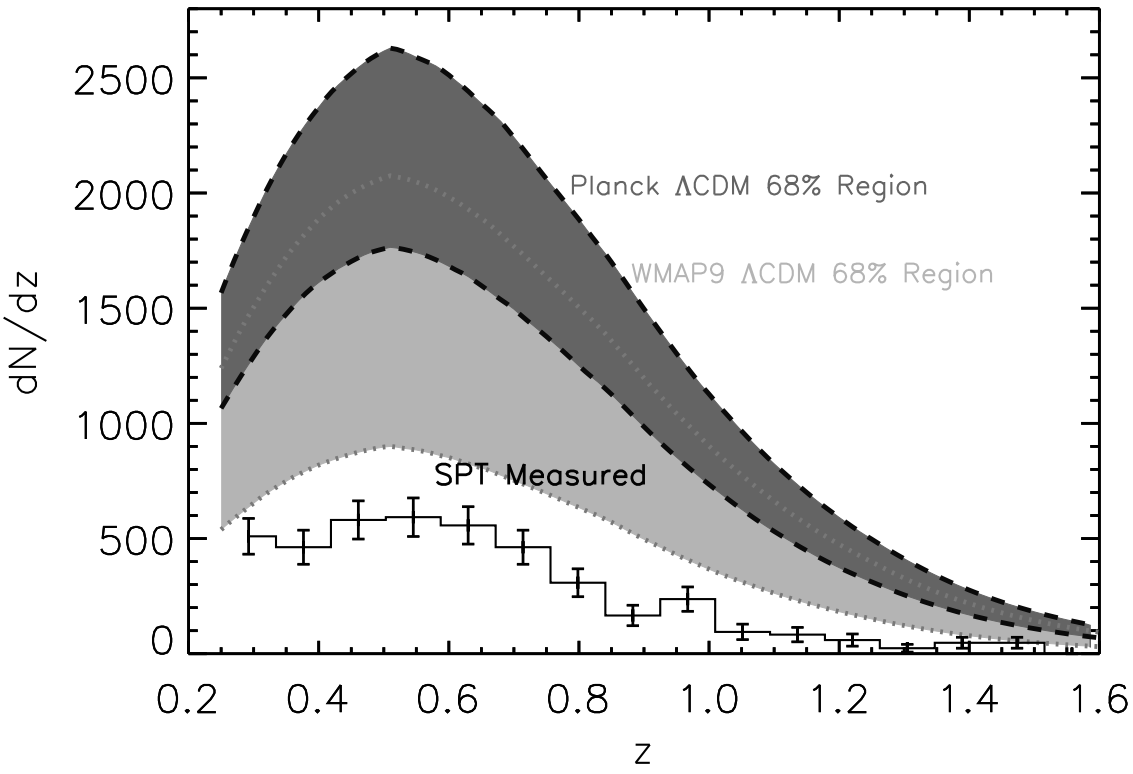

Figure 7.1 The cluster abundance as measured in the SPT-SZ 2500 square degree survey. The predicted abundances given either the *Planck* or WMAP9 constraints on the ΛCDM cosmology is shown in the gray bands. Here, we assume the nominal scaling relation parameters from the AGN8.0 simulation. Note the strong discrepancy between the measured and predicted cluster abundances when using the *Planck* constraints.

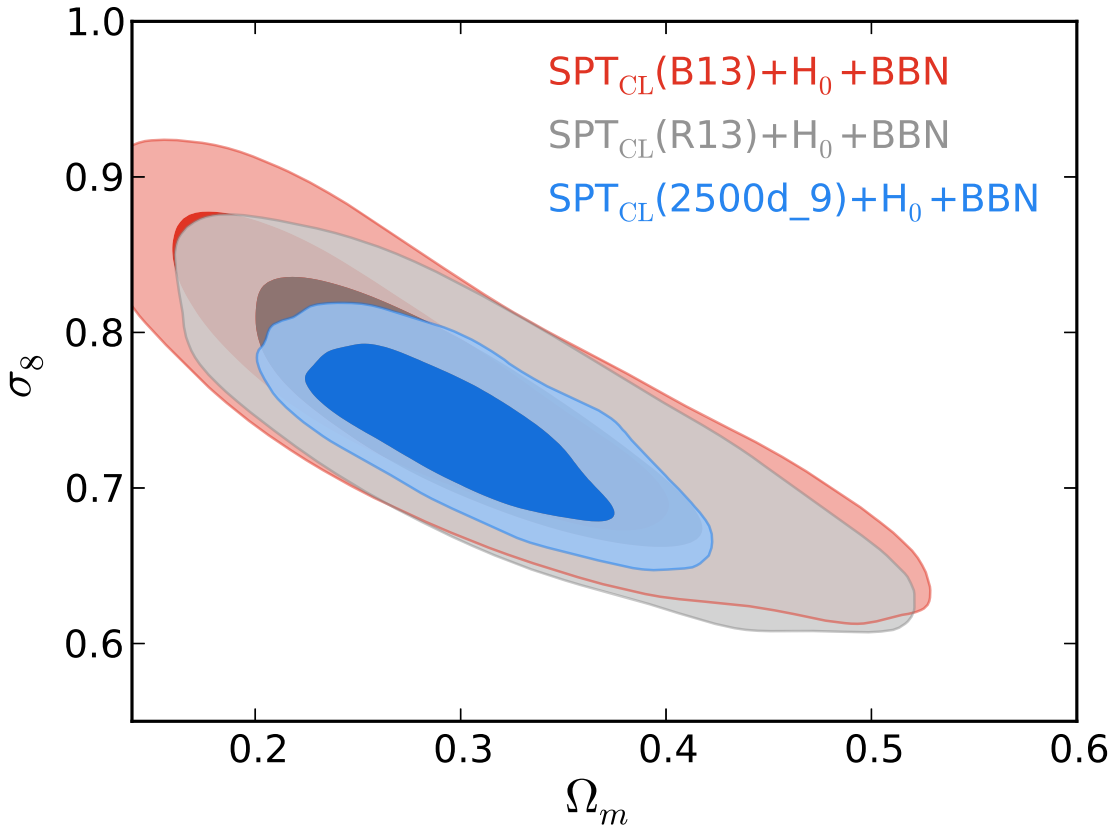

Figure 7.2 Comparison of the cosmological constraints presented in this work to those presented in previous SPT publications. B13 presented 18 SZ-selected clusters, 14 of which had X-ray data. R13 expanded the cluster catalog to 100 cluster candidates, keeping the amount of X-ray clusters to 14. This work presents a 375-cluster catalog, 83 of which have X-ray Y_X data. Note the σ_8-Ω_M contours become significantly smaller.

over scaling relation parameters in order to obtain proper parameter constraints. In order to visualize this constraint on the combination of three cosmological parameters directly, we can apply a prior on one of these parameters. Here, we choose the local measurement of $H_0 = 73.8 \pm 2.4$ km/s/Mpc by Riess et al. (2011) and show the resulting constraints in the σ_8-Ω_M plane in Figure 7.2. The same analysis is shown in previous SPT works, which were based on smaller subsamples of the cluster sample and X-ray follow-up data. We also compare this constraint to measurements of the primary CMB power spectrum in Figure 7.3. We find that datasets are consistent when comparing to WMAP9, though there is $\gtrsim 2\,\sigma$ tension between the *Planck*+WP and $\mathrm{SPT_{CL}}$+$\mathrm{H_0}$ datasets. The tension between *Planck* CMB powerspectrum and cluster-based results was also found by the *Planck* team in Planck Collaboration et al. (2013b).

The constraint on the combined parameter $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71\ \mathrm{km/s/Mpc}}\right)^{-0.1}$ is the primary ΛCDM result in this work. A direct comparison to CMB datasets is shown in Figure 7.4. Upon investigating swapping in and out the HST prior, the BAO prior and the conservative CMB-based prior on θ_s, we find that this cluster constraint is remarkably robust to these priors. We do not show constraints without any external datasets since the ΛCDM model cannot be constrained to physically reasonable values with clusters alone.

7.2.2 Constraints on the Neutrino Sector

In order to reconcile the *Planck* data with the SPT cluster abundance, we can extend the cosmological parameter space. A reasonable extension to ΛCDM is to allow neutrinos to have mass. Neutrino mass differences from terrestrial measurements imply a minimum sum of neutrino masses of $\Sigma m_\nu \approx 0.06$ eV. See Gonzalez-Garcia et al. (2012) for a review of recent terrestrial measurements of the neutrino sector. Neutrino mass has negligible impact at the surface of last scattering ($z \sim 1100$), since they remain relativistic then. At $z \lesssim 1$, where cluster abundance is measured with

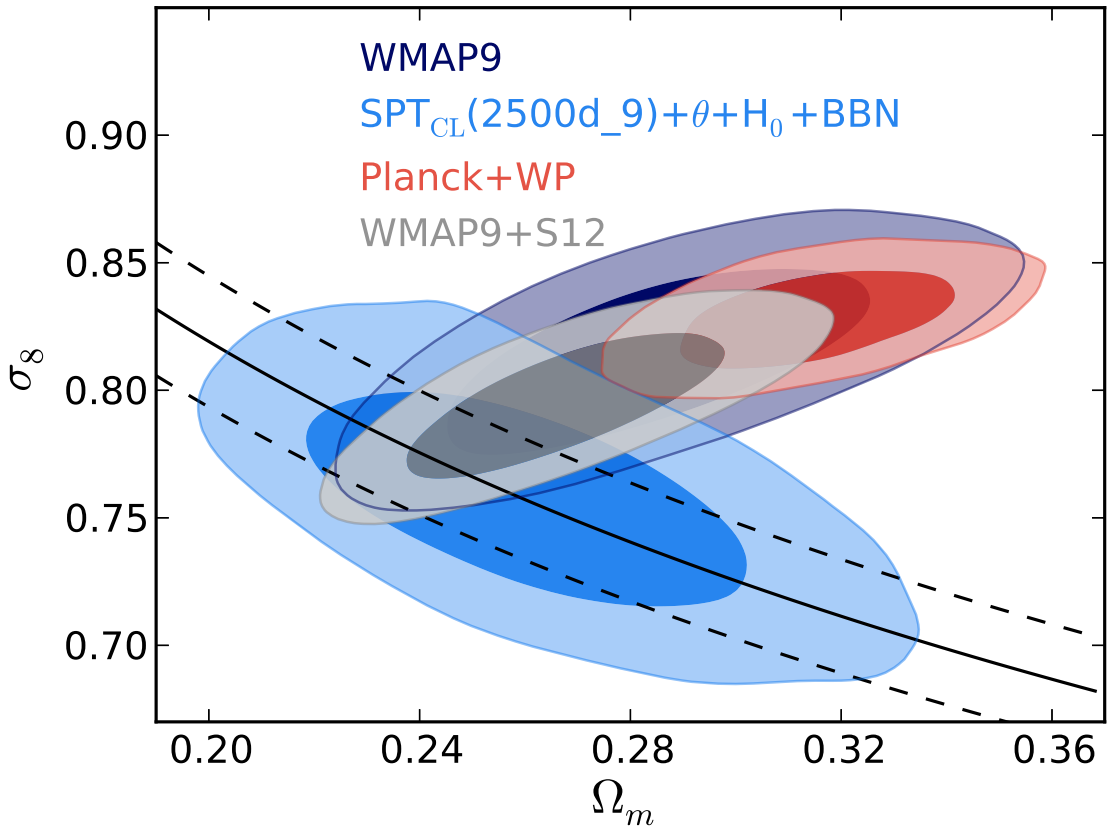

Figure 7.3 The 68% and 95% confidence intervals in the σ_8-Ω_M plane from the SPT cluster abundance measurement presented in this work are shown in blue. The WMAP7 bandpowers from Larson et al. (2011) combined with the SPT bandpowers from Story et al. (2012a) are shown in gray, where the red contours show the constraints implied from the *Planck* CMB power spectrum with WMAP polarization information (Planck Collaboration et al., 2013a). The black curves show the primary constraint coming from the cluster abundance measurement.

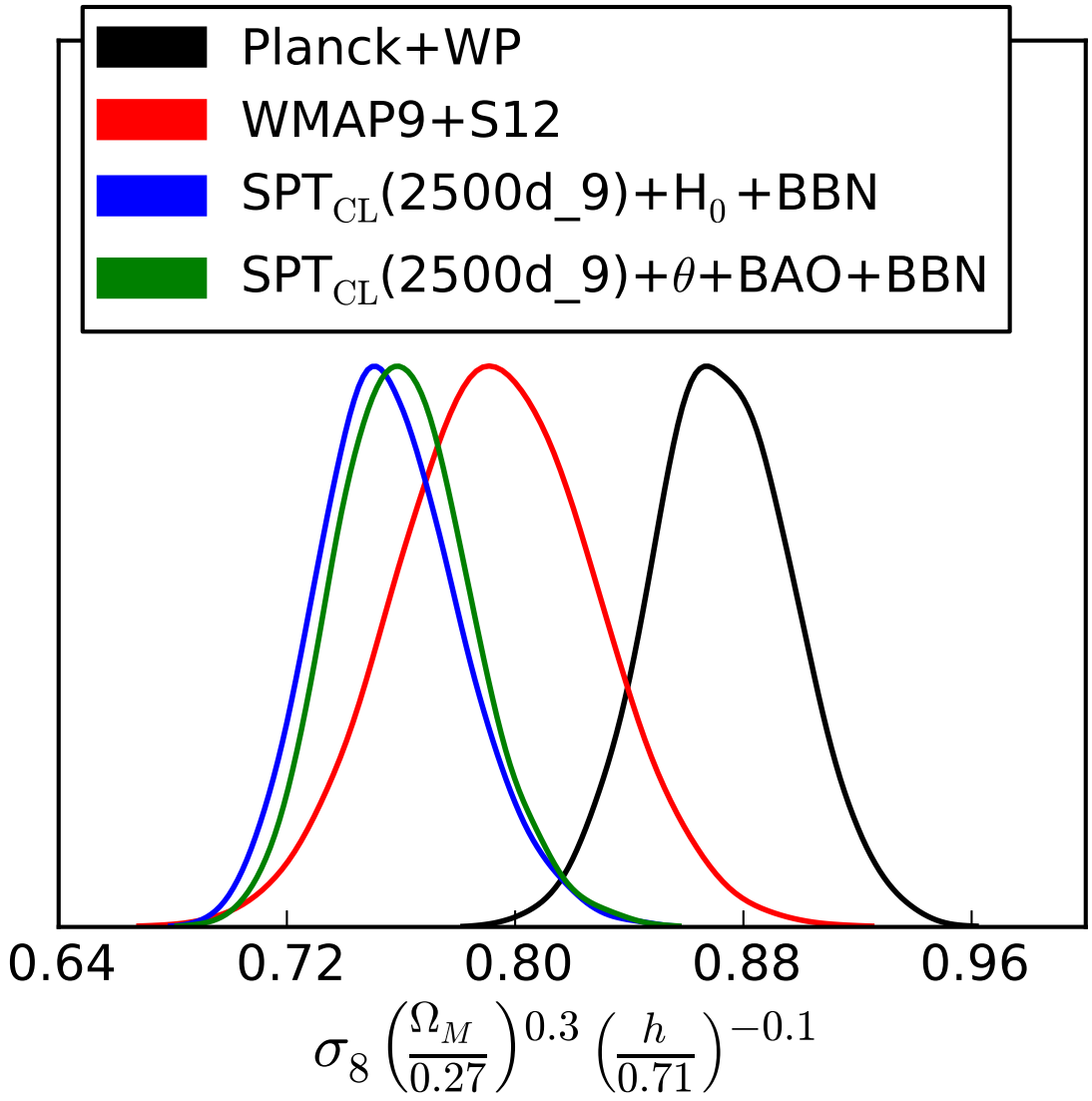

Figure 7.4 The one-dimensional primary constraint implied by the SPT cluster abundance measurements in concert with the measurements of the observable-to-mass scaling relations are shown in comparison to the CMB constraints given the ΛCDM model. Note that there is tension with the *Planck*+WP-implied parameters.

SPT, assigning a significant rest mass to neutrinos suppresses structure formation, since neutrinos act as warm dark matter.

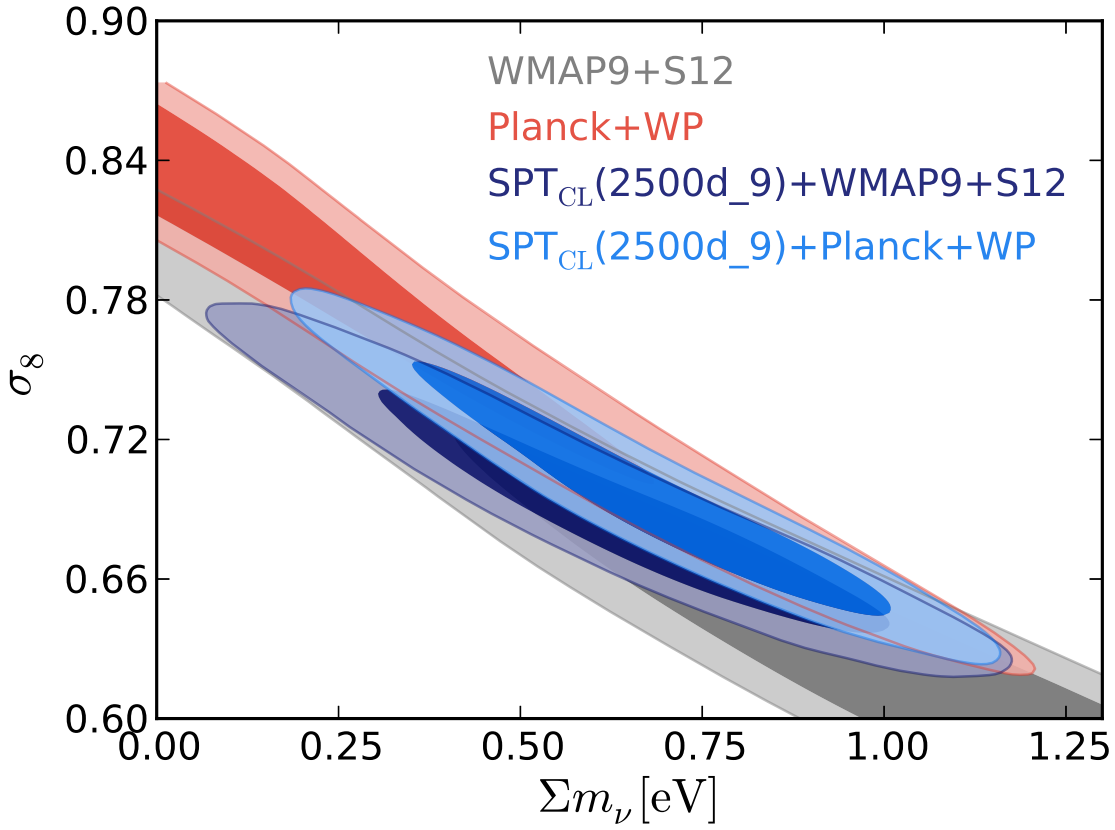

Figure 7.5 Cosmological constraints in the ΛCDM+Σm_ν cosmological model. The combination of the *Planck*+WP CMB powerspectrum data with the SPT cluster abundance measurement is shown in blue. This combination of datasets and cosmological parameters yields a 3.3σ detection of non-zero neutrino mass.

Figure 7.5 shows the *Planck*+WP constraints in the σ_8-Σm_ν plane, and the effect of combining with the cluster abundance measurement. This nominally implies a detection of non-zero neutrino mass at 3.3σ. However, the combination of the CMB powerspectrum with the SPT cluster data in this ΛCDM + Σm_ν cosmology does not agree with either direct H_0 measurements or measurements of the baryon acoustic scale r_{drag}/D_V at $z = 0.57$. We therefore disfavor this way of resolving the apparent *Planck*-SPT$_{\rm CL}$ discrepancy.

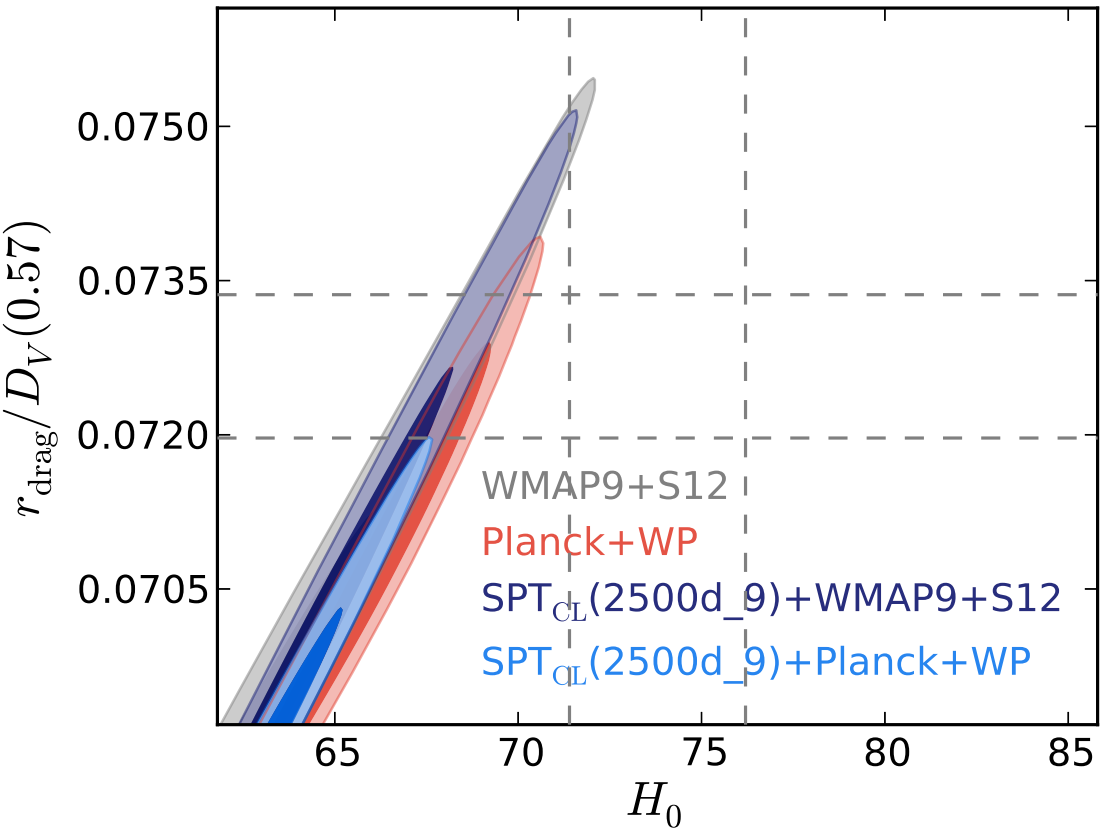

Figure 7.6 Cosmological constraints in the ΛCDM+Σm_ν cosmological model. The horizontal lines denote the $\pm$ 1σ constraints from the BAO DR11 CMASS constraints, whereas the direct H_0 measurement is shown as vertical lines. While the free neutrino mass parameter reduces tension between the cluster abundance information and *Planck*+WP, this comes at the cost of being discrepant with BAO and H_0 measurements.

One way to resolve the tension with the direct H_0 and BAO measurements is to free the effective number of relativistic species N_{eff} together with the sum of neutrino masses. This parameterization allows for new particle physics such as an additional sterile neutrino, as explored in Dvorkin et al. (2014), though this parameter need not be near an integer (e.g. Garcia-Cely et al. (2014)). When freeing N_{eff}, the joint *Planck*+WP+SPT$_{\rm CL}$ likelihood includes the parameter space allowed by direct measurements of H_0 and BAO, as shown in Figure 7.7. We show the likelihood contours in the N_{eff}-Σm_ν plane in Figure 7.8.

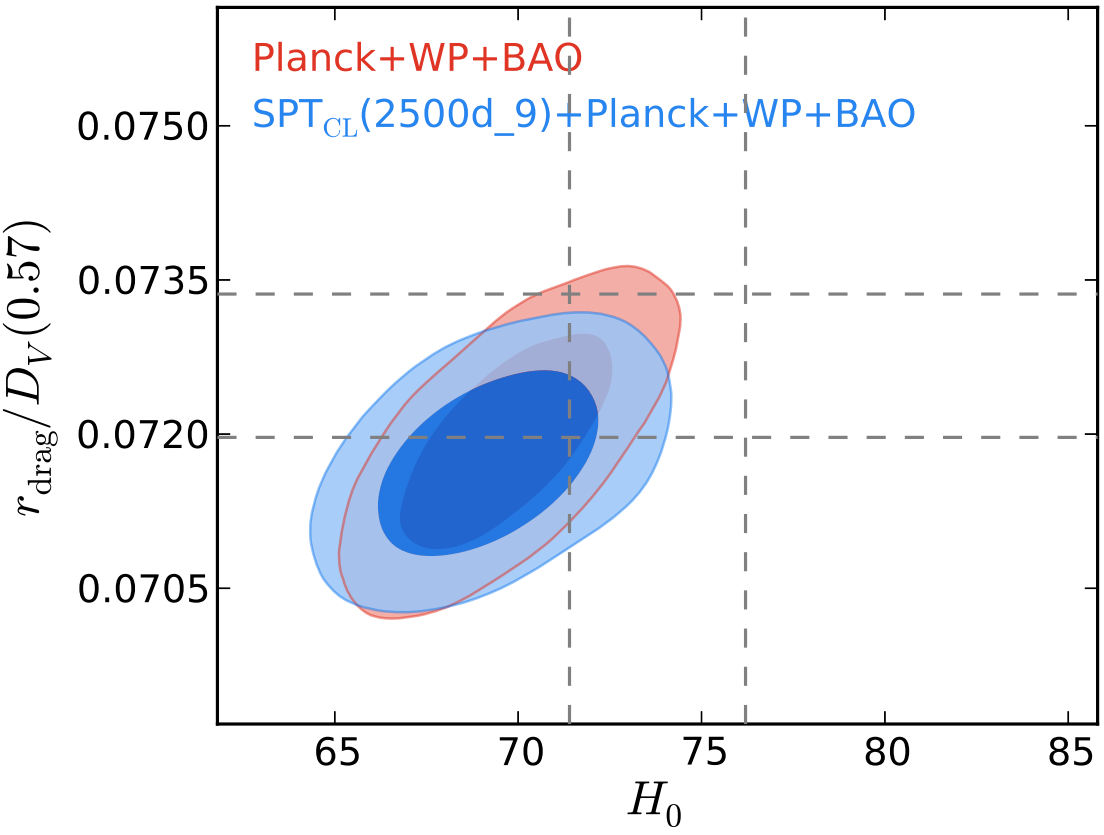

Figure 7.7 Cosmological constraints in the ΛCDM+Σm_ν+N_{eff} cosmological model. Note that the tension between SPT$_{\rm CL}$+*Planck* and BAO+H$_0$ is significantly reduced by allowing $N_{eff} > 3.046$.

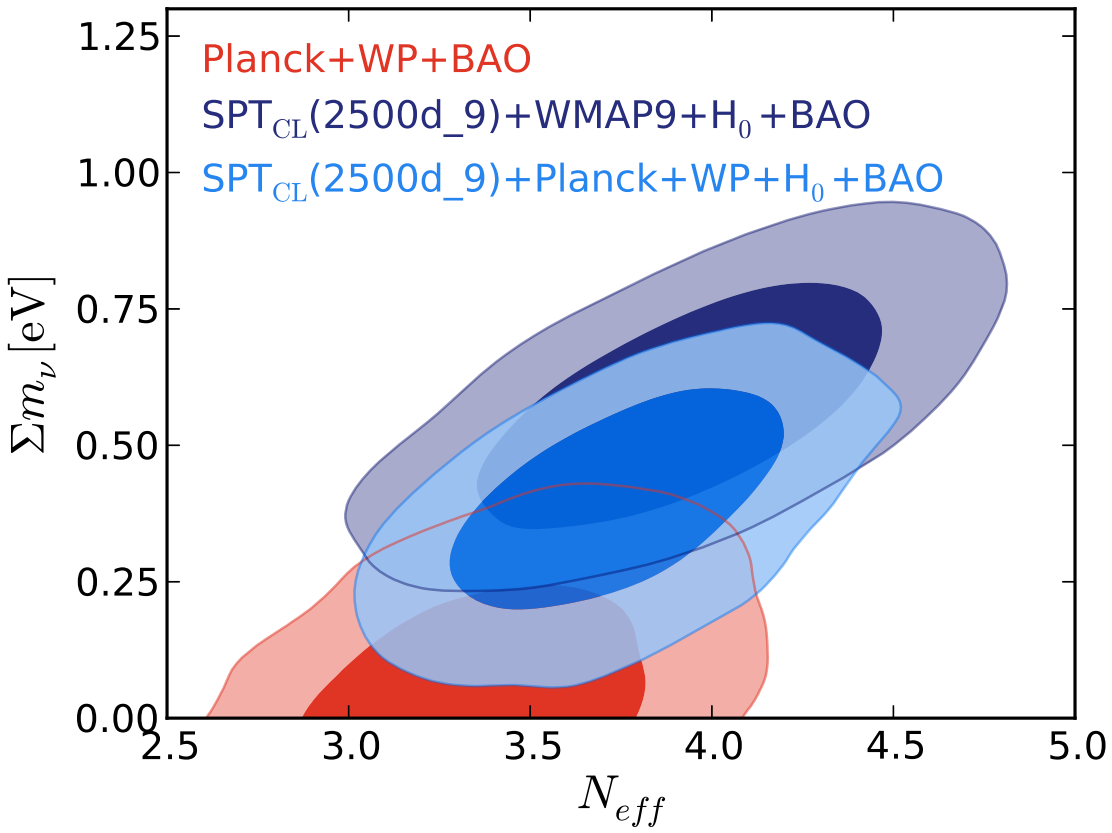

Figure 7.8 The constraints on the sum of neutrino masses Σm_ν and effective number of relativistic species N_{eff} are shown for various combinations of datasets.

7.2.3 Error Budget and Future Prospects

The most important systematic uncertainty driving the cosmological constraints from cluster abundance is knowledge of the overall mass scale. In this work, the parameter A_X captures this dominant systematic. Figure 7.9 shows the effect of doubling the assumed A_X uncertainty on the $\mathrm{SPT_{CL}}$+θ+BAO constraints. The nominal X-ray based assumed uncertainty on cluster mass scale is approximately 15% when averaged over the SPT redshift range. Doubling the assumed A_X uncertainty changes this to be $\sim 22\%$. This is comparable to the assumed prior on A_{SZ} of 30% in ζ. Therefore, the resulting uncertainty on cosmological parameters does not double when doubling the assumed value for the 1σ Gaussian prior width on A_X.

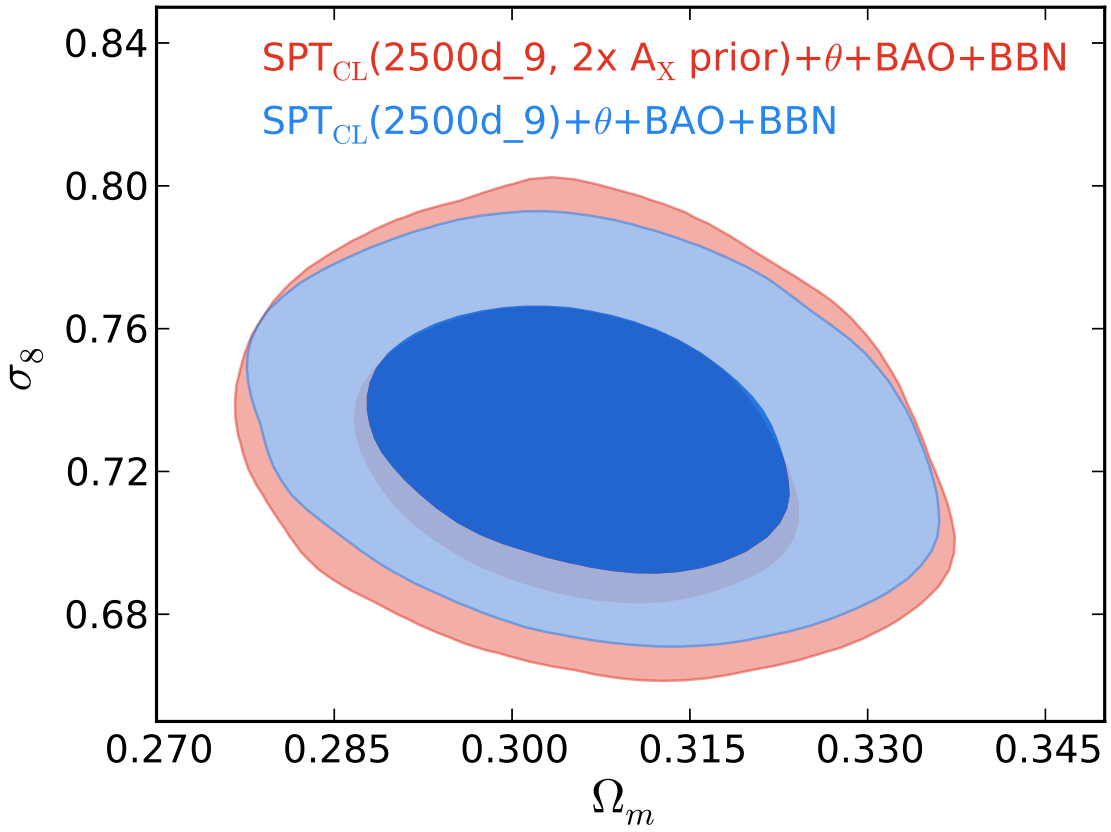

Figure 7.9 ΛCDM constraints when placing twice as conservative a prior on the normalization of the Y_X-M relation A_X. Note the constraints degrade only slightly.

A more cosmological prior-independent statement can be made by looking at how the $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71 \text{ km/s/Mpc}}\right)^{-0.1}$ constraint depends on A_X. This is shown in Figure 7.10. Note the degeneracy is strong, implying improvement in our knowledge of the overall mass scale of the SPT cluster sample would greatly improve the cluster-based cosmological constraints.

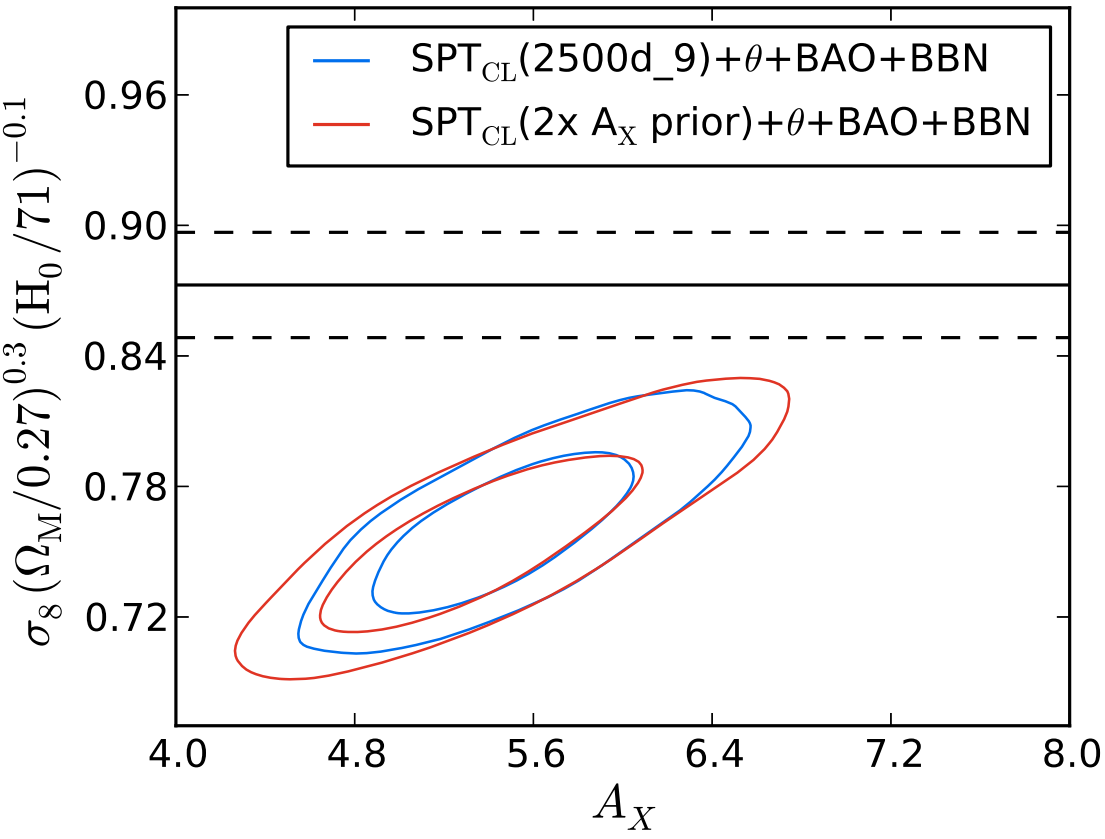

Figure 7.10 Constraints on $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71 \text{ km/s/Mpc}}\right)^{-0.1}$, the primary parameter measured using cluster abundance, is plotted against A_X. The blue contours shows the nominal constraints, where for the red contours we have increased the assumed uncertainty on the normalization of the Y_X-M relation by a factor of two. The black lines show the 68% confidence region for *Planck*+WP. Note that the tension is not significantly relieved by doubling the width of the prior.

Changes in Ω_M and σ_8 affect the matter power spectrum (and hence the cluster mass function) differently as a function of redshift. Therefore, measuring the cluster abundance in more than one redshift bin can break the σ_8-Ω_M degeneracy. However,

this redshift trend in abundance is degenerate with evolution in the observable-mass scaling relation. This can be seen in the C_{SZ} column of Figure 7.11.

This improvement in our knowledge of the mass scale of the SPT sample is expected soon. The SPT team has obtained high-fidelity weak lensing data for dozens of SPT-selected clusters. The analysis is currently ongoing and expected to result in significant improvements to the constraints presented in this dissertation.

7.2.4 wCDM

The local energy density of dark energy was assumed to be constant in 1.1.3. This assumption has little theoretical foundation, so testing it is important. We follow the standard notation and substitute an arbitrary dark energy component in 1.8 to obtain

$$H(z) = H_0\sqrt{\Omega_\gamma(1+z)^4 + \Omega_M(1+z)^3 + \Omega_\Lambda(1+z)^{3(1+w)}}\ . \tag{7.3}$$

The cluster abundance is sensitive to the dark energy equation of state parameter w through several mechanisms. First, the epoch when the dark energy density starts to dominate the global energy density ($\Omega_M(z) \sim \Omega_\Lambda(z)$) is affected. The growth of structure is also altered by dark energy. Finally, the angular diameter distance as a function of redshift changes, which affects the X-ray and SZ observables differently. Since we require clusters to only have one true mass in the likelihood expression (Equation 6.4), having independent information on the absolute calibration of the X-ray and SZ mass proxies provides information on dark energy.

In this work, the prior knowledge on the normalization of the SZ-mass relation is taken very conservatively, such that dark energy is not significantly constrained by the angular diameter distance. In addition, the dark energy equation of state parameter w is highly degenerate with both H_0 and the redshift evolution in the SZ-mass scaling relation. Without adding more cosmological datasets, we are therefore unable to extract meaningful constraints when extending the ΛCDM model by the

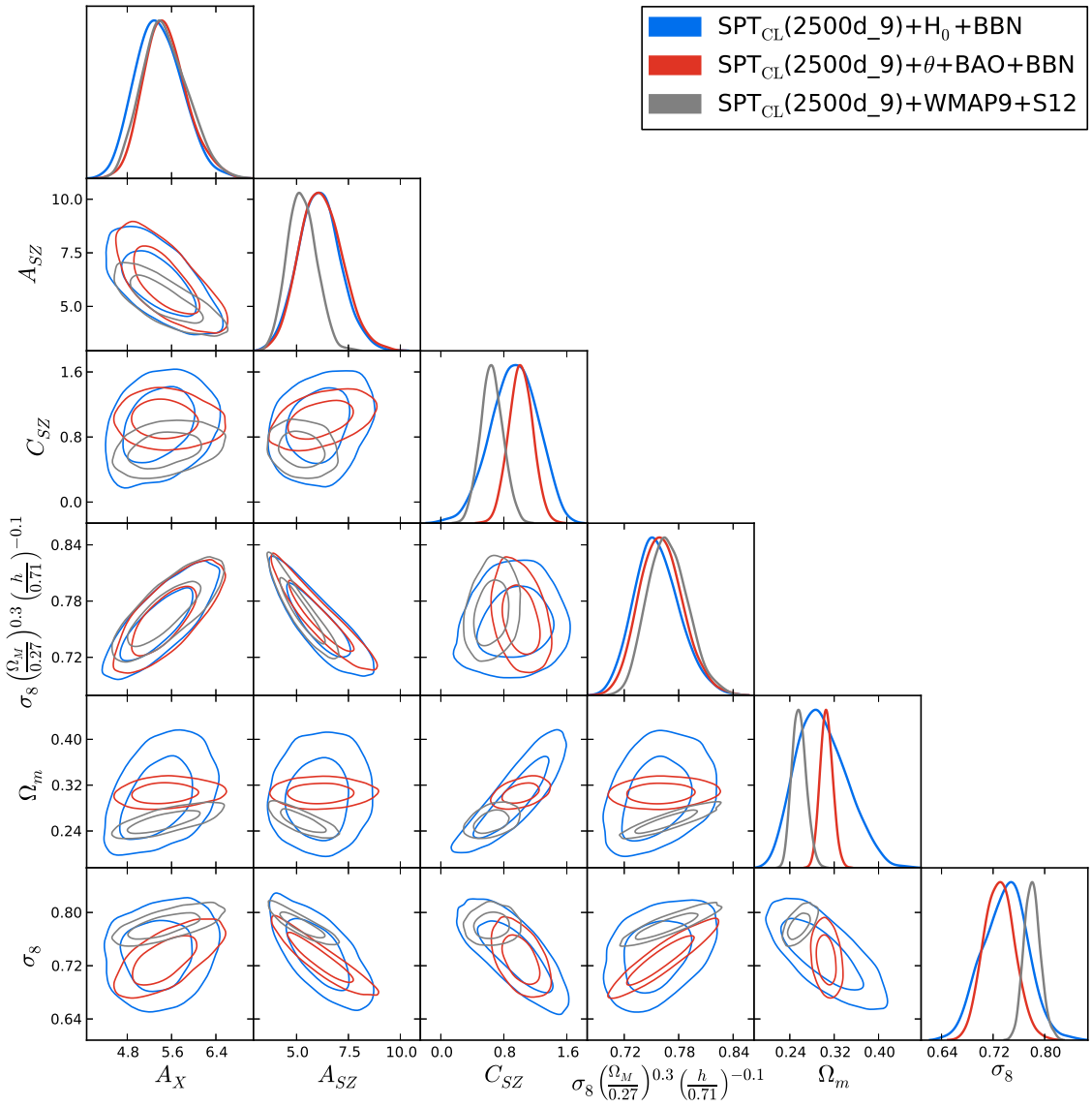

Figure 7.11 The degeneracies between key scaling relation parameters and cosmological parameters are shown. The constraints on $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71 \text{ km/s/Mpc}}\right)^{-0.1}$ are relatively independent of the additional cosmological data assumed. However, the constraints on σ_8 and Ω_M differ significantly depending on the choice of external dataset. Note that improved knowledge of the redshift evolution scaling relation parameter C_{SZ} will break the σ_8-Ω_M degeneracy.

parameter w. However, to highlight the potential constraining power of the cluster abundance measurement, we adopt the nominal values of the mass-observable scaling relations and find the cosmological parameters to be strongly constrained, as shown in Figure 7.12. We find $w = -1.010 \pm 0.038$ given this assumption.

In this chapter, we have presented cosmological constraints using cluster abundance measurements from the SPT-SZ survey. When applying conservative scaling relation priors, we find significant tension with *Planck* measurements of the primary CMB anisotropy, in a ΛCDM cosmology. The neutrino sector can help resolve this tension. Allowing the sum of the neutrino masses to be ~ 0.5 eV rather than the assumed 0.06 eV in ΛCDM brings the cluster abundance measurement in agreement with *Planck*. However, the resulting parameter space has significant tension with measurements of BAO and direct measurements of H_0. Allowing the effective number of relativistic species N_{eff} to be a free parameter brings these datasets into agreement. Replacing the *Planck* CMB data with the state-of-the-art results from before the *Planck* publications (WMAP9+S12) does not show significant tension with cluster abundance. However, when combined with local probes, the preference for neutrino physics beyond ΛCDM remains.

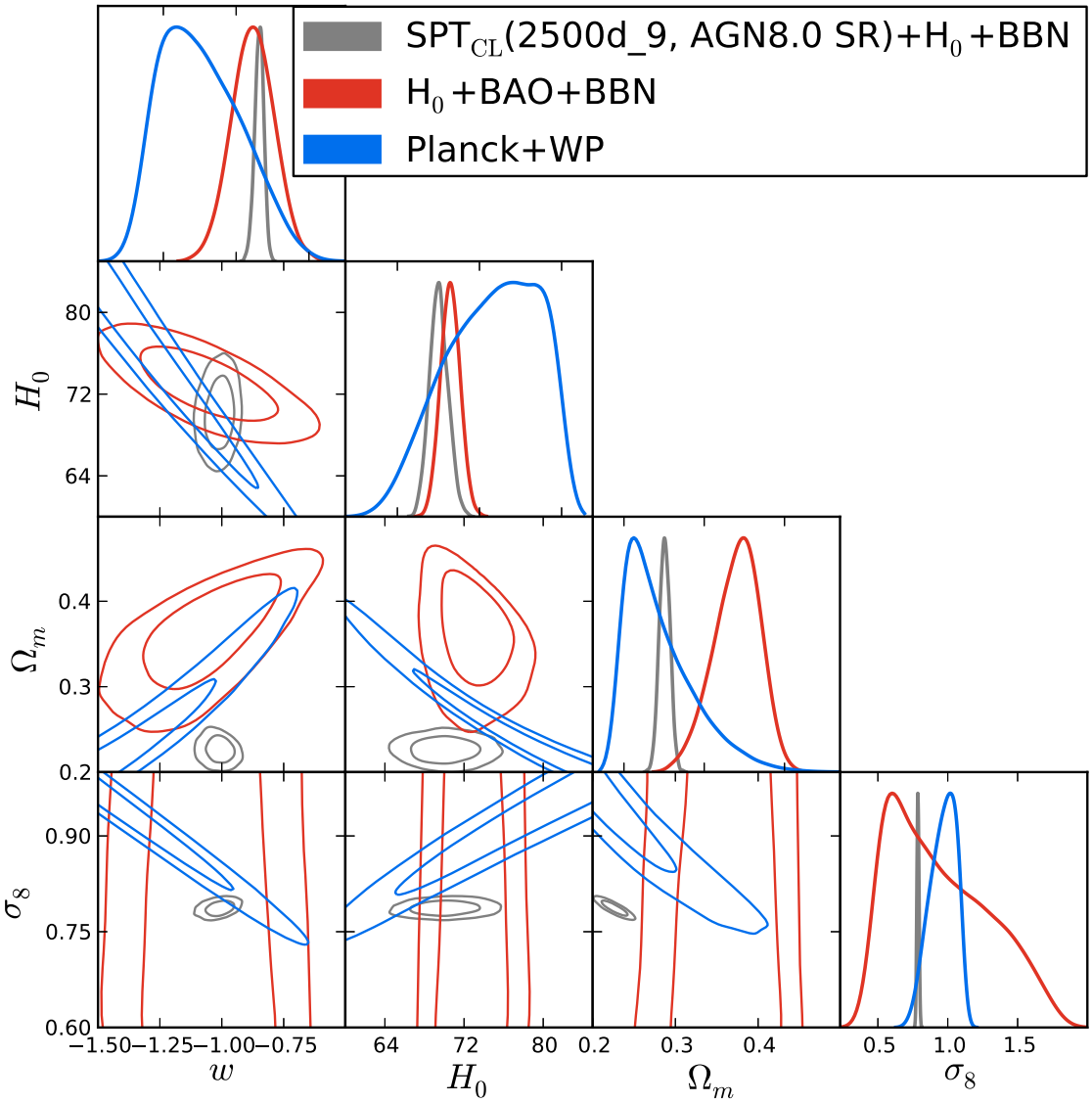

Figure 7.12 Constraints on the wCDM model when assuming the nominal values for the mass-observable relation. The cluster abundance measurement is able to constrain the five-parameter model with higher precision than either the H_0+BAO or *Planck*+WP datasets.

8

Conclusion

In this dissertation, we have presented several results. New results require new technological advances. We have developed DAN, a digital algorithm for providing feedback to the SQUID amplifier at relevant frequencies, enhancing the capability of the frequency domain multiplexed readout to read more channels over a single wire. DAN is fully implemented, tested, and is used in the operation of the South Pole Telescope. In addition, both DEVB and its resistance-feedback counterpart are digital algorithms to enhance bolometer stability, linearity, and ease of calibration. These algorithms are still under development.

We also presented constraints on cosmological parameters from galaxy cluster number counts, as measured with the SPT. We designed and implemented a computationally efficient method of obtaining the cluster likelihood in the presence of an arbitrary number of mass-observables, while allowing each mass proxy to follow an arbitrary dependence on cosmological parameters, have an arbitrary uncorrelated probability distribution in addition to a log-normal scatter with true mass. The log-normal components of the scatter in the observable-mass relations can be arbitrarily correlated between mass proxies, and these correlation coefficients can be marginalized over.

Using the likelihood code we developed, we used the 2500 square-degree SPT-SZ cluster catalog, which consists of 375 galaxy cluster candidates above the selection threshold of $z = 0.25$ and a detection significance exceeding five. All but 18 clusters have measured optical/near infrared redshifts, where we show that we expect $\sim$18.5 cluster candidates to be false from simulations. The cluster catalog is com-

bined with 82 observations from the *Chandra* X-ray telescope, providing a follow-up mass proxy. We find that, given the mass-observable relations, finding only 375 clusters in this area, with our noise level, is surprising. Using the ΛCDM model as constrained by data from the *Planck* satellite, we might rather expect to have found ~ 1000 clusters. This manifests in tension, most prominently in the parameter $\sigma_8 \left(\frac{\Omega_M}{0.27}\right)^{0.3} \left(\frac{H_0}{71 \text{ km/s/Mpc}}\right)^{-0.1}$. We show that the tension with the *Planck* data can be resolved by allowing the sum of neutrino masses to be ~ 0.5 eV. This comes at the cost of tension with the BAO and H_0 datasets, which in turn can be resolved by extending the model with a free parameter for the effective number of relativistic species N_{eff}.

Ongoing observations and analysis of weak gravitational lensing data should provide more insight in the dominant systematic for this work: the cluster mass scale. With this data in hand, we should be able to reduce the uncertainties on the cosmological parameters discussed in this work. The 2500d SPT-SZ cluster abundance measurement is a powerful cosmological dataset. We show that, once the observable-mass scaling relation parameter uncertainties are known, the sample is able to place tight constraints on the wCDM model, measuring the dark energy equation of state parameter w to $\sim 3.8\%$.

A

Stability criterion for series impedance to the TES

The left hand panel of Figure 2.6 shows that the reason for the TES latching due to series impedance is well understood. The total impedance is seen to have a zero derivative with respect to the total voltage, so further lowering the total bias voltage causes the voltage across the TES to fall in a runaway manner. This instability occurs when

$$\frac{d|z_s + R_{\mathrm{TES}}|}{dV_{\mathrm{OUT}}} = 0\,. \tag{A.1}$$

We substitute Equation 2.3 which describes the model, and assume that R_{TES} is a monotonically decreasing function of V_{TES}. We find that this condition is met when

$$\frac{d\log V_{\mathrm{TES}}}{d\log R_{\mathrm{TES}}} = \frac{|z_s|}{|z_s + R_{\mathrm{TES}}|} \tag{A.2}$$

and conclude that the TES will latch due to series impedance when the fractional stray impedance exceeds the logarithmic slope of the TES transition.

B

Glossary

The following terms were used throughout, and are again defined here for your convenience:

- **AGN8.0:** The nominal hydrodynamic SZ sky simulation using feedback from Active Galactic Nuclei, version 8.0.
- **CMB:** Cosmic Microwave Background. The radiation emitted when the universe cooled sufficiently for neutral hydrogen to form.
- **DAN:** Digital Active Nulling. A method in frequency multiplexed readout to keep the SQUID amplifier nuller at bolometer resonant frequencies.
- **DEVB:** Digitally Enhanced Voltage Bias. A method in frequency multiplexed bolometer readout where the bias carrier voltage amplitude is modulated digitally in order to keep the voltage drop across the TES constant, despite parasitic stray impedance.
- **ETF:** Electrothermal Feedback. The process whereby a voltage-biased TES bolometer is resistant to changes in incident optical power due to strong negative feedback in the joint electrical-thermal system.
- **FDR:** False Detection Rate. The expected number of false positive cluster candidates per unit area.

- **FRW metric:** Friedman-Robertson-Walker metric. An exact solution to the Einstein equation describing a homogeneous, isotropic universe that may be expanding or contracting.

- **FSF:** Field Scaling Factor. The numerical correction factor that is used to compensate for the slightly different noise levels of the 19 SPT-SZ survey fields.

- H_0**:** The present-day value of the Hubble constant. Measured in km/s/Mpc, this encodes the rate at which objects in the "Hubble flow" are receding as a function of their distance.

- **ΛCDM:** The cosmological model where the universe is assumed to be spatially flat, and consist of a cosmological constant, cold dark matter, and baryons. The abundances of cold dark matter and baryons, in addition to the amplitude and power-law slope of the primordial density fluctuations, and the angular scale to the sound horizon are the five parameters that describe the model. A sixth parameter, the optical depth to reionization, may be added when CMB data is considered.

- **MCMC:** Markov chain Monte Carlo. A popular numerical algorithm for obtaining a sample of points drawn from an unknown probability distribution. The method is computationally inexpensive, particularly for high dimensional likelihood surfaces.

- **Noise PSD:** Noise Power Spectral Density. Used in this work for the two-dimensional expected Fourier-domain amplitudes which model the noise in SPT maps.

- Ω_M**:** The present-day value of the sum of energy density of all pressureless components. Typically Ω_M is modelled as a baryonic component (value around 5%) and cold dark matter (value around 30%).

- ***Planck*:** Modern CMB experiment on a satellite platform. The data used in this paper is from the first year-release (March 2013).
- **RCW38:** A mm-wave bright galactic H-II source used for SPT calibration.
- σ_8**:** The present-day amplitude of the matter power spectrum on 8 Mpc scales.
- **SPT:** South Pole Telescope. The 10-m mm-wave telescope located at the geographic south pole which is at the focus of this dissertation.
- **SPT-SZ:** The first generation camera deployed on the SPT.
- **SPTpol:** The second generation camera deployed on the SPT. The number of detectors approximately doubled relative to SPT-SZ and the experiment is polarization-sensitive.
- **SQUID:** Superconducting Quantum Interference Device. An extremely sensitive magnetometer based on superconducting loops and Josephson junctions. Used as a low-input impedance low-noise amplifier.
- **SZ:** Sunyaev-Zel'dovich. The SZ effect described the upscattering of CMB photons as they travel through a hot plasma and is used to detect and study galaxy clusters.
- **TES:** Transition Edge Sensor. A superconducor held at the edge of superconductivity, where small changes in temperature lead to large changes in resistance. This results in a very sensitive thermometer.
- **WMAP:** Wilkinson Microwave Anisotropy Probe. Satellite designed to measure the properties of the CMB. Took data in the period 2001-2010.
- ξ**:** The symbol denoting the significance of an SPT cluster candidate using the matched filter technique from §3.3.

- ζ**:** The unbiased significance of an SPT cluster candidate. This is typically measured using the matched filter in the absence of non-SZ sources of mm-wave anisotropy or noise.

Bibliography

M. Arnaud, G. W. Pratt, R. Piffaretti, H. Böhringer, J. H. Croston, and E. Pointecouteau. The universal galaxy cluster pressure profile from a representative sample of nearby systems (REXCESS) and the Y_{SZ} - M_{500} relation. *A&A*, 517:A92+, July 2010.

B. A. Benson, T. de Haan, J. P. Dudley, C. L. Reichardt, K. A. Aird, K. Andersson, R. Armstrong, M. L. N. Ashby, M. Bautz, M. Bayliss, et al. Cosmological Constraints from Sunyaev-Zel'dovich-selected Clusters with X-Ray Observations in the First 178 deg^2 of the South Pole Telescope Survey. *ApJ*, 763:147, February 2013.

E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *A&AS*, 117: 393–404, June 1996.

S. Bhattacharya, K. Heitmann, M. White, Z. Lukić, C. Wagner, and S. Habib. Mass Function Predictions Beyond ΛCDM. *ApJ*, 732:122–+, May 2011.

L. E. Bleem, B. Stalder, T. de Haan, K. A. Aird, S. W. Allen, D. E. Applegate, M. L. N. Ashby, M. Bautz, M. Bayliss, B. A. Benson, et al. Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-square-degree SPT-SZ survey. *ArXiv e-prints*, September 2014.

P. Bode, J. P. Ostriker, J. Weller, and L. Shaw. Accurate Realizations of the Ionized Gas in Galaxy Clusters: Calibrating Feedback. *ApJ*, 663:139–149, July 2007.

M. Brodwin, J. Ruel, P. A. R. Ade, K. A. Aird, K. Andersson, M. L. N. Ashby, M. Bautz, G. Bazin, B. A. Benson, L. E. Bleem, et al. SPT-CL J0546-5345: A Massive z>1 Galaxy Cluster Selected Via the Sunyaev-Zel'dovich Effect with the South Pole Telescope. *ApJ*, 721:90–97, September 2010.

J. E. Carlstrom, G. P. Holder, and E. D. Reese. Cosmology with the Sunyaev-Zel'dovich Effect. *ARA&A*, 40:643–680, 2002.

K. Coble, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, C. R. Contaldi, B. P. Crill, P. de Bernardis, P. Farese, et al. Observations of Galactic and Extra-galactic Sources From the BOOMERANG and SEST Telescopes. *ArXiv Astrophysics e-prints*, January 2003. astro-ph/0301599.

T. M. Crawford, K. K. Schaffer, S. Bhattacharya, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. Cho, A. T. Crites, et al. A measurement of

the secondary-CMB and millimeter-wave-foreground bispectrum using 800 square degrees of South Pole Telescope data. *submitted to ApJ*, March 2013.

T. de Haan, G. Smecher, and M. Dobbs. Improved performance of TES bolometers using digital feedback. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8452 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, September 2012.

G. De Zotti, R. Ricci, D. Mesa, L. Silva, P. Mazzotta, L. Toffolatti, and J. González-Nuevo. Predictions for high-frequency radio surveys of extragalactic sources. *A&A*, 431:893–903, March 2005.

R. den Hartog, J. Beyer, D. Boersma, M. Bruijn, L. Gottardi, H. Hoevers, R. Hou, M. Kiviranta, P. de Korte, J. van der Kuur, et al. Frequency Domain Multiplexed Readout of TES Detector Arrays With Baseband Feedback. *IEEE Transactions on Applied Superconductivity*, 21:289–293, June 2011.

R. den Hartog, M. D. Audley, J. Beyer, D. Boersma, M. Bruijn, L. Gottardi, H. Hoevers, R. Hou, G. Keizer, P. Khosropanah, et al. Low-Noise Readout of TES Detectors with Baseband Feedback Frequency Domain Multiplexing. *Journal of Low Temperature Physics*, page 266, February 2012.

M. Dobbs, E. Bissonnette, and H. Spieler. Digital Frequency Domain Multiplexer for Millimeter-Wavelength Telescopes. *IEEE Transactions on Nuclear Science*, 55: 21–26, 2008.

M. Dobbs, F. Aubin, T. de Haan, S. Hanany, N. Harrington, W. Holzapfel, J. Hubmayr, A. Lee, M. Lueker, K. MacDermid, et al. Digital Frequency Multiplexer for TES Detectors - Path to Flight. *Journal of Low Temperature Physics*, page 163, January 2012.

M. A. Dobbs, M. Lueker, K. A. Aird, A. N. Bender, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H.-M. Cho, J. Clarke, et al. Frequency Multiplexed SQUID Readout of Large Bolometer Arrays for Cosmic Microwave Background Measurements. *ArXiv e-prints*, December 2011.

C. Dvorkin, M. Wyman, D. H. Rudd, and W. Hu. Neutrinos help reconcile Planck measurements with both Early and Local Universe. *ArXiv e-prints*, March 2014.

R. J. Foley, K. Andersson, G. Bazin, T. de Haan, J. Ruel, P. A. R. Ade, K. A. Aird, R. Armstrong, M. L. N. Ashby, M. Bautz, et al. Discovery and Cosmological Implications of SPT-CL J2106-5844, the Most Massive Known Cluster at z>1. *ApJ*, 731:86–+, April 2011.

C. Garcia-Cely, A. Ibarra, and E. Molinaro. Cosmological and astrophysical signatures of dark matter annihilations into pseudo-Goldstone bosons. *Journal of Cosmology and Astroparticle Physics*, 2:032, February 2014.

M. C. Gonzalez-Garcia, M. Maltoni, J. Salvado, and T. Schwetz. Global fit to three neutrino mixing: critical look at present precision. *Journal of High Energy Physics*, 12:123, December 2012.

M. G. Haehnelt and M. Tegmark. Using the Kinematic Sunyaev-Zeldovich effect to determine the peculiar velocities of clusters of galaxies. *MNRAS*, 279:545+, March 1996.

I. Harrison and S. Hotchkiss. A consistent approach to falsifying ΛCDM with rare galaxy clusters. *Journal of Cosmology and Astroparticle Physics*, 7:022, July 2013.

D. Herranz, J. L. Sanz, R. B. Barreiro, and E. Martínez-González. Scale-adaptive Filters for the Detection/Separation of Compact Sources. *ApJ*, 580:610–625, November 2002a.

D. Herranz, J. L. Sanz, M. P. Hobson, R. B. Barreiro, J. M. Diego, E. Martínez-González, and A. N. Lasenby. Filtering techniques for the detection of Sunyaev-Zel'dovich clusters <br/> in multifrequency maps. *MNRAS*, 336:1057–1068, November 2002b.

K. D. Irwin. An application of electrothermal feedback for high resolution cryogenic particle detection. *Applied Physics Letters*, 66(15):1998–2000, 1995. URL http://scitation.aip.org/content/aip/journal/apl/66/15/10.1063/1.113674.

A. Jenkins, C.S̃. Frenk, S. D.M̃. White, J.M̃. Colberg, S. Cole, A.Ẽ. Evrard, H. M.P̃. Couchman, and N. Yoshida. The mass function of dark matter haloes. *MNRAS*, 321:372–384, February 2001.

D. Kirkman, D. Tytler, N. Suzuki, J. M. O'Meara, and D. Lubin. The Cosmological Baryon Density from the Deuterium-to-Hydrogen Ratio in QSO Absorption Systems: D/H toward Q1243+3047. *ApJS*, 149:1–28, November 2003.

D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, M. R. Nolta, C. L. Bennett, B. Gold, M. Halpern, R. S. Hill, N. Jarosik, et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-derived Parameters. *ApJS*, 192:16–+, February 2011.

A. M. C. Le Brun, I. G. McCarthy, J. Schaye, and T. J. Ponman. Towards a realistic population of simulated galaxy groups and clusters. *ArXiv e-prints*, December 2013.

A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538:473–476, August 2000.

M. Lueker, B. A. Benson, C. L. Chang, H.-M. Cho, M. Dobbs, W. L. Holzapfel, T. Lanting, A. T. Lee, J. Mehl, T. Plagge, et al. Thermal Design and Characterization of Transition-Edge Sensor (TES) Bolometers for Frequency-Domain Multiplexing. *IEEE Transactions on Applied Superconductivity*, 19:496–500, June 2009.

M. Lueker, C. L. Reichardt, K. K. Schaffer, O. Zahn, P. A. R. Ade, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, et al. Measurements of Secondary Cosmic Microwave Background Anisotropies with the South Pole Telescope. *ApJ*, 719:1045–1066, August 2010.

Martin Lueker. *Measurements of Secondary Cosmic Microwave Background Anisotropies with the South Pole Telescope*. April 2011.

I. G. McCarthy, A. M. C. Le Brun, J. Schaye, and G. P. Holder. The thermal Sunyaev Zel'dovich effect power spectrum in light of Planck. *ArXiv e-prints*, December 2013.

M. McDonald, B. A. Benson, A. Vikhlinin, B. Stalder, L. E. Bleem, H. W. Lin, K. A. Aird, M. L. N. Ashby, M. W. Bautz, M. Bayliss, et al. The Growth of Cool Cores and Evolution of Cooling Properties in a Sample of 83 Galaxy Clusters at $0.3 < z < 1.2$ Selected from the SPT-SZ Survey. *Submitted to ApJ*, May 2013.

J.-B. Melin, J. G. Bartlett, and J. Delabrouille. Catalog extraction in SZ cluster surveys: a matched filter approach. *A&A*, 459:341–352, November 2006.

M. J. Mortonson, W. Hu, and D. Huterer. Simultaneous falsification of ΛCDM and quintessence with massive, distant clusters. *Phys. Rev. D*, 83(2):023015–+, January 2011.

T. Murphy, E. M. Sadler, R. D. Ekers, M. Massardi, P. J. Hancock, E. Mahony, R. Ricci, S. Burke-Spolaor, M. Calabretta, R. Chhetri, et al. The Australia Telescope 20 GHz Survey: the source catalogue. *MNRAS*, 402:2403–2423, March 2010.

D. Nagai, A. V. Kravtsov, and A. Vikhlinin. Effects of Galaxy Formation on Thermodynamics of the Intracluster Medium. *ApJ*, 668:1–14, October 2007.

J. P. Ostriker, P. Bode, and A. Babul. A Simple and Accurate Model for Intracluster Gas. *ApJ*, 634:964–976, December 2005.

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, et al. Planck 2013 results. XVI. Cosmological parameters. *ArXiv e-prints*, March 2013a. Paper XVI.

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, et al. Planck 2013 results. XX. Cosmology from Sunyaev-Zeldovich cluster counts. *ArXiv e-prints*, March 2013b.

Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, et al. Planck 2013 results. VIII. HFI photometric calibration and mapmaking. *ArXiv e-prints*, March 2013c.

C. L. Reichardt, L. Shaw, O. Zahn, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, et al. A measurement of secondary cosmic microwave background anisotropies with two years of South Pole Telescope observations. *ApJ*, 755:70, November 2012.

C. L. Reichardt, B. Stalder, L. E. Bleem, T. E. Montroy, K. A. Aird, K. Andersson, R. Armstrong, M. L. N. Ashby, M. Bautz, M. Bayliss, et al. Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the First 720 Square Degrees of the South Pole Telescope Survey. *ApJ*, 763:127, February 2013.

B. Reichborn-Kjennerud, A. M. Aboobaker, P. Ade, F. Aubin, C. Baccigalupi, C. Bao, J. Borrill, C. Cantalupo, D. Chapman, J. Didier, et al. EBEX: a balloon-borne CMB polarization experiment. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7741 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2010.

P. L. Richards. Bolometers for infrared and millimeter waves. *Journal of Applied Physics*, 76:1–24, July 1994.

A. G. Riess, L. Macri, S. Casertano, H. Lampeitl, H. C. Ferguson, A. V. Filippenko, S. W. Jha, W. Li, and R. Chornock. A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3. *ApJ*, 730: 119–+, April 2011.

J. Ruhl, P. A. R. Ade, J. E. Carlstrom, H.-M. Cho, T. Crawford, M. Dobbs, C. H. Greer, N. W. Halverson, W. L. Holzapfel, T. M. Lanting, et al. The South Pole Telescope. In J. Zmuidzinas, W. S. Holland, and S. Withington, editors, *Proc. SPIE, Vol. 5498, Millimeter and Submillimeter Detectors for Astronomy II*, pages 11–29, Bellingham, October 2004. SPIE Optical Engineering Press.

K. K. Schaffer, T. M. Crawford, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, A. T. Crites, T. de Haan, et al. The First Public Release of South Pole Telescope Data: Maps of a 95 deg^2 Field from 2008 Observations. *ApJ*, 743:90, December 2011.

J. Schaye, C. Dalla Vecchia, C. M. Booth, R. P. C. Wiersma, T. Theuns, M. R. Haas, S. Bertone, A. R. Duffy, I. G. McCarthy, and F. van de Voort. The physics driving the cosmic star formation history. *MNRAS*, 402:1536–1560, March 2010.

D. Schwan, P. A. R. Ade, K. Basu, A. N. Bender, F. Bertoldi, H.-M. Cho, G. Chon, J. Clarke, M. Dobbs, D. Ferrusca, et al. Invited Article: Millimeter-wave bolometer array receiver for the Atacama pathfinder experiment Sunyaev-Zel'dovich (APEX-SZ) instrument. *Review of Scientific Instruments*, 82(9):091301, September 2011.

L. D. Shaw, G. P. Holder, and P. Bode. The Impact of Halo Properties, Energy Feedback, and Projection Effects on the Mass-SZ Flux Relation. *ApJ*, 686:206–218, October 2008.

E. Shirokoff, C. L. Reichardt, L. Shaw, M. Millea, P. A. R. Ade, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, et al. Improved Constraints on Cosmic Microwave Background Secondary Anisotropies from the Complete 2008 South Pole Telescope Data. *ApJ*, 736:61–+, July 2011.

G. Smecher, F. Aubin, E. Bissonnette, M. Dobbs, P. Hyland, and K. MacDermid. Biasing and Demodulation Firmware for Kilopixel TES Bolometer Arrays. *ArXiv e-prints*, August 2010.

G. Smecher, F. Aubin, E. George, T. de Haan, J. Kennedy, and M. Dobbs. An Automatic Control Interface for Network-Accessible Embedded Instruments. *ACM SIGBED Review, Special Interest Group on Embedded Systems*, 2012.

B. Stalder, J. Ruel, R. Šuhada, M. Brodwin, K. A. Aird, K. Andersson, R. Armstrong, M. L. N. Ashby, M. Bautz, M. Bayliss, et al. SPT-CL J0205-5829: A z = 1.32 Evolved Massive Galaxy Cluster in the South Pole Telescope Sunyaev-Zel'dovich Effect Survey. *ApJ*, 763:93, February 2013.

K. Story, E. Leitch, P. Ade, K. A. Aird, J. E. Austermann, J. A. Beall, D. Becker, A. N. Bender, B. A. Benson, L. E. Bleem, et al. South Pole Telescope software systems: control, monitoring, and data acquisition. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8451 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, September 2012a.

K. Story, E. Leitch, P Ade, K.A. Aird, J.E. Austermann, J.A. Beall, D. Becker, B.A. Benson, L.E. Bleem, J. Britton, et al. Software Systems for Control, Monitoring and Data. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2012b.

Y. Takei, N. Y. Yamasaki, W. Hirakoso, S. Kimura, and K. Mitsuda. SQUID multiplexing using baseband feedback for space application of transition-edge sensor

microcalorimeters. *Superconductor Science Technology*, 22(11):114008, November 2009.

The Polarbear Collaboration, J. Errard, P. A. R. Ade, A. Anthony, K. Arnold, F. Aubin, D. Boettger, J. Borrill, C. Cantalupo, M. A. Dobbs, et al. The new generation CMB B-mode polarization experiment: POLARBEAR. *ArXiv e-prints*, November 2010.

J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *ApJ*, 688:709–728, December 2008.

K. Vanderlinde, T. M. Crawford, T. de Haan, J. P. Dudley, L. Shaw, P. A. R. Ade, K. A. Aird, B. A. Benson, L. E. Bleem, M. Brodwin, et al. Galaxy Clusters Selected with the Sunyaev-Zel'dovich Effect from 2008 South Pole Telescope Observations. *ApJ*, 722:1180–1196, October 2010.

J. D. Vieira, T. M. Crawford, E. R. Switzer, P. A. R. Ade, K. A. Aird, M. L. N. Ashby, B. A. Benson, L. E. Bleem, M. Brodwin, J. E. Carlstrom, et al. Extragalactic Millimeter-wave Sources in South Pole Telescope Survey Data: Source Counts, Catalog, and Statistics for an 87 Square-degree Field. *ApJ*, 719:763–783, August 2010.

A. Vikhlinin, A. V. Kravtsov, R. A. Burenin, H. Ebeling, W. R. Forman, A. Hornstrup, C. Jones, S. S. Murray, D. Nagai, H. Quintana, et al. Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. *ApJ*, 692:1060–1074, February 2009.

M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro. Precision Determination of the Mass Function of Dark Matter Halos. *ApJ*, 646:881–885, August 2006.