

# The Missing Baryon Component: A search for the Cosmic Web using the CMB

Mitchell de Zylva

Supervised by  
Dr Christian Reichardt

School of Physics  
Faculty of Science  
University of Melbourne



November 8, 2019

Submitted in fulfillment of the requirements of the degree of  
Master of Science (Physics)

**Statement of contribution:**

This is to certify that:

- This thesis entitled “The Missing Baryon Component: A search for the Cosmic Web using the CMB” comprises only my original work except where indicated otherwise.
- Due acknowledgement has been made in the text to all other material used.
- The thesis is no longer than 50 pages in length, inclusive of tables, figures, bibliographies and appendices.

.....

Mitchell de Zylva

**Acknowledgements:**

I would like to acknowledge and thank both Dr. Christian Reichardt, and Dr. Federico Bianchini, for all their help and support throughout the production of this work.

I would also like to dedicate this thesis to my father, without whom I would not have made it through this

It matters not how strait the gate,  
How charged with punishments the scroll,  
I am the master of my fate,  
I am the captain of my soul.

## Abstract

Modern cosmology is in a golden age. Access to larger and larger datasets have allowed astronomers to construct and constrain some of the most detailed models of the universe available. However, as more data becomes available, there has been an increase in tension between measurements of fundamental cosmological parameters in different areas of the universe. This opens the door for the possibility that our models are wrong, and for the development of new physics. One such tension is the missing baryon problem, a discrepancy between the measured amount of ordinary baryonic matter at low ( $z < 2$ ) and high redshifts ( $z > 2$ ). At high redshift, sources from the Cosmic Microwave Background (CMB), baryon acoustic oscillations, gravitational lensing, and quasar absorption spectrum all place a very firm constraint on the amount of matter that exists in the universe. However, when we search for this proportion in the near universe, we find our estimates missing between 30 and 50 percent of the expected matter. By making use of a spectral distortion in the CMB, the thermal Sunyaev-Zel'dovich effect (tSZ), which acts as a tracer for the large scale structure of the universe, we construct an algorithm which makes use of both CMB and near universe galaxy catalogues to search for this missing baryon fraction.

This work takes the redMaGiC luminous red galaxy catalogue from the *Dark Energy Survey*, and high resolution images of the CMB from the *South Pole Telescope* and the *Planck* satellite, and searches for the correlation between these galaxies and the tSZ signal. A set of 787,058 galaxy pairs were constructed, with line of sight separations  $< 20 h^{-1} \text{ Mpc}$  and transverse separations  $4 h^{-1} \text{ Mpc} - 20 h^{-1} \text{ Mpc}$ .

We stack pairs by rescaling and rotating the positions of the galaxies so that they sit within a common reference frame. A circularly symmetric halo model is then subtracted from the stacked data in order to obtain a residual signal between the galaxies.

This revealed a  $2.05\sigma$  detection of the existence of the missing baryons in a filamentary structure, with a mean Compton- $y$  measurement of  $1.29 \times 10^{-8}$ . This is in agreement with other measurements of the missing baryon component, and opens the possibility for future work in constraining the missing baryon component at low redshift.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Modern Cosmology, and the Cosmic Microwave Background . . . . .	5
1.2	Cosmological Parameters . . . . .	7
<b>2</b>	<b>The Missing Baryon Problem</b>	<b>9</b>
2.1	A Census of Baryons at Low Redshift . . . . .	10
2.1.1	Stellar Baryons . . . . .	10
2.1.2	Cold Interstellar Medium . . . . .	10
2.1.3	Quasar Absorbers . . . . .	11
2.1.4	Summary . . . . .	13
2.2	Warm-Hot Interstellar Medium . . . . .	13
2.3	Sunayev-Zeldovich Effect . . . . .	15
2.3.1	Atomic Physics . . . . .	15
2.3.2	CMB Signal . . . . .	17
<b>3</b>	<b>Stacking Methodology</b>	<b>20</b>
3.1	Null Tests . . . . .	25
3.1.1	Un-Physical Pairs . . . . .	25
3.1.2	Random Stack . . . . .	26
<b>4</b>	<b>Results</b>	<b>27</b>
4.1	Null Tests . . . . .	38
4.1.1	Un-Physical Pairs . . . . .	38
4.1.2	Random Stack . . . . .	41
<b>5</b>	<b>Discussion and Conclusion</b>	<b>43</b>
5.1	Discussion . . . . .	43
5.1.1	Conclusion . . . . .	44

# Chapter 1

## Introduction

### 1.1 Modern Cosmology, and the Cosmic Microwave Background

Modern cosmology has been experiencing a golden age since the 1990s. Access to larger and larger datasets has allowed astronomers to develop increasingly more detailed models of the universe. These models rely on several fundamental assumptions stemming from observation, the chief of which is the Big Bang Model.

In 1929, Edwin Hubble confirmed a relation earlier found by Georges Lematre, between distances to galaxies and their recessional velocities: galaxies appear to be moving faster away from us, the further away they are (Hubble, 1929). These observations gave rise to the Hubble Law,

$$v \propto d \tag{1.1}$$

where  $v$  is the velocity of a galaxy in question, and  $d$  is its distance from us.

To the best of our knowledge, there is nothing that suggests that our position in the universe is special or unique, a concept known as the Copernican Principle (Peacock & Peacock, 1998). This leads to the logical extension that if every point in the universe is moving away from every other point, the universe must be expanding. If we reverse the flow of time, the universe must have been an incredibly hot, dense environment at some point in the past, a singularity known as the *Big Bang*.

This singularity then expanded outwards, bringing into existence the early universe. Small fluctuations present in quantum fields, resulting from Heisenberg's Uncertainty Principle, created small inhomogeneities which seeded the universe with elementary quantum objects, such as quarks, gluons, photons, and dark matter. These particles swirled in a primordial fluid, mixing together, until eventually the universe cooled to a point that all the elements of the mixture stopped interacting with (i.e. de-couple from) each other.

This early exponential expansion allowed the universe to mix, and so achieve a very high level of isotropy. Observationally, there appears to be no favoured direction in the universe, since distributions of distant galaxies and other extragalactic sources seem to be even across the sky. Perhaps the most spectacular example of this isotropy is the presence of the *Cosmic Microwave Background* (CMB).

Discovered in 1964 (Penzias & Wilson, 1965) as a source of noise in another microwave ex-

periment, over the next 30 years, it was noticed that there was isotropic black-body radiation everywhere in the universe, at a temperature of  $T \approx 2.7\text{K}$ . This radiation is most intense in the microwave section of the electromagnetic spectrum, it was named the cosmic microwave background.

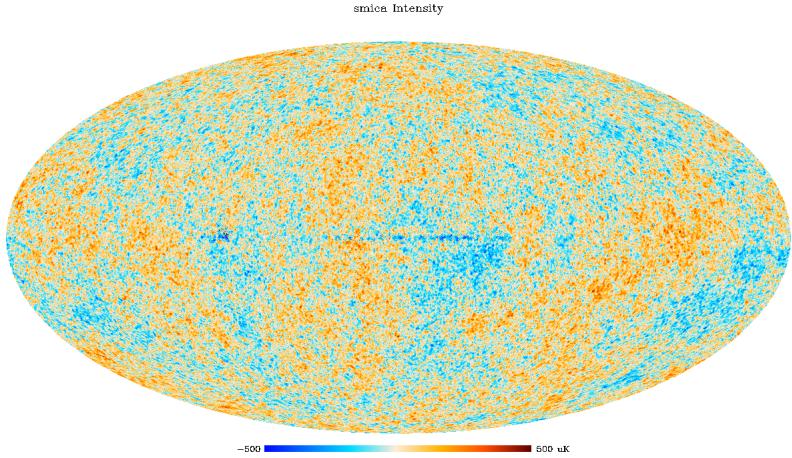


Figure 1.1: *Planck* Satellite Full Sky CMB Map, extracted using the SMICA method (Planck Collaboration et al., 2018b). This map was produced by linearly combining the full spectrum of frequencies observed by the *Planck* satellite from 30 GHz to 857 GHz in frequency space. Each map is first converted to its power spectrum, and then weighted at various multipoles, in order to account for contamination which appears at characteristic scales in different frequencies.

The CMB is a near-perfectly uniform field of background radiation which permeates the universe. Initially believed to be featureless, today the CMB is known to have very specific inhomogeneities, shown in 1.1. The uniformity of the CMB is the closest measurement we have to a perfect black-body, but with variations of approximately one part in 100,000, and an RMS of  $18\,\mu\text{K}$  (Wright, 2004).

Theory holds that a very short time after the Big Bang ( $\sim 10^{-37}$  seconds), the universe underwent an exponential growth period now termed *inflation*. This phase is necessary to ensure that the universe is isotropic, flat, and gaussian, whilst also still taking into account the fact that opposite sides of the observable universe appear to be too far apart to ever have been causally connected. During this period, the universe was smoothed out, only leaving behind very small irregularities. These irregularities are what eventually gives rise to the large scale structure of the universe, what eventually would become the *Cosmic Web*.

As the universe adiabatically cooled, there was a considerable period of time, until approximately 380,000 years after the Big Bang, where photons were coupled to the other components of the universe, such as baryonic matter and dark matter. This period is what allowed for the quark-gluon plasma to mix appropriately, and develop waves in the fluid, ultimately resulting in the characteristic pattern in the CMB. This pattern is highly dependant on statistical parameters, and so the characteristic angular size of the pattern of the CMB is incredibly sensitive to the relative proportions of the universe's matter-energy density. These characteristic inhomogeneities can be decomposed into an angular power spectrum, which exhibits a shape highly dependant on universal parameters.

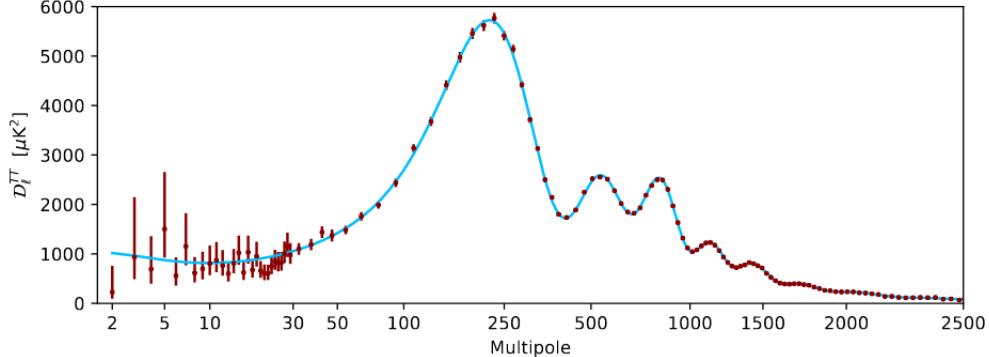


Figure 1.2: *Planck* 2018 Angular Power Spectrum (Planck Collaboration et al., 2018a). The location, and relative heights of the first three peaks are sensitive to the overall energy density of the universe, as well as the relative amounts of baryonic matter and dark matter.

Shown in Figure 1.2, this angular power spectrum is very sensitive to cosmological parameters. Since the discovery of its anisotropy by *Cosmic Background Explorer* (COBE) (Levi, 1992), and then its later refinement by *Wilkinson Microwave Anisotropy Probe* (WMAP) (Spergel et al., 2007), the power spectrum has been the gold standard by which astronomers measure cosmological parameters. The precision to which we know the CMB also makes it very good as constraining these parameters, leading many cosmologists to hold it as one of the most accurate measurements ever made in physics.

## 1.2 Cosmological Parameters

The current accepted model, the  $\Lambda$ CDM model, contains six independent parameters which describe the evolution and behaviour of the universe: the physical baryon density  $\Omega_b h^2$ , the physical dark matter density  $\Omega_c h^2$ , the age of the universe  $t_0$ , the scalar spectral index  $n_s$ , the acoustic scale  $100\theta_*$ , and the reionisation optical depth  $\tau$ , the values of which are reported in Table 1.1. These primary parameters give rise to other parameters, such as the Hubble constant  $H_0$  and its unitless form, known as the reduced Hubble constant  $h$ . It serves as a measure of the current rate of expansion of the universe, and carries units of  $\text{km s}^{-1} \text{Mpc}^{-1}$ . The Hubble constant is usually quoted for a number of reasons. Firstly, it contains both the age, and the size, of the universe within it, and secondly, it was one of the first cosmological parameters to be determined, back when Hubble was making his initial measurements, so it has some historical value. Because it is so ubiquitous, astronomers use different forms depending on what they are analysing, which becomes important for making direct comparisons between different measurements. When using the unitless form, if it lacks a subscript, it refers to the definition  $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . Alternatively, if it carries one, it refers to replacing the number in the denominator with the number in the subscript, e.g.  $h_{70} = H_0/70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

Currently, the highest precision measures of these features from the CMB come from the Planck Collaboration et al. (2018a) paper, which details that baryonic matter only comprises  $\approx 5\%$  of the universe's energy density.

In principle, this component of the universe should be directly measurable. At just three minutes after the Big Bang, deuterium can be used as a tracer for this abundance (Steigman, 2007), and at a redshift  $z \geq 2$ , the baryon fraction can be found in the absorption lines of quasars

passing through the diffuse, photo-ionised intergalactic medium, known as the Lyman- $\alpha$  forest (Weinberg et al., 1997).

The baryon content has been confirmed very well at high redshift by the Planck collaboration, and the agreement between the CMB, and other independant high redshift measurements, such as baryon acoustic oscillations, and gravitational lensing reconstructions .

However as the universe evolved, this gas became sparser, both as space expanded, and as it became more ionised by processes in the universe. This makes searching for the entirety of the baryon fraction at low redshift difficult, since high density objects are usually more easily detected. When this fraction is calculated for the local universe directly from observations, it shows only one tenth of the baryonic content shown in high redshift measurements is contained in galactic structures (Persic & Salucci, 1992). This is troubling, because tensions in cosmological parameters between their high and low redshift measurements forces us to ask questions about the validity of cosmological models at all times.

Parameter	Value	Error
$\Omega_c h^2$	0.120	$\pm 0.001$
$\Omega_b h^2$	0.0224	$\pm 0.0001$
$n_s$	0.965	$\pm 0.004$
$\tau$	0.054	$\pm 0.007$
$100\Theta_*$	1.0411	$\pm 0.0003$
$H_0$ (km s $^{-1}$ Mpc $^{-1}$ )	67.4	$\pm 0.5$

Table 1.1: Latest Reported Values for Cosmological Parameters from the *Planck* Satellite (Planck Collaboration et al., 2018a).

# Chapter 2

## The Missing Baryon Problem

The *Missing Baryon problem* is that we see more baryons at high redshift, than at low redshift. At high redshift ( $z > 2$ ), the Lyman- $\alpha$  forest provides a good measure of the proportion of baryons, since at these redshifts the majority of baryons in the universe are contained in diffuse, low-density gas. These analyses give a reported value of

$$\Omega_b \geq 0.035.$$

Observed light-element ratios and standard nucleosynthesis allows for direct computation of the expected baryon densities (Burles & Tytler, 1998), which is in agreement with the above figure, giving a result of

$$\Omega_b = (0.019 \pm 0.001)h^{-2} = 0.039 \pm 0.002.$$

The  $\sim 2 - 3\sigma$  agreement between these two measures of baryon density, and the measurement obtained from the CMB ( $\Omega_b = (0.0224 \pm 0.0001)h^{-2}$ ) lends confidence to the value obtained by Planck. The CMB measurement is far more precise than the other measurements due to it requiring inherently fewer assumptions, and so the systematics for a CMB measurement are often much easier to quantify. In contrast, at low redshifts all analysis indicates that the summing over all observed contributions gives a value of

$$\Omega_{\star} + \Omega_{HI} + \Omega_{H_2} + \Omega_{X-Ray,cl} \approx 0.0068 \leq 0.011,$$

where  $\Omega_{\star}$  refers to the density of stars,  $\Omega_{HI}$  is the density of neutral hydrogen,  $\Omega_{H_2}$  is the density of molecular hydrogen, and  $\Omega_{X-Ray,cl}$  is the density derived from X ray sources, and other measurements of clusters.

This severe discrepancy between measurements at high and low redshifts suggests that either the majority of the baryons at low redshifts are yet to be detected, there are fundamental errors in numerous independent measures of the baryon density at high redshift, or there is yet-to-be uncovered new physics. Occams razor suggests that new physics is unlikely, given how self-consistent both sets of measurement are, and the agreement between independent high redshift measurements, such as the Lyman  $\alpha$  forest, and the CMB, suggests that the most likely problem is that we have not made an accurate account of baryons in the local universe.

## 2.1 A Census of Baryons at Low Redshift

### 2.1.1 Stellar Baryons

The most obvious location to search for baryons are in the stellar populations of galaxies. At a broad level, we can imagine that there are two distinct stellar populations which can be considered to be found in high density galaxies; a class of old stars which exists in the bulge of a galaxy ( $\Omega_{*,\text{Bulge}}$ ), and a class of young stars in the disk ( $\Omega_{*,\text{Disk}}$ ), as well as a third population existing in irregular galaxies ( $\Omega_{*,\text{Irr}}$ ).

Estimating the proportion of stellar baryons therefore becomes an exercise in galactic morphology and luminosity density function computation (Fukugita et al., 1998). Performing this calculation gives mean mass density numbers for these three classes of stars of

$$\begin{aligned}\Omega_{*,\text{Bulge}} &= (0.00180^{+0.00121}_{-0.00085})h^{-1} \\ \Omega_{*,\text{Disk}} &= (0.00060^{+0.00030}_{-0.00024})h^{-1} \\ \Omega_{*,\text{Irr}} &= (0.000048^{+0.000033}_{-0.000026})h^{-1}\end{aligned}$$

These numbers depend on the mass-to-light ratio for age estimation, and so in turn are dependent on the cosmological parameters in a complex way. Even if efforts were made to remove this dependency by changing the methodology used to calculate the mass-to-light function, the necessity for the new methodology to hold consistent with other measurements would force the dependency regardless.

### 2.1.2 Cold Interstellar Medium

We also know that some of the baryons in the local universe are stored in the cold interstellar medium, a term for neutral and molecular gas, primarily consisting of unionised hydrogen (HI). The HI present in gas-rich galaxies is the best tracer for the neutral hydrogen mass content in the near universe, since it can be shown that very little neutral gas is present in objects outside the galaxy population in these regions. At redshift  $z \approx 0$ , there have not been any free HI clouds detected by radio surveys that have not yet subsequently been detected in optical bands (Rao & Briggs, 1993), and so optical detection in galaxies can be considered to be adequate when taking a census of the total baryons contained in the cold interstellar medium.

The HI content of a given galaxy can be calculated from the optical luminosity functions of given galaxy morphological types. For a given optical luminosity function  $\phi^T(M)$ , and a given HI mass content  $M_{HI}^T(M)$ , as functions of absolute magnitude  $M$ , the total HI mass content of a given galaxy morphological type  $T$  can be found by computing an integral over all luminosities

$$M_{HI} = \int \phi^T(M) M_{HI}^T(M) dM$$

Adding all the contributions for all morphological types will therefore give an estimate of the total HI at  $z = 0$ . At these distances, spiral galaxies are the primary contributors to the neutral hydrogen content, containing  $\sim 89\%$  of the hydrogen mass. Initial estimates (Rao & Briggs, 1993) place the mass density of neutral hydrogen at the present time at

$$\Omega_{HI} = (2.5 \pm 0.6) \times 10^{-4} h_{75}^{-1}$$

Later surveys, such as the *HI Parkes All-Sky Survey* (HIPASS), refined this measurement further (Zwaan et al., 2003). HIPASS is a blind survey of the southern sky south consisting of approximately 7000 galaxies. The much higher number of galaxies in this survey necessitated a different method of computing the neutral hydrogen fraction. Taking a statistical approach the probability that a galaxy with a given HI mass is

$$p(M_{HI,i}|D) = \frac{\theta(M_{HI,i})}{\int_{M_{HI,\text{lim}(D_i)}}^{\infty} \theta(M_{HI}) dM_{HI}},$$

where  $M_{HI,\text{lim}(D_i)}$  is the minimum detectable HI mass at some distance  $D_i$ . This essentially gives the fraction of galaxies in the survey with a given HI mass sufficient to be detected. The parent distribution,  $\theta$ , can then be maximised by finding which product of probabilities is also maximal. The statistical nature of this method requires accounting for various forms of bias, such as selection bias in the survey, the Eddington effect, self-absorption of the neutral hydrogen, and cosmic variance (Shull et al., 2012). This methodology gives a measure of the mass density as

$$\Omega_{HI} = (3.8 \pm 0.6) \times 10^{-4} h_{75}^{-1}$$

### 2.1.3 Quasar Absorbers

One source that can act as a tracer for intervening matter at all redshifts is quasar absorption spectra. Quasars are extremely bright active galactic nuclei that emit light when matter falls into the potential of a black hole, and is ripped apart. The photons that result from this then propagate outwards, until they intersect with some intervening structure. As such, we can look for known absorption lines in their spectra, and determine the redshift at which they interacted.

#### Photoionised Lyman $\alpha$

At higher redshifts, essentially the entirety of the baryon content in the universe can be found in large quantities of gas that have not yet collapsed into galaxies, which shows up very clearly in the Lyman  $\alpha$  (Ly $\alpha$ ) forest.

The Lyman  $\alpha$  forest is a series of absorption lines that are visible in the spectra of distant quasars, which intersect with clouds of neutral hydrogen along their line of sight to us. Since the light emitted by the quasar passes through many hydrogen clouds, and is redshifted by cosmological expansion, there will be a characteristic absorption signal in the spectrum which corresponds to the position of the clouds relative to the quasar. These absorption signals make the Quasar spectrum look like a forest, hence, the name Lyman  $\alpha$  forest.

At low redshifts, we can search for the residual Ly $\alpha$  signal using neutral hydrogen (HI) absorber frequencies in distant quasars. By making the assumption that these absorbers are isothermal spheres, and choosing a given impact parameter, the total cloud mass can be essentially inferred from the HI column density.

One model for estimating the contribution of the local Lyman  $\alpha$  forest is outlined by Penton et al. (2000). From big bang nucleosynthesis, the total baryons contributed from the low- $z$  Ly $\alpha$  is given by

$$\Omega_{Ly\alpha} = \int_{N_{min}}^{N_{max}} \frac{\phi_0(N_{HI}, p) M_{cl}(N_{HI}, p, J_0)}{\rho_{crit}} dN_{HI},$$

where  $M_{cl}(N_{HI}, p, J_0)$  is the mass of an individual cloud,  $\phi_0(N_{HI}, p)$  is space density of the clouds,  $N_{HI}$  is the column density of the clouds,  $J_0$  is the specific intensity of the metagalactic ionizing

radiation field,  $\rho_{crit} = 3H_0^2/8\pi G$  is the critical density at the present day which is necessary to maintain flatness.

Making more assumptions about the composition of the spheres, and that the gas is in photoionising equilibrium, and the impact parameter, this reduces the integral to

$$\Omega_{Ly\alpha} = (0.008 \pm 0.001) \left[ J_{-23} p_{100} \left( \frac{4.8}{\alpha_s + 3} \right) \right]^{1/2} h_{70}^{-1},$$

where  $\alpha_s$  is the spectral index of the radiation field,  $J_{-23} = J_0/10^{-23}$ ,  $p_{100} = p/(100 \text{ kpc})$  and  $h_{70} = H_0/(70 \text{ Mpc})$ . This corresponds to approximately 20% of the baryons, but it inherently makes a number of rather significant assumptions, which are highly dependant on the number of Lyman  $\alpha$  absorbers detected at low redshifts.

Further accounting for the fact that the clouds are gravitationally bound, and that their densities are typical over a characteristic Jeans length allows for further refinement of this number (Schaye, 2001). Doing so yields a measure of the number of baryons in the residual Lyman  $\alpha$  forest of  $\Omega_{Ly\alpha} = 29\% \pm 4\%$  (Danforth & Shull, 2008; Penton et al., 2004)

### OVI and BLA Absorbers

Because gas at low redshift exists at a large range of temperatures, and at a higher metalicity than the same gas at higher redshifts, distant quasars will exhibit absorption from higher Lyman alpha lines, such as oxygen (OVI), and broad Lyman  $\alpha$  lines (BLA). The OVI absorption line probes gas in temperature ranges from  $10^5 - 10^6 \text{ K}$ , which has been shock heated as a result of gravitational instability during structure formation. Gas hotter than this is only sensitive to very weak absorption lines from higher ions of oxygen, neon, or nitrogen (such as OVII, OVIII, NeX, NVI, and NVII), which are only detectable in weak X-Rays (Danforth & Shull, 2005).

Broad Lyman  $\alpha$  also act as a tracer of the gas in these temperature ranges. Theory suggests that the ionisation equilibrium of gas should contain a very small portion of neutral gas, typically  $f \sim 10^{-5} - 10^{-6}$ , so there should be some Ly  $\alpha$  emission, thermally broadened by the intervening gas (Richter et al., 2006). It is not entirely clear if the OVI and BLA absorbers trace the same underlying gas phase, so they have been included here together.

The method of calculating these numbers is a comparatively simply integral, over the number of absorbers per redshift bin. This does make several assumptions which are sensitive to the visibility, as well as the number of absorbers in a given survey.

Danforth & Shull (2005) found the proportion of matter contained in OVI absorbers to be approximately 5%, extrapolated from  $\sim 50$  absorbers. However, this number is highly dependant on the model chosen to describe the spread of oxygen/hydrogen metalicities, the redshift range of the survey, and OVI ionisation fractions, leading to large errors on this measurement.

Richter et al. (2006) found the proportion from BLAs to be  $\sim 15\% - 150\%$  of the total baryon fraction, from a sample of between 20 and 50 sources. The reason for the large uncertainty is the inherent uncertainty in the detection of BLA sources. In fact, the unreasonably high measurement of the baryon fraction says that this method is inherently counting sources which are not actually good tracers of the free baryons. There is also some confusion regarding sources which appear as both OVI and BLA absorbers, since the connection between the two types is not clear.

## 2.1.4 Summary

These methods all carry some sources of error with them, some so large that they call into question their validity entirely. Calculations of stellar baryons and cold interstellar medium are limited by the brightness that a given survey instrument is sensitive to, and so real world considerations like integration time, survey depth, and survey area make using them as an accurate measure difficult. The limiting factor for absorption studies is the incredibly low number of observed absorbers. The areas of phase space that absorption studies probe contains requires bright sources such as quasars to align with intervening matter. This low likelihood, and the unknowns associated with models of the source behaviour makes this a vague probe of the baryon fraction at best. There is also the very real possibility that they probe the same phase space of baryons, leading to an unknown level of degeneracy in the overall baryon fraction.

If we include some smaller contributions, such as from cluster contributions ( $\Omega_b^{(cl)} \sim 4\%$ ) (Fukugita & Peebles, 2004), or the circumgalactic medium ( $\Omega_b^{(CGM)} \sim 5\% \pm 3\%$ ), these still do not account for the entirety of the baryon component known from the CMB. Some revised estimates of the total baryon content at low redshift considered that the limitations of observations were primarily to blame for the discrepancy, and not inherently new physics (Bristow & Phillipps, 1994; Fukugita et al., 1998).

If we examine the latest estimates of the local baryon content in its entirety by Shull et al. (2012), shown in Table 2.1, it becomes clear that up to 30% of the universe's baryons are not being located by direct observational methods.

Source	$\Omega_b$	Percentage	Error
Photoionised Lyman $\alpha$ Forest	28%	$\pm 11\%$	
X Ray Absorbers (OVI and BLA)	31%	$\pm 11\%$	
Galaxies	7%	$\pm 2\%$	
Circumgalactic Medium	5%	$\pm 3\%$	
Intercluster Medium	4%	$\pm 1.5\%$	
Cold Gas	1.7%	$\pm 0.4\%$	
Missing	29%	$\pm 13\%$	

Table 2.1: Table of Current Baryon Census

This is incredibly unsatisfactory, because so many models rely on fundamental parameters to hold consistent at all redshifts, and despite our best efforts, it seems like we cannot get the baryon content at low redshifts to agree with the baryon content at high redshift. We must therefore ask the question, are we looking in the right place?

## 2.2 Warm-Hot Interstellar Medium

High resolution hydrodynamical simulations allow us to predict the overall structure of the cold dark matter in the universe (Cen & Ostriker, 1999, 2006; Davé et al., 2001; Smith et al., 2011). Because dark matter is so much more prevalent than baryonic matter, and it only interacts gravitationally, it stands to reason that the baryons will fall into the potential wells produced by the gravitationally bound dark matter, and so will trace the underlying dark matter structure, but with less density. These hydrodynamical simulations can therefore be used to estimate the baryon distribution at low and moderate redshifts, by informing us where we should be searching for our

missing baryons.

It is clear that by the current era, hierarchical structure formation collects baryons in gravitational potential wells formed by the dark matter, which moves a significant portion of the baryon component that was previously located in the intergalactic medium at higher redshifts, into structure, such as stars, galaxies, groups, and clusters.

These simulations indicate that the baryons at low redshift fall into four general phases, defined by the overdensity  $\delta \equiv \rho/\bar{\rho} - 1$  (where  $\bar{\rho}$  is the mean density of baryons).

- Diffuse Gas:  $\delta < 1000$ ,  $T < 10^5 K$ , Photoionised gas which is visible in Lyman- $\alpha$  absorption spectra
- Condensed:  $\delta > 1000$ ,  $T < 10^5 K$ , Stars and cool galactic gas
- Hot:  $T > 10^7 K$ , Galaxy Clusters and Groups
- Warm-Hot:  $10^5 K < T < 10^7 K$ , Warm-Hot Intergalactic Medium (WHIM)

Simulations (Cen & Ostriker, 1999; Davé et al., 2001) indicate that at redshift  $z = 0$  approximately 30–40% of baryonic mass is contained within the last category, in the WHIM. WHIM gas seems to primarily trace filamentary large scale structures, and clusters around sites of galaxy formation. Because the gas is not bound or virialised, it is apparent that the mechanism which heats it to such high temperatures is shock-heating, caused by gas accreting onto large scale structure. This is consistent with measurements from the soft X-ray background.

Because the temperature and density of the WHIM are correlated, and the WHIM is in turn correlated with the large scale structure, we can use the presence of other tracers of large structure, temperature and density, to search for the baryons contained in the WHIM.

The WHIM is so highly ionised, and so sparse (with average densities of the order of 10 particles per cubic meter), that they can only emit or absorb far-ultraviolet or soft x-ray photons. These photons are primarily at highly ionised lines of C, O, Ne, and Fe (Cen & Fang, 2006).

Tracking the baryons contained in the WHIM can only be done by exploiting both experimental multiwavelength observations and theoretical calculations. Given this, we can turn to tracers we know probe some area of the baryon phase space, our spectroscopic absorbers. X-Ray and UV spectroscopic surveys measure the mass of WHIM using the relative and absolute metal content, and the ionisation correction. These can be then combined with optical and infrared photometry and spectroscopy which measure dark matter concentrations by measuring galaxy density around WHIM filaments. These observations then feed into simulations, which allow for more detailed study of virialised structure and the intergalactic medium.

The intensity of the signals obtained from direct observation is low in both the UV and the X-Ray bands, both as a result of the density and the relatively small size of the filaments (1 - 10 Mpc). Direct detection ideally requires large field of view and effective area imagers-spectrometers, which is not currently available. The strategy that can best be utilised with current technology involves searching for discrete absorption lines in the spectra of bright, featureless background astrophysical sources.

The key feature necessary to detect an absorption line is the ratio between the line wavelength and its equivalent width, called its transition contrast. For the dominant absorption lines of oxygen, in this case OVI, the current resolving power of UV spectrometers is sufficient to measure

its transition contrast, but for the X-Ray band it is worse, and so searches for the WHIM have proven more fruitful in the UV band than in the X-Ray (Danforth & Shull, 2005; Richter et al., 2006). Using hydrodynamical simulations to replicate the observed absorption per unit redshift, it can be shown that if the WHIM was based solely on the OVI absorption, it would only account for approximately 10 percent of the missing baryon component.

An alternative method for searching for the missing mass is to look for hydrogen absorption in broad Ly- $\alpha$  absorbers (BLAs). At the temperatures that the WHIM is thought to exist at, most of the hydrogen will be ionised, but left-over neutral hydrogen can still imprint Lyman series absorption onto the ultraviolet spectra of background objects. These lines will be very broad, given that the temperatures of the WHIM create a Doppler parameter of the order of  $b \approx 40 \text{ km/sec}$ . This technique again gives a similar measurement to that done with OVI absorbers, and so suggests that BLAs and OVI absorbers can be considered to be good tracers of the WHIM. However, they aren't sufficient to probe the entirety of the missing mass due to the majority of it existing in temperatures only probed by the X-Ray band.

Comprehensive studies of the WHIM therefore require both the UV and X-Ray bands, since the X-Ray is crucial to detect the WHIM, and provide an accurate ionisation correction, and the UV is necessary to measure the associated amount of HI and hence the baryonic mass of the system.

According to theory, the chances of finding a WHIM filament along an arbitrary line of sight increases with the path length crossed between the observer and the beacon used to obtain the X-Ray images of intervening space, and the inverse of the baryon column density in the filament in question. This tells us that the larger the amount of baryons in the filament, the lower the probability of finding one. It can be shown that the detection of the WHIM is within the range of instrumentation sensitivity currently, but it requires long observation times, making it untenable as a practical survey. Searching for an alternative tracer for the WHIM is therefore necessary to accurately locate the missing baryon content.

## 2.3 Sunayev-Zeldovich Effect

### 2.3.1 Atomic Physics

The thermal Sunyaev-Zel'dovich effect is one possible tracer. The Sunyaev-Zel'dovich effect (SZ) is one type of spectral distortion in the CMB, which refers to the inverse compton scattering of CMB photons off of hot electrons in the WHIM. In order for a spectral distortion to exist in the CMB, it must occur sufficiently late in the cosmological timeline that the radiation doesn't thermalise and regain a pure Planck spectrum (approximately  $z < 2700$ ).

Compton Scattering is a form of inelastic scattering between light and free charged particles, such as electrons. There is a momentum transfer between the photon in the interaction, and the charged particle, and so the photon's wavelength changes as a result of the scattering. The mechanisms by which it occurs are described in detail by quantum field theory, and the allowed Feynmann diagrams for the interaction are shown in Figure 2.1.

It was first described in the context of X-rays interacting with electrons in atoms (Compton, 1923), and so regular Compton Scattering is taken to describe the interaction between a high energy photon, and a low energy particle. By applying principles of conservation of momentum, and conservation of energy, the formula for the shift in wavelength as a result of this scattering is

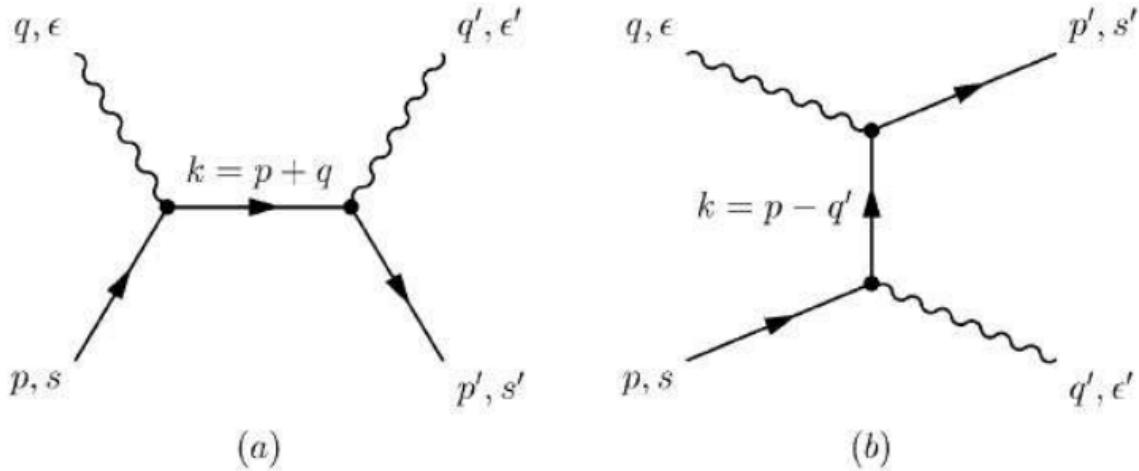


Figure 2.1: Leading order diagrams for Compton Scattering

given by

$$\Delta\lambda = \frac{h}{m_e c} (1 - \cos\theta), \quad (2.1)$$

where  $m_e$  is the mass of the electron, and  $\theta$  is the angle between the incident and scattered trajectories.

However, for the SZ effect that is observed in the CMB, the energies of the photons in question are much lower than the energies of the electrons involved, so the frequency shift is parametrised by something called the Compton- $y$  parameter. The expression given by (2.1) is functional for a single interaction, but given the number of interactions and the statistical nature of the CMB, we have to consider the more broad Compton- $y$  parameter, which is given by:

$$y = \int \frac{k_B T_e}{m_e c^2} n_e \sigma_T dl, \quad (2.2)$$

where  $m_e c^2$ ,  $k_B$ , and  $\sigma_T$  are the electron rest mass energy, Boltzmann constant, and Thompson Cross Section respectively. These are all well defined constants, and so have no effect on the integration.

The  $y$ -parameter therefore amounts to the line-of-sight integration over  $n_e T_e$ , which are the electron gas density and temperature. The degeneracy between temperature and pressure can be broken in principle by obtaining measurements of one of the two quantities, which we take from hydrodynamical simulations.

This  $y$  parameter can be calculated in the CMB from frequency and intensity information at those frequencies. Starting from Kompaneets equation (Kompaneets, 1957), the time rate of change of the photon occupation number  $\bar{n}$  due to Compton Scattering by non-relativistic, isotropic Maxwellian electron gas is given by a non-relativistic Fokker-Planck Equation, (Rephaeli, 1995)

$$\frac{\partial \bar{n}}{\partial t} = \frac{kT}{mc} \frac{\sigma_T n_e}{x^2} \frac{\partial}{\partial x} \left[ x^4 \left( \frac{T_e}{T} \frac{\partial \bar{n}}{\partial x} + \bar{n} + \bar{n}^2 \right) \right], \quad (2.3)$$

where  $x = h\nu/kT$  is the non-dimensional frequency,  $T$  is the temperature of the radiation,  $n_e$  and  $T_e$  are the number density and temperature of the electrons, and  $\sigma_T$  is the Thompson cross section. Because  $T_e \gg T$ , the first term in the parenthesis dominates, reducing the above to

$$\frac{\partial \bar{n}}{\partial t} = \frac{kT_e}{mc} \frac{\sigma_T n_e}{x^2} \frac{\partial}{\partial x} \left( x^4 \frac{\partial \bar{n}}{\partial x} \right). \quad (2.4)$$

Since the incident radiation is only weakly scattered, an approximate solution to 2.4 can be found by substituting the occupation number of a Planckian radiation field

$$\bar{n}_p(x) = \frac{1}{e^x - 1}, \quad (2.5)$$

By integrating this along a given path length, we can determine the change in spectral intensity along that path, and therefore obtain the spectral form of the SZ effect:

$$g(x) = \frac{x^4 e^x}{(e^x - 1)^2} \left[ \frac{x(e^x + 1)}{e^x - 1} - 4 \right]. \quad (2.6)$$

The shape of (2.6) is shown in Figure 2.2, illustrating how relative frequencies must be rescaled to extract the Compton- $y$  parameter. This spectral form assumes that all the velocity flow of the cluster is given by the thermal energy of the particles, and so there is often a distinction made between the thermal SZ effect (tSZ) and the kinetic SZ effect (kSZ).

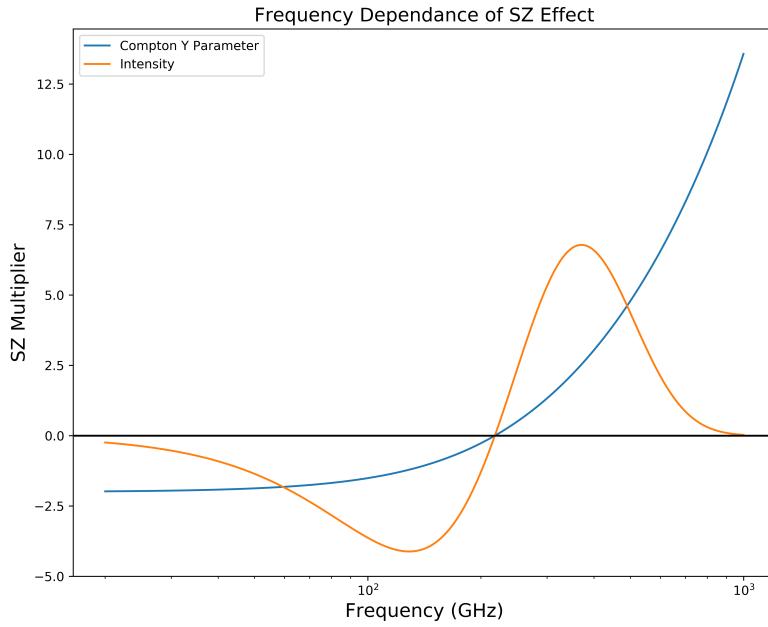


Figure 2.2: Sunyaev-Zel'dovich effect Intensity and Frequency Scaling Factors

### 2.3.2 CMB Signal

The method for constructing a map of the SZ effect is given by variants on the ‘Internal Linear Combination’ (ILC) method (Remazeilles et al., 2011). This method presumes very little about

the form of the data itself, assuming that it can be written in the form

$$\vec{x}(p) = \vec{a}s(p) + \vec{n}(p), \quad (2.7)$$

where  $\vec{x}$  is a vector of  $N_{obs}$  observations,  $s(p)$  is a single map,  $\vec{a}$  is a mixing vector, which does not depend on  $p$  but is known, and  $\vec{n}$  is some noise term, containing both instrument and astrophysical noise. It is also assumed that the maps are the same resolution.

The method then provides an estimator of the mean map,  $\hat{s}_{ILC}(p)$  of the map  $s$ , by making a linear combination  $\hat{s}(p) = \vec{w}\vec{x}(p)$ :

$$\hat{s}_{ILC} = \frac{\vec{a}^t \hat{R}^{-1}}{\vec{a}^t \hat{R}^{-1} \vec{a}} \hat{x} \quad (2.8)$$

where  $\hat{R}$  is the covariance matrix of the observations.

This method then allows for the map be linearly added together in such a way the the variance is minimised, and so the resulting map is as close to the best fit as possible.

One main advantage of the ILC component separation method is that it doesn't assume a model for the components that are not under direct consideration, they are simply collected in a catch-all nuisance term  $\vec{n}(p)$ . Unfortunately, if any of these components are correlated with the signals we are looking for, this method is unable to directly separate the signal from the noise without some a priori knowledge of the components  $a_i$ .

An updated version of the algorithm, known as the modified internal linear combination algorithm (MILCA) (Hurier et al., 2013), takes into account three changes to the above methodology. It accounts for localisation in pixel and spherical harmonic spaces to take into account variations in spatial spectral laws. These laws describe the behaviour of different processes at different frequencies, and so are very important to consider when adding maps taken at different wavelengths. It also modifies the definition of the variance being minimised by an action of the covariance matrix, to account for the possible correlation between noise and astrophysical sources.

Now, this technique can be applied to both the CMB as a whole, as well as the SZ effect specifically, because the characteristic scale and frequency dependance of the SZ is well known. The size of the anisotropy caused by the SZ effect can be determined from the number of clusters sampled from the observational beam. A detector beam is essentially an instrument's response to the patch of sky it is measuring. It is unique to each instrument, and is a function which convolves everything within its detecting area, effectively smoothing out areas of the sky smaller than its characteristic size.

The size of the anisotropy is therefore highly dependant on the constraints on cluster evolution and their subsequent properties. The range of angular scales of this anisotropy depends on the angular extent of the clusters, but ultimately sits within ranges of approximately  $1' - 10'$  for a typical cluster with a radial extent of approximately  $0.5h_{-1}\text{Mpc}$ , and a mass of  $2 \times 10^{15} M_\odot$ . For this system, we expect that the size of the anisotropy to be roughly  $\Delta T/T \approx 10^{-6} - 10^{-5}$  (Rephaeli, 1995). This makes sense, since the upper measure for the total *comptonisation* of the CMB is only  $y \approx 4 \times 10^{-3}$ . However, it does pose an issue for detecting the WHIM, since the average density of a filament is so much lower than the density of a cluster. Since the  $y$  parameter is sensitive to the temperature of the gas and the number density, reducing the number density by a factor of  $10^5 - 10^8$  and the gas temperature by  $\sim 10^2$ , will result in a corresponding reduction in the measured  $y$  parameter.

Given the signal-to-noise ratio expected for the thermal Sunyaev-Zel'dovich effect of a single filament, many such filaments must be co-added, so as to drive the signal-to-noise to a detectable

level. Initially outlined in Clampitt et al. (2016) for application to weak gravitational lensing maps, it was found that stacking  $\sim 135,000$  pairs yielded a filament mass at  $\sim 4.5\sigma$  confidence. Further follow up using the Canada France Hawaii Telescope Lensing Survey (CFHTLenS), and the Sloan Digital Sky Survey’s (SDSS) Luminous Red Galaxy (LRG) catalogue by Epps & Hudson (2017) detected the weak lensing signal from stacked filaments at  $5\sigma$  confidence.

Investigations by Van Waerbeke et al. (2014), Ma et al. (2015), and Hojjati et al. (2015) established firmly that there is a correlation between weak gravitational lensing from CFHTLenS and tSZ signals from *Planck*, which suggests that we can use tSZ in the same way as weak lensing, without having to be careful about the peculiarities associated with weak lensing, such as sufficiently nulling spherical components. This was further reinforced by Hill & Spergel (2014), who reported a  $6.2\sigma$  correlation between the *Planck* lensing potential and the *Planck* tSZ map.

# Chapter 3

## Stacking Methodology

Given the signal from a single filament is well below the level of noise in a given  $y$ -map, in order to effectively detect the signal for the filaments, we need to add many individual sources together. Because the noise and the signal in the CMB is presumed to be gaussian, the expectation is that correlated signals would coadd, and uncorrelated signals would be driven to zero. This means that by adding together thousands of signals lower than the signal-to-noise ratio of a given  $y$  map, it should show the filaments as a correlated signal, along with their corresponding galactic halos.

The CMB data used in this work is sourced from the South Pole Telescope, a 10 meter diameter telescope located at the geographic south pole. The data was collected primarily during the SPT-SZ observing run, with a primary instrument of a 960-element bolometer array of superconducting transition edge sensors. It was sensitive to three frequency bands, 95 GHz, 150 GHz, and 220 GHz.

Initially outlined in Clampitt et al. (2016), the stacking algorithm involves creating a list of galaxy pairs, which we would expect to see a filament between, and then using those pairs, forming a normalised two-dimensional image, with the galaxies of the pair being placed on two points in the image, and stacking them until the signal to noise is sufficient to be measurable.

The intial problem faced involves generating the galaxy pairs from the Dark Energy Survey redMaGiC Catalogue. The Year 1 Catalogue consists of 650 thousand red-sequence galaxies in the redshift range  $0.15 < z < 0.9$ . The algorithm used by the Dark Energy Survey to select these red-sequence galaxies provides redshift estimates of very high quality and very low bias ( $\lesssim 0.5$  percent). They also have very low scatter, and a very low rate of catastrophic outliers. The algorithm yields superior photo-z performance than the colour-cut methodology used to define the Sloan Digital Sky Survey CMASS catalogue (Rozo et al., 2016).

The redMaGiC data obtains redshifts by fitting every galaxy to a red sequence template derived from the Sloan Digital Sky Survey (SDSS) spectroscopic galaxies, along with a goodness of fit parameter. It then takes this photometric redshift, and determines the galaxy luminosity. If it is sufficiently bright, and it falls below a threshold for goodness of fit, it is included in the redMaGiC catalogue. This process results in a mean photometric error of  $\Delta z \approx 0.01(1 + z)$ . When the redMaGiC algorithm is compared to the spectroscopic redshifts from SDSS, the measure of the spread doesn't exceed 2% for outliers (Rozo et al., 2016).

The DES catalogue has a larger footprint than the SPTpol viewing footprint, so we first have to exclude any galaxies that lie outside of the SPTpol area. The SPTpol observing area consists of a 500 square degree patch of sky, extending from an Right Ascension (RA) of 22h to 2h, and a Declination (Dec) of -52 degrees to -67 degrees, shown in Figure 3.1. This is done in order to test

the algorithms without excessive computation burden. Excluding any galaxies outside this region yields approximately 100,000 galaxies.

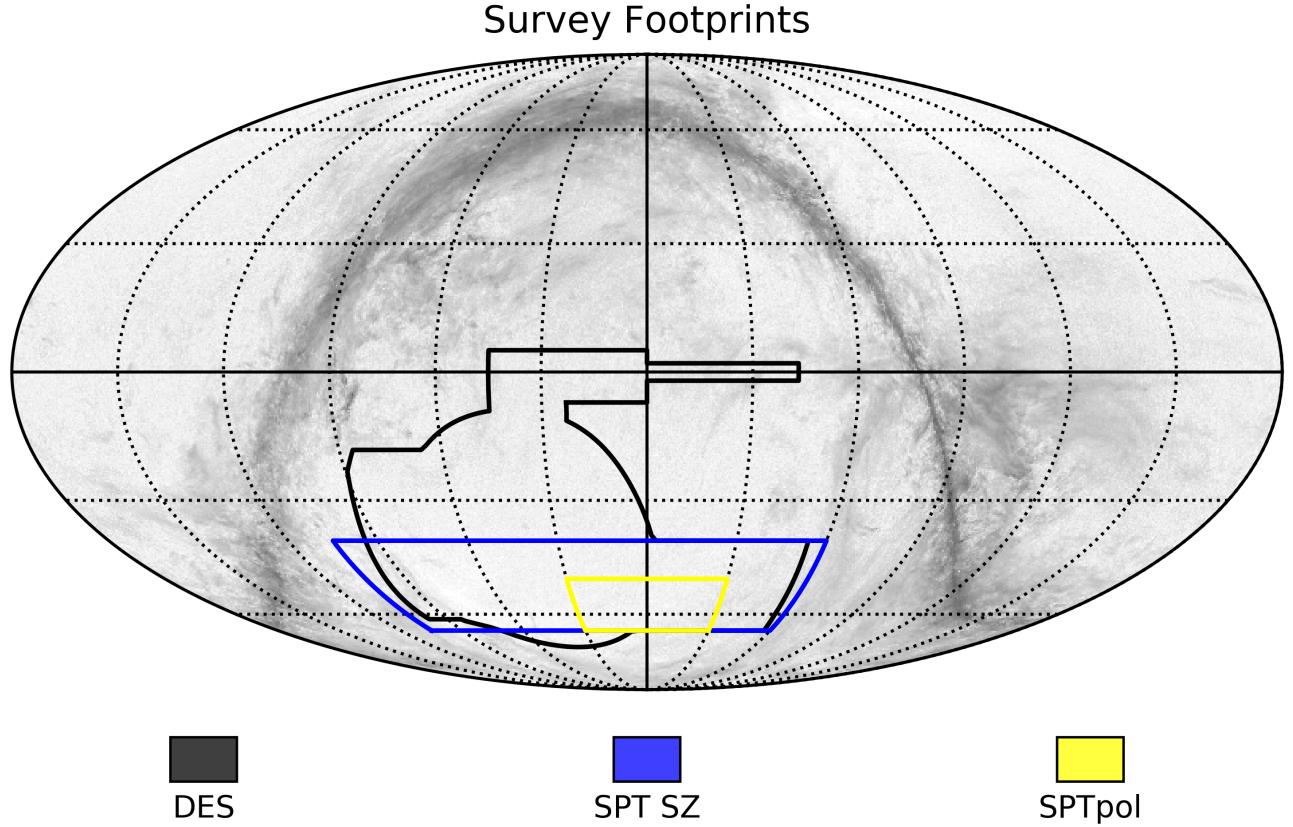


Figure 3.1: Outlines of SPTpol, SPT-SZ, and DES surveys

Pairs were generated by making use of kD-trees. A generalisation of a binary tree, the kD tree is one where every leaf node is representative of a  $k$  dimensional point. Each non-leaf node is one that 'splits' the space into two parts, whereby points to the left of this hyperplane are represented by the left sub-tree, and to the right are represented by the right sub-tree. When we apply this to our galaxy catalogue, we have to locate any two galaxies which satisfy certain conditions regarding line of sight, and transverse separations. All galaxies in the footprint that is being considered are then placed in a kD tree, and any pair which has a radial comoving separation of less than  $20h^{-1}$  Mpc, and transverse comoving separation range of  $4 - 20h^{-1}$  Mpc is considered to have a filament (Clampitt et al., 2016; Colberg et al., 2005).

### Data Preparation Algorithm

The algorithm for producing the catalogue of galaxy pairs is as follows:

1. Convert DES Right Acensions to usable format.
2. Remove any galaxies that are not contained in a given footprint
3. Store Locations of Galaxies (as Right Acensions and Declinations) in kD-Tree.

4. Search kD-Tree for any point within some chosen distance from a given galaxy (e.g. within  $\sim 1\text{deg}^2$ ).
5. Construct a list of pairs which satisfy this approximate distance condition .
6. Convert DES Photometric Redshifts into Comoving Distances.
7. Calculate the difference in the comoving distance between a given pair.
8. Make cuts to catalogue of galaxy pairs based on chosen line-of-sight separation condition.
9. Compute the mean redshift of a given pair
10. Compute the actual angular separation of a given pair
11. Use 9 to determine the equivalent transverse comoving distance at the redshift given by 8.
12. Make cuts to catalogue of galaxy pairs base on chosen transverse separation conditions.

Step 1 is necessary, because there are multiple ways to report one of the two coordinates for sky position, Right Ascension (RA), and in order for the kD-Tree algorithm to work, they need to fall into a specific format. They can either be expressed on the domain from  $[0, 360)$ , or  $(-180, 180]$ , and the data as obtained reports the RAs of galaxies in the first format, whilst we need the second. If we tried to run the algorithm on the original data, we would inadvertently exclude any pairs in one half of the footprint with step 2. It would also prevent pairs from being formed in the left half of the image, since the domain for sky locations would now be discontinuous, and so would register two galaxies which may be right next to each other as having a separation of 360 degrees.

Step 2 ensures that we are not including galaxies which are not contained in the relevant footprint. We also have to be careful to avoid edge effects in the  $y$ -map. When constructing CMB data products, the edges of the map typically have significantly more noise, since there are fewer passes made of the edges in a given observing run. Typically these pixels are de-weighted when compared to pixels in the centre of the map, but if sufficient observing run maps are co-added together, the effects of the noise should be mitigated. The map we are using has been made by combining two independent data sources, the SPT-SZ, and Planck datasets. This should serve to appropriately reduce edge effects, but it is still something to keep in mind.

Step 3 and step 4 are important, because the only way that the kD-Tree functions is if there is some ‘distance’ in kD space from which an initial calculation can be defined. The algorithm starts at a given node and searches down the tree until it moves beyond its defined ‘distance’, then it constructs a pair of every node which falls below this condition. These steps do make an assumption that on small enough scales, the sky is flat, which isn’t true in a general sense, but is necessary to construct an initial set of pairs.

Step 6 follows the process outlined by Hogg (1999), where the comoving distance at a given redshift is given by

$$D_C = D_H \int_0^z \frac{dz'}{E(z')} \quad (3.1)$$

where  $E(z') = \sqrt{\Omega_M(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda}$  defines the scale factor at a given redshift, and  $D_H \equiv \frac{c}{H_0}$  is the Hubble distance. The scale factor depends on  $\Omega_M$ ,  $\Omega_\Lambda$ , the dimensionless density parameters describing the proportion of matter and dark energy in the universe, and  $\Omega_k = 1 -$

$\Omega_M - \Omega_\Lambda$ , which describes the curvature of the universe. This calculation takes into account the curvature effects of the various constituents of the universe, as well as the changing scale factor as you look back.

This calculation allows us to make the rather significant cuts to the list of potential pairs quite early. The size of the dataset means that performing all operations on all possible pairs would be very computationally expensive, so making the line of sight cuts early is preferred.

Computing the transverse separation takes a little bit more computation effort, because our earlier assumption of sky-flatness cannot be taken as true. The separation between two events at the same redshift, but separated by some angle on the sky  $\delta\theta$  is given by  $D_M\delta\theta$ , where  $D_M$  is the comoving transverse distance at the redshift in question. For a flat universe ( $\Omega_k = 0$ ), the comoving transverse distance is equivalent to the comoving line of sight distance (i.e.  $D_M = D_C$ ), but this number is fundamentally dependent on the chosen cosmology. Because of the changing scale factor, objects at different redshifts with the same size will have different angular extents, so this must be taken into account. We do have to make one rather significant approximation here however, because it is very unlikely that both galaxies in a pair might be at the same redshift. As such, we have to approximate their transverse separation at the mean redshift of a given pair. While this approximation is not entirely accurate, it is sub-dominant to the uncertainty on photometric redshifts provided by DES.

## Data Analysis Algorithm

Functionally, this algorithm follows some elementary primary operations:

1. A pair is located in the CMB
2. A slice is taken around the pair
3. The slice is then rotated, so that the pair is aligned along the same axis
4. The slice is then rescaled, so that each element of the pair is situated at the same point in pixel space
5. These slices are then mirrored, in both the  $x$  and  $y$  axes separately and then together, and the average is taken over all four mirror images
6. These averages are then co-added, and the average is taken over the number of pairs added together

We perform operations 4 and 5 because we make the assumption that the dark matter halos that host the LRGs we are considering are spherically symmetric. Mirroring the halos about two axes essentially allows us to stack multiple versions of each individual halo. Doing so reduces irregularities introduced by any halos that happen to be non-symmetric, whilst correlating their spherical structure.

Steps 5 and 6 are done by applying an affine transformation to the array of values, because we are seeking to preserve the functional position of all points, straight lines, and ratios in the array, as close to the original as possible. This transformation effectively involves applying a transformation matrix to a given image or array, and applying a spline interpolation if there isn't a perfect transformation between the initial and final state.

We tested this algorithm by creating mock simulated datasets, where an artificial signal inserted in noise that is 3 to 4 orders of magnitude higher than it.

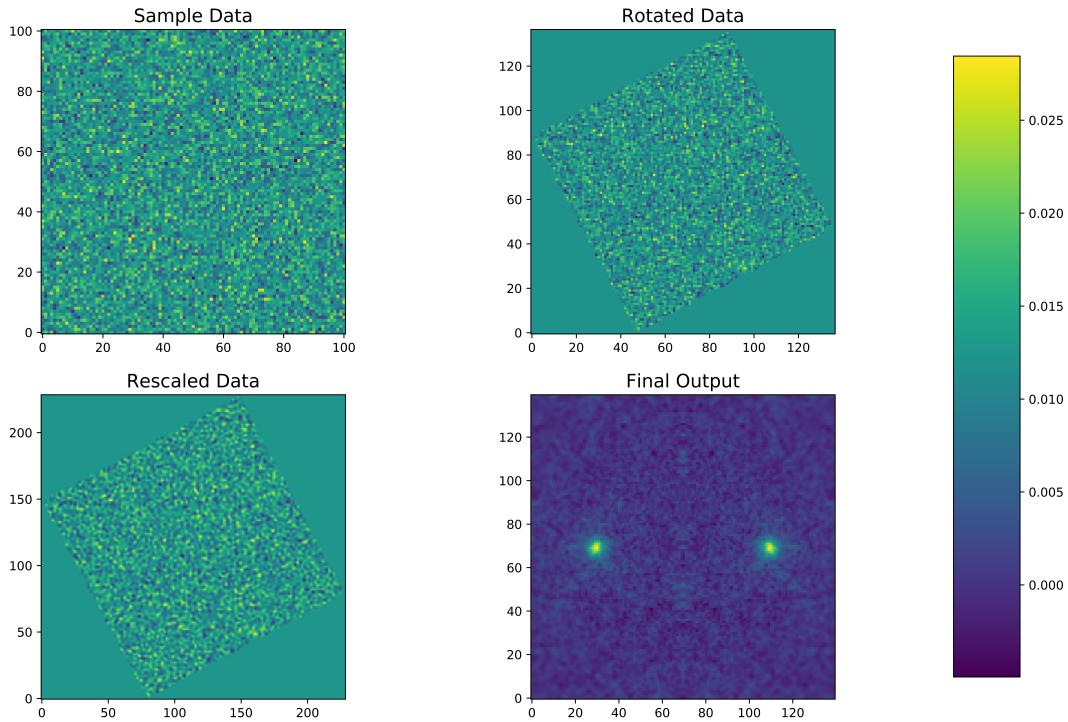


Figure 3.2: Stacking Algorithm as applied to Dummy Data. Top Left: Slice is taken of a given pair, Top Right: Slice is rotated, aligning both elements of the pair with the  $x$  axis, Bottom Left: Slice is rescaled. Axes scales have been left in to illustrate the relative scaling between Top Right and Bottom Left, Bottom Right: Second slice is taken of the manipulated data

As can be clearly seen, there appears to be no signal in the individual slices of data, but when they are coadded and mirrored, the signals clearly appear as two bright halos.

Some things to note about this method. First, there are some interesting artefacts along the lines of symmetry. Because both signal and noise are getting mirrored about two axes, any data that lies along one of these axes will get flipped, but will not move in place. This introduces some measure of artificial correlation, because any noise that lies along these axes will look like it is correlated with the un-mirrored data. This is clear in artefacts that lie along the vertical axis. Secondly, it is sensitive to the separation of the galaxy pair in pixel space. If a galaxy is relatively close together, the rescaling operation will stretch out the pair until they are located at the normalised positions chosen. This also has the effect of stretching out their halo as a whole. This means that pairs that are more separated in pixel space will be rescaled less, and so the halos will be more compact, and pairs that have a lower separation will be rescaled more, and so have a larger angular extent after rescaling. This essentially means that the halo sizes will not be evenly distributed, they will follow the distribution of the pair separations on the sky.

Once we have the pairs co-added, we need to subtract the halo contribution. There are a number of possible ways to do this. We can analytically solve for the halo contribution in the map

by using a model of the form

$$y_h(p) = y_{L,i} + y_{R,j}, \quad (3.2)$$

where  $p$  represents a given pixel,  $L$  and  $R$  indicate the left and right halos respectively, and  $i$  is the  $i$ th radial bin from the centre of the left halo, and  $j$  is the  $j$ th radial bin from the centre of the right halo.

We now have to consider our model choice. Naively, we can model the halos simply by computing the radial profile of one of the halos, and excluding regions which will have a minimal contribution from the secondary halo. We can then assume some background level, and subtract that model away from our signal.

If we want to truly take into account both halo contributions at a given pixel however, we need to calculate the model in a 2D space. There are two primary models that we can consider here, with a number of possible variations. Given that we have assumed that the halos are mostly spherically symmetric, it follows that they may take the form of Gaussian profiles, so we can consider a model which places two gaussians in the positions of the two halos.

The second model we consider is a fourth order polynomial model, with some form of suppression. This second model takes into account the possibility that the functional form of the halos needs to decay faster than a polynomial ordinarily would, but allows for some flexibility of shape that is not afforded by the gaussian model.

We initially consider the models in one dimension, along the effective centre of the galaxy halo pairs. This makes sense if the assumption holds that the galaxy halos are spherical to some leading order. If this assumption doesn't hold for any reason, either physical, or one introduced by our algorithm, we should consider performing our fit over the entirety of the two dimensional image produced by the stack. Whether this will be effective is ultimately dependant on the level of symmetry in the stack, and how well the two dimensional fit performs. In order to avoid introducing bias from the potential filament, we exclude the central region of the stack, as shown in Figure ??

Once these models have been fitted, and subtracted from our detected signal appropriately, we hope to be left with a filament detection. This will have an average comptonisation, and so will act as a means to measure the amount of matter contained within the WHIM filaments connecting galaxy pairs.

## 3.1 Null Tests

In order to test the validity of the detection, and make an estimate of its uncertainty, we perform several different types of Monte-Carlo based tests.

### 3.1.1 Un-Physical Pairs

The first test involves stacking the  $y$  map against 'pseudo-pairs' of galaxies. These pairs satisfy the transverse separation condition, but do not satisfy the radial separation condition, having instead radial separations between  $100 h^{-1}$  and  $200 h^{-1}$  Mpc. Pairs with these conditions are not expected to have any filament connecting them, and the radial separation conditions have been calculated to take into account the errors associated with the photometric redshifts found in the Dark Energy Survey catalogue.

In order to produce this pair set, we first calculate the total number of pairs that sit within

approximately  $1^\circ$  on the sky, and calculate their line of sight separations. We then determine the transverse separation of the pair, and apply the same cuts as for the physical pairs.

Once these pairs are stacked, they should show a map that is similar to the physical pairs, but with visibly less signal between the two halos, where we expect to see the signal.

If we perform the same halo fit as for the physical pairs, we should be able to subtract the halo contributions for this dataset and see no visible filament signal.

### 3.1.2 Random Stack

The second test involves creating a random stack of CMB slices, to estimate the RMS of the background. At first glance, this would suggest just randomly sampling points inside the footprint of both surveys. However, because there is contamination in the CMB, in the form of various foregrounds, we have to ensure that the noise introduced by this is going to be the same as for the physical pairs. Given that the noise in the CMB is approximately constant at a given galactic latitude, we can therefore randomly select a point that is at the same galactic latitude as a given pair, but with some random longitude. We would expect, given that the fluctuations in the CMB are approximately gaussian, that this stack would have no discernable structure.

# Chapter 4

## Results

From the process outlined above, 787,058 galaxy pairs were constructed with a mean angular separation of  $\sim 27.6$  arcmins, and a mean comoving separation of  $11.9 h^{-1}$  Mpc.

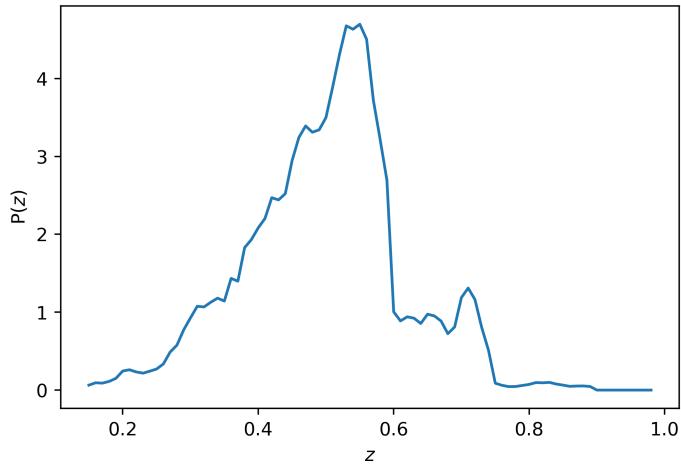


Figure 4.1: Redshift Distributions of Physical Pairs. The pairs range from redshifts  $z \approx 0.15$  to  $z \approx 0.90$ , with a mean redshift of  $z \approx 0.50$ .

Figure 4.1 shows the overall PDF of galaxy pairs as a function of redshift. It shows that the mean redshift for the pairs is  $z = 0.468$ , with a minimum redshift of  $z = 0.15$ , and a maximum redshift of  $z = 0.90$ . There is a rather drastic drop in the galaxy population after redshift  $z \sim 0.58$ . This is due to there being fewer galaxies in the higher redshift bins in the DES Year 1 Catalogue. Figures 4.2 and 4.3 show the line-of-sight and transverse separation PDFs of the pairs.

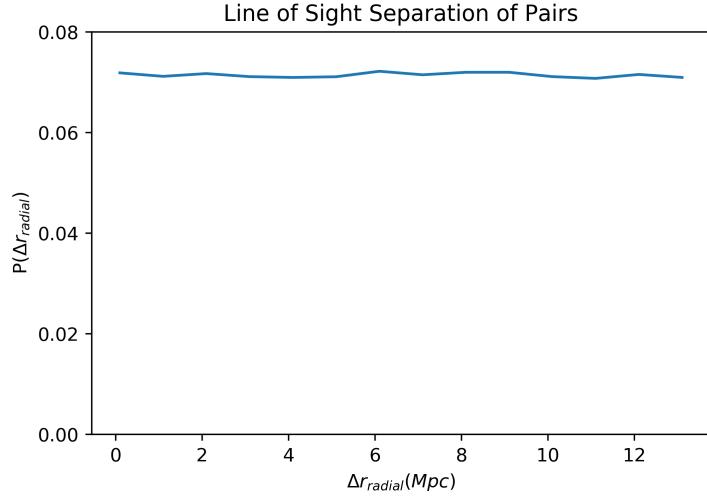


Figure 4.2: Histogram of Line of Sight Separations of Galaxy Pairs. The distribution is relatively flat, with a minimum separation of 0 Mpc, and a maximum separation of 14.8 Mpc.

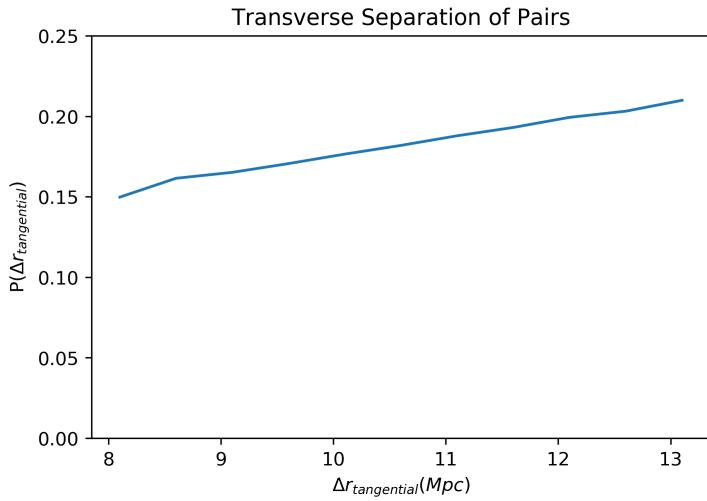


Figure 4.3: Histogram of Transverse Separations of Galaxy Pairs. The distribution climbs steadily, with a minimum separation of 8.85 Mpc and a maximum separation 20.7 Mpc.

There are more pairs in higher transverse separation bins than in lower ones, which suggests that there will be a tendency to rescale pairs by smaller amounts in the algorithm. This will also effect the resulting halo shape, since pairs that are closer together will also by nature be scaled to larger effective halo sizes.

The stacking procedure was performed on a Compton  $y$ -map produced from a combination of the South Pole Telescope SZ observing run, and the *Planck* datasets. It made use of the 90 GHz, 150 GHz, and 220 GHz maps from SPT-SZ, and combined them with the 100 GHz – 350 GHz and dust maps from *Planck*, in the same manner as is described in Crawford et al. (2016) (Bleem, in prep.). The algorithm for producing this map also took the half survey and half mission power spectra from *Planck* as inputs. It minimised the contribution of the primary CMB, Cosmic Infrared

Background (CIB), Instrumental, and Atmospheric sources as the primary sources of noise.

We performed the stack on a map with a resolution of  $0.25'$  per pixel, but the effective resolution, after combining the beam sizes of the various raw data maps, is closer to  $2\text{ arcmin}$ , so there is likely some interpolation in the output, which would have introduced a source of noise.

Stacking these pairs returns an average  $y$  map, which is shown in figure 4.4. The signal is dominated by the contribution of the galaxy halos, and so these will need to be effectively subtracted in order to evaluate the significance of the filament signal. To the eye, there appears to be some residual signal in between the two halos, which has not been driven to zero as a result of the stacking process.

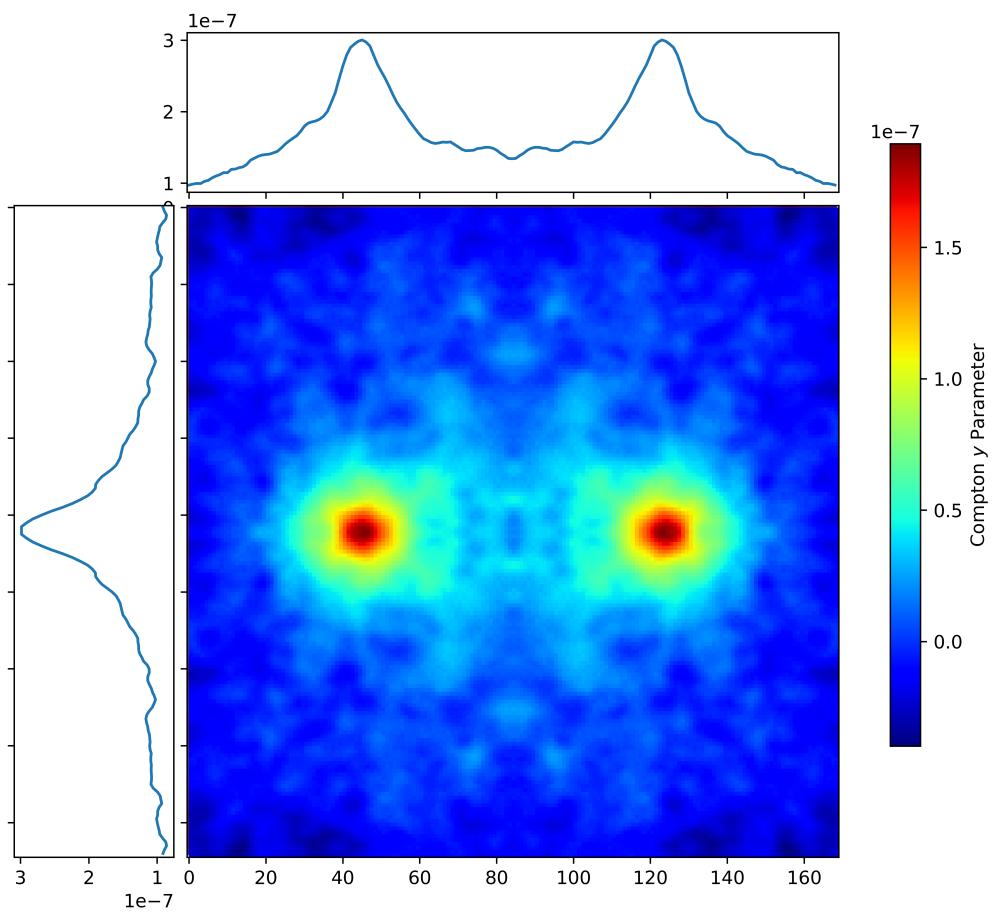


Figure 4.4: Stacked Image of galaxy pairs. The upper panel shows the slice through the centre of the stack. The left panel displays a vertical slice through the left halo, which by the mirroring procedure, will be the same as the slice through the right halo.

Once we have stacked the pairs, we perform our fit, as described in 3. Beginning with our naive fit, we considered only a single halo, because the mirroring of the image essentially forces the two halos to look identical under this calculation. We also only consider the half of the halo that is on the other side to the filament, to prevent contamination from the filament. Doing so yields a halo shape shown in Figure 4.5.

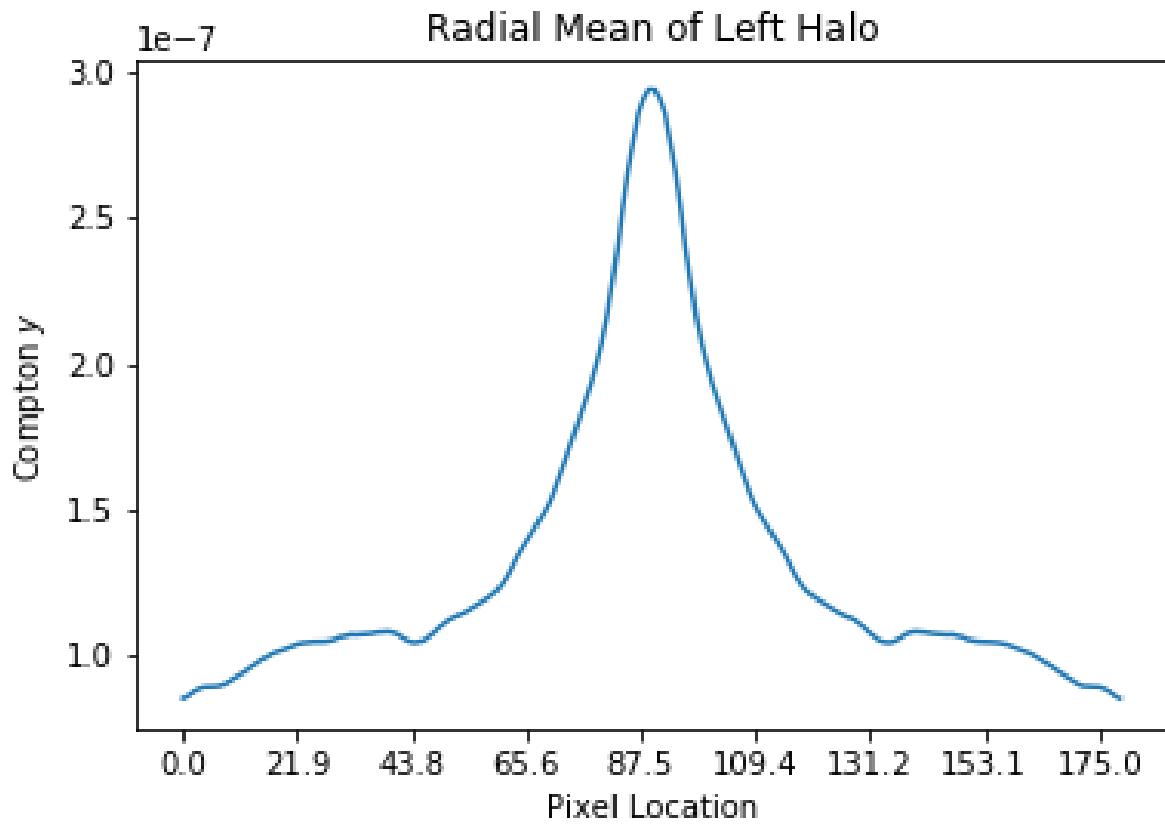


Figure 4.5: Radial Mean of Left Halo from figure 4.4. The shape of the halo does not appear to be gaussian in nature, perhaps owing to the uneven scaling depending on transverse separations.

Taking this halo shape, and assuming that there is some constant background signal of approximately  $y = 1 \times 10^{-7}$ , we can combine this into our naive model, and find a measure of the residual filament. Assuming this constant background signal is functionally the same as taking a high bandpass filter of the stack.

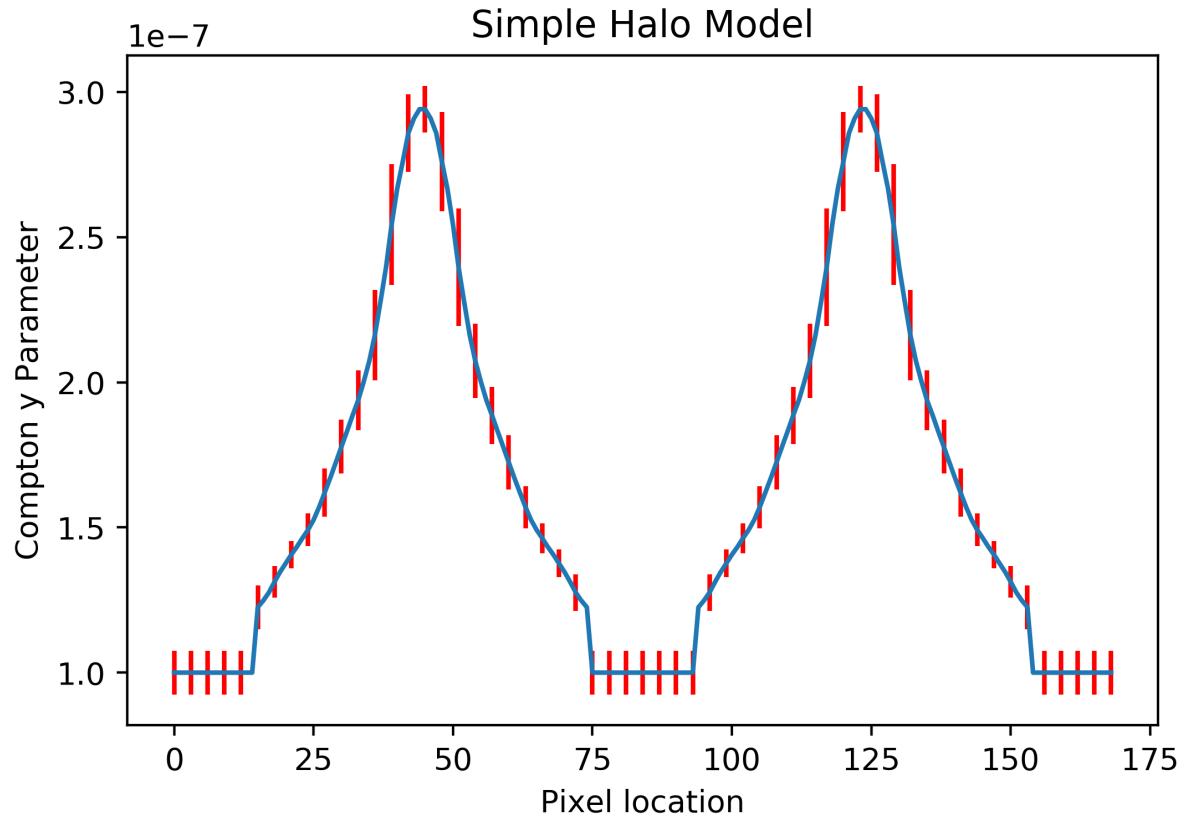


Figure 4.6: Basic Halo Model, constructed from the radial mean of a single halo, with error bars given by the standard deviation of the halo in the same radial bins

Figure 4.6 shows this model, with errors derived from computing the standard deviation of the radial halo in the same halo bins as the mean was calculated. Subtracting this from the slice through the halos gives us the graph shown in figure 4.7.

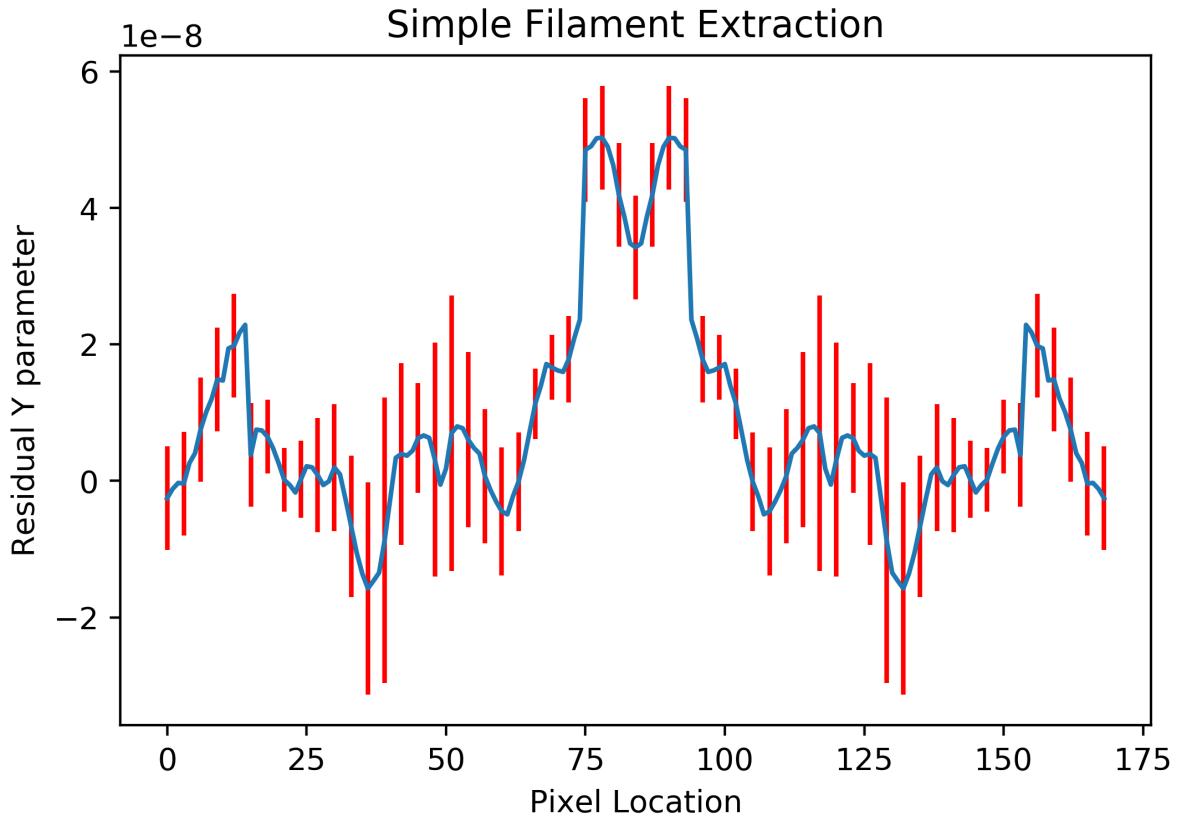


Figure 4.7: Basic Filament Model, with filamentary residual shown between Pixel locations 60 and 110.

This naive calculation shows what is visible to the eye when looking at the stack in figure 4.4, that there is some residual signal in between the two halos which should be from the filaments that connect the galaxy halos. The mean compton y parameter for this region is  $\bar{y} \approx 1.9 \times 10^{-8}$ , with a mean error of  $8.31 \times 10^{-9}$ .

If we consider our more complex models in 1 dimension (along the central slice through both halos), and fit to our data, we get the results shown in figure 4.8. All models were fit using standard definitions, and using a damped least-squares method known as the Levenberg-Marquardt algorithm.

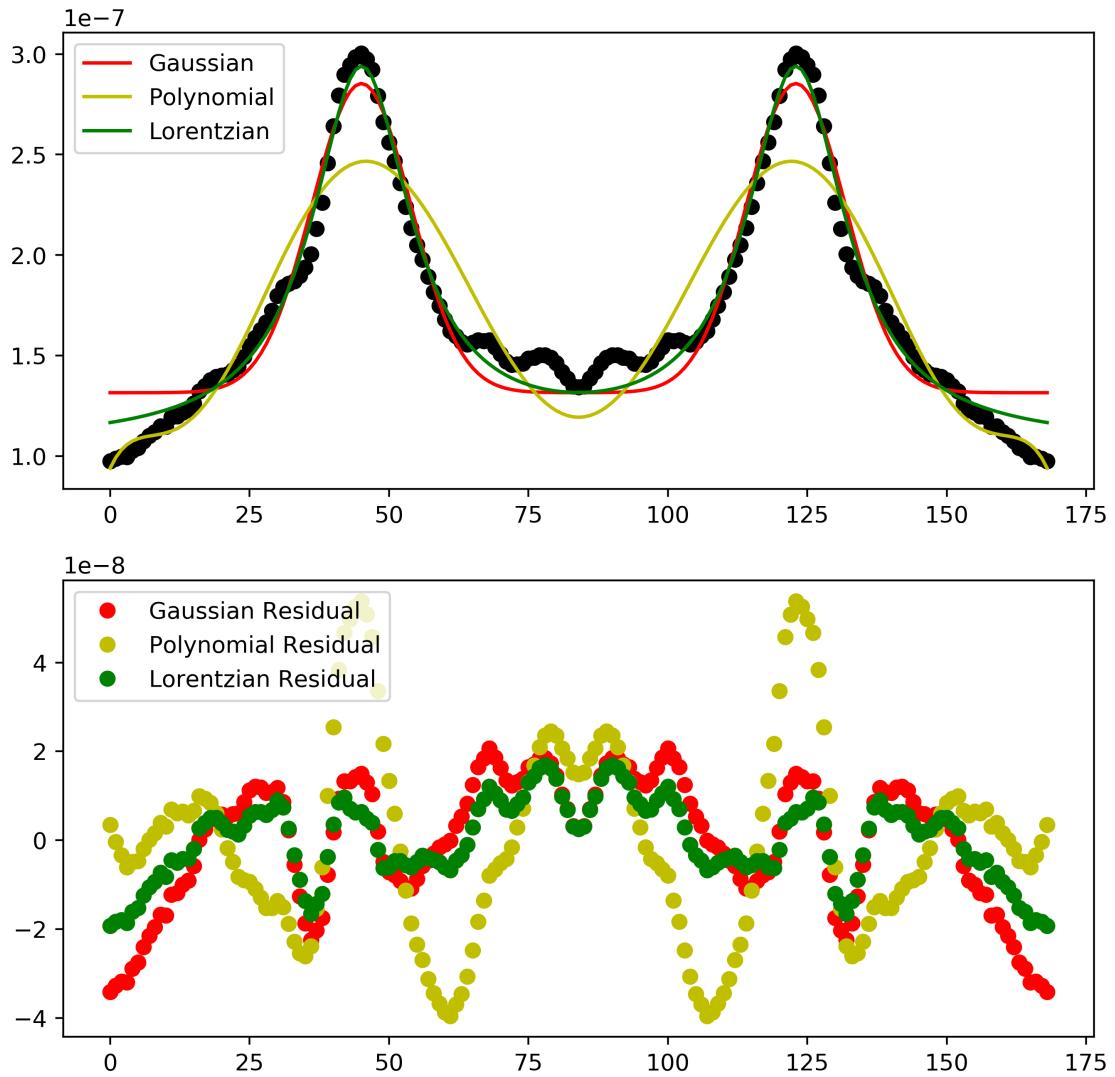


Figure 4.8: Multiple Halo Models fitted to Central Slice of Figure 4.4

These are noticeably worse than our simple model. The polynomial fit doesn't subtract the halos properly, so we discard it here. The gaussian model does subtract most of the halos away, but the left-over signal is significantly lower, with a mean residual Compton  $y$  of  $\tilde{y} \approx 8.68 \times 10^{-9}$ . The fact that the gaussian models fail to match the curves at the centres of the halos does suggest that it is a poor model.

Given the apparent shape was non-gaussian, we applied a Lorentzian fit to the shape of the halos. This fit the halo shapes much better, and dies away much quicker than the gaussian model, but also removes more of the signal we are looking for. It is possibly that this is simply the best fit for the algorithm we have implemented, and a change in our process would result in another model

fitting better. It is also incredibly difficult to explain why the halo shape would be lorentzian physically, further reinforcing the idea that it is the result of introduced scaling effects in the algorithm.

Considering that the model of the halo seems to have a non-gaussianity to it, we performed a number of two dimensional model fits to our two dimensional image of the stacks. The models we chose to fit to the data were

(I) 4 Independent Gaussians

(II) 2 sets of 2 coupled Gaussians

(III) 4th order polynomial

(IV) Spherical Gaussians

We chose to use variations of multiple gaussians in order to take into account the distorted shape of the halos. In model (I), we allow 4 gaussians to vary independantly. In model (II), we couple the underlying gaussians together, so that one pair of them will extract the spread of the halo, and another will extract the height. We chose model (III) to maintain consistency with our fit from the simple 1D case. In model (IV), we force the gaussians to contain our assumption of spherical halos.

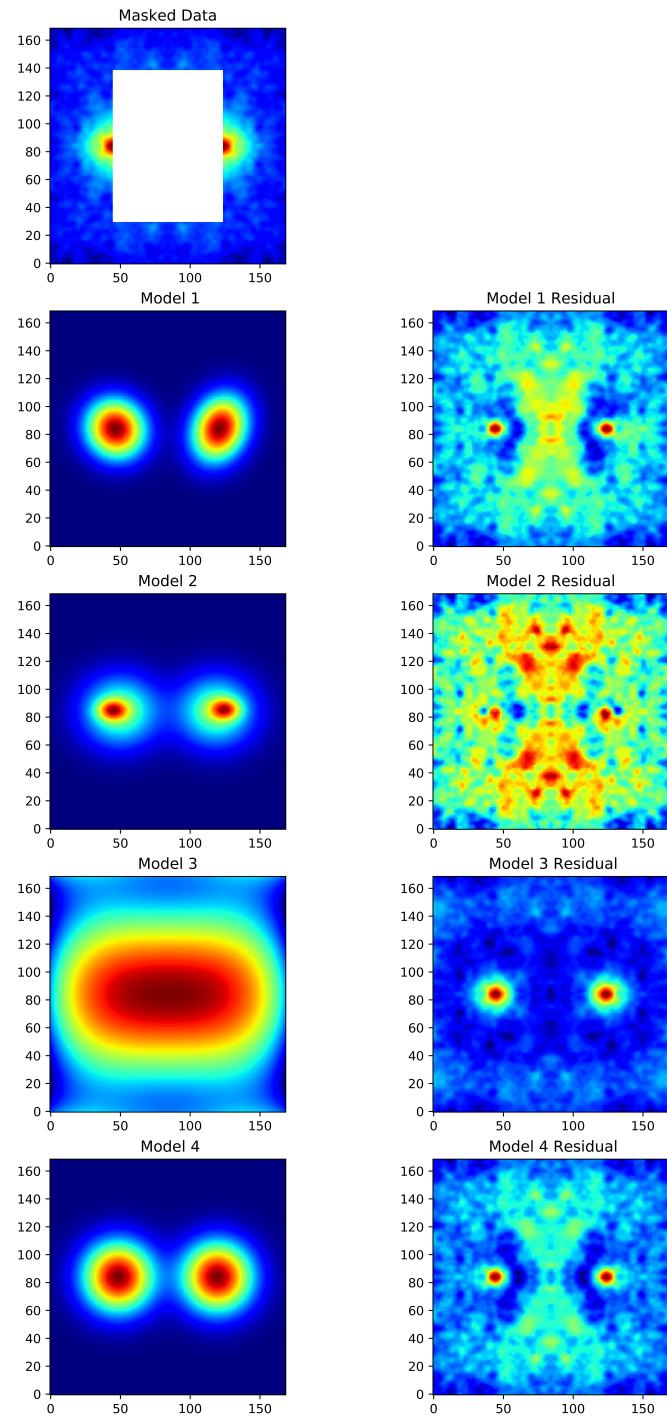


Figure 4.9: Multiple Halo Models fitted to 2D array of Figure 4.4

When we look at these models, none of them seem to appropriately subtract the halo contributions.

Model (I) appears to do the best, removing the majority of the halo, and leaving behind a reasonably strong signal where we would expect to see a filament. What is noticeable however, is that the right halo in the model is distorted along some diagonal axis, which is strange, given the data that is being fitted to has been made symmetric about two axes. This results in greater subtraction in the residual of the right halo than the left, and so introduces some error in our measurement of the filament. It also leaves a very large amount of signal in the central region, in what looks like a hourglass shape. It is unclear if this is indicative that filaments don't necessarily link galaxies in a straight line, or if it is an artefact of the mirroring process.

If we correct for the perceived issues with the first model, model (II) forces the gaussians into sets of two, which each obey the same statistics, therefore accounting for the mirroring in the image. This form of the model does succeed in extracting nearly all of the halo, but leaves so much residual signal everywhere that it looks like we are extracting residual baryons from areas surrounding the galactic halos.

Unfortunately, (III) extracts too much of the signal, and not enough of the halos. This may be due to the way that the model is being fit to the data. Since the data is symmetric about the central  $x$  axis, the model fitting will see the same value on both sides of the masked area. This forces the fit to assume that the data holds the same value across the masked area, removing the possibility that the halos fall away, and so preventing us from detecting the filament signal. If we perform the fit on the whole array of data, the polynomial overfits the data, and subtracts the entirety of the filament signal anyway. The current fit doesn't subtract the whole halo signal either, meaning that the data isn't being fit properly with the current mask.

Model (IV) also extracts too much of the filament signal. It appears that forcing the spherical shape of the halos causes too much crossover in the modelled halos. This effectively leaves us with no filament signal, allowing us to discard this model. Doing so does raise some interesting points however, regarding our assumption of spherical symmetry.

These two dimensional models do not perform as well as the one dimensional ones, either in terms of accuracy or strength of signal. This is likely due to the increased amount of noise in the two dimensional image, as opposed to the one dimensional slice. The higher resolution in the  $y$  map we use could also be introducing this higher level of noise, since it would not necessarily be smoothed out by the detector beam.

The results for all models which are considered sensible are included in Table 4.1.

Model	Mean Residual $y$	Error	Sigma
Radial Mean 1D Fit	$2.5 \times 10^{-8}$	$\pm 7.72 \times 10^{-9}$	3.24
1D Gaussian Fit	$1.29 \times 10^{-8}$	$\pm 6.29 \times 10^{-9}$	2.05
2D Model 1	$1.38 \times 10^{-8}$	$\pm 1.77 \times 10^{-8}$	-
2D Model 2	$0.82 \times 10^{-9}$	$\pm 1.59 \times 10^{-8}$	-

Table 4.1: Results for residual mean filament for various fitting methods

Given the necessity of accounting for both halo contributions at all points in the stack, we hold that the best result to quote is the 1D gaussian fit. Our fiducial value of the average Compton- $y$

of the filament is

$$\bar{y} = 1.29 \times 10^{-8} \pm 6.29 \times 10^{-9} \quad (4.1)$$

## 4.1 Null Tests

### 4.1.1 Un-Physical Pairs

We constructed a catalogue of unphysical pairs from the galaxies contained in both the SPTpol and DES survey footprints, with radial separations between  $100 h^{-1}$  Mpc and  $200 h^{-1}$  Mpc, and transverse separations between  $6 h^{-1}$  Mpc and  $14 h^{-1}$  Mpc.

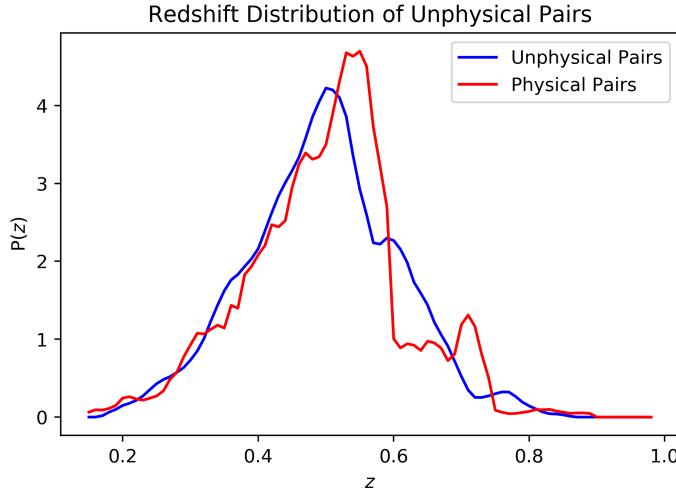


Figure 4.10: Redshift Distributions of Unphysical Pairs. The pairs range from redshifts  $z \approx 0.24$  to  $z \approx 0.87$ , with a mean redshift of  $z \approx 0.55$ . When compared to the physical pairs, this set has a higher mean redshift, and more pairs in higher redshift bins. . The distribution of physical pairs is overlaid in red.

As can be seen in figure 4.10, the distribution of redshifts is generally skewed towards higher redshifts than for the physical pairs. This is likely due to the inherent errors associated with the photometric redshifts in the DES redMaGiC survey. Lower redshifts are more likely to be more precise, since the photometric errors for redshift are given by  $\Delta z = 0.01(1+z)$ , and so be excluded by the line of sight cuts. The other effect which may influence this is the fact that there are far more pairs in the unphysical dataset than in the physical dataset.

This produced a set of unphysical pairs containing  $\sim 670,000$  pairs, with a mean line of sight separation of  $220$  Mpc, and a mean transverse separation of  $12.5$  Mpc.

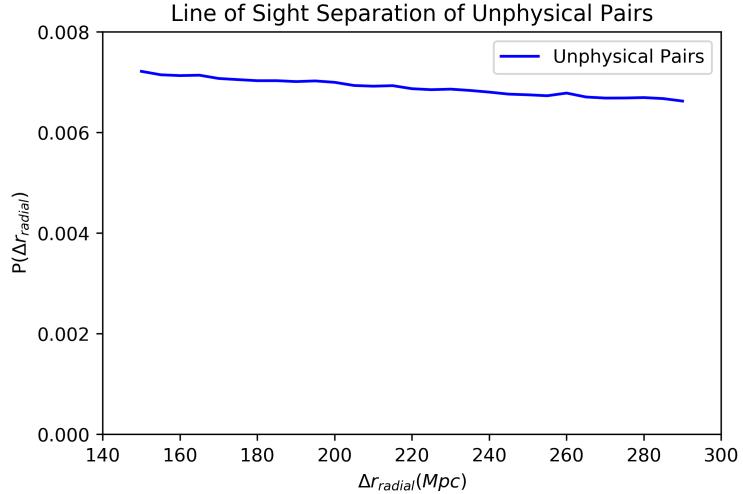


Figure 4.11: Histogram of Line of Sight Separations of Unphysical Galaxy Pairs. The distribution is relatively flat, with the same shape as the physical pairs, a minimum separation of 147 Mpc, and a maximum separation of 295 Mpc.

The choice was made to consider unphysical pairs with a line of sight separation in excess of  $100 h^{-1}$  Mpc because at large redshifts, the errors associated with the photometric redshifts can place a very large range on the possible distance to a given galaxy, and we want to be very careful to exclude galaxy pairs that might have a connecting filament between them.

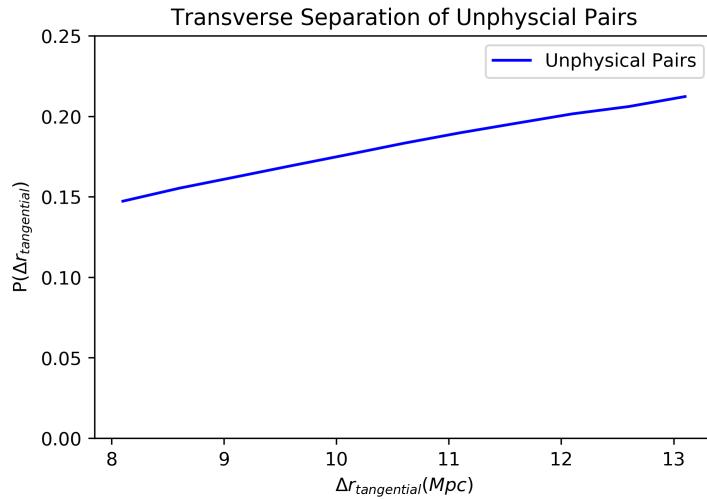


Figure 4.12: Histogram of Transverse Separations of Unphysical Galaxy Pairs. The distribution is relatively flat, with a minimum separation of 8.85 Mpc and a maximum separation 20.7 Mpc. This has the same overall shape as the physical pairs data set, except the relative population in the higher separation bins decreases, rather than increases.

Performing the stacking procedure on the unphysical dataset gives the stack shown in figure 4.13. To the eye, in both the 2D slice and the 3D colour map, it appears that there is less filamentary signal than for the physical pairs.

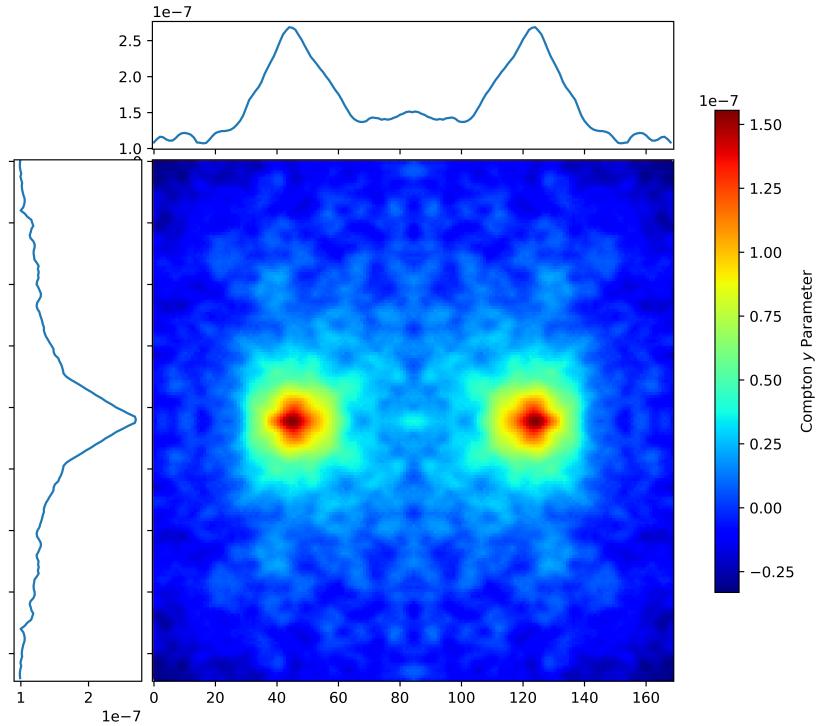


Figure 4.13: Stacked Image of Unphysical pairs.

If we fit the two halos for gaussian profiles (shown in figure 4.7) , along with some constant offset, we can see very clearly that there appears to be no residual signal left. The mean of this residual is  $3.48 \times 10^{-18}$ , with a variance of  $1.13 \times 10^{-16}$ , which is consistent with a zero measurement.

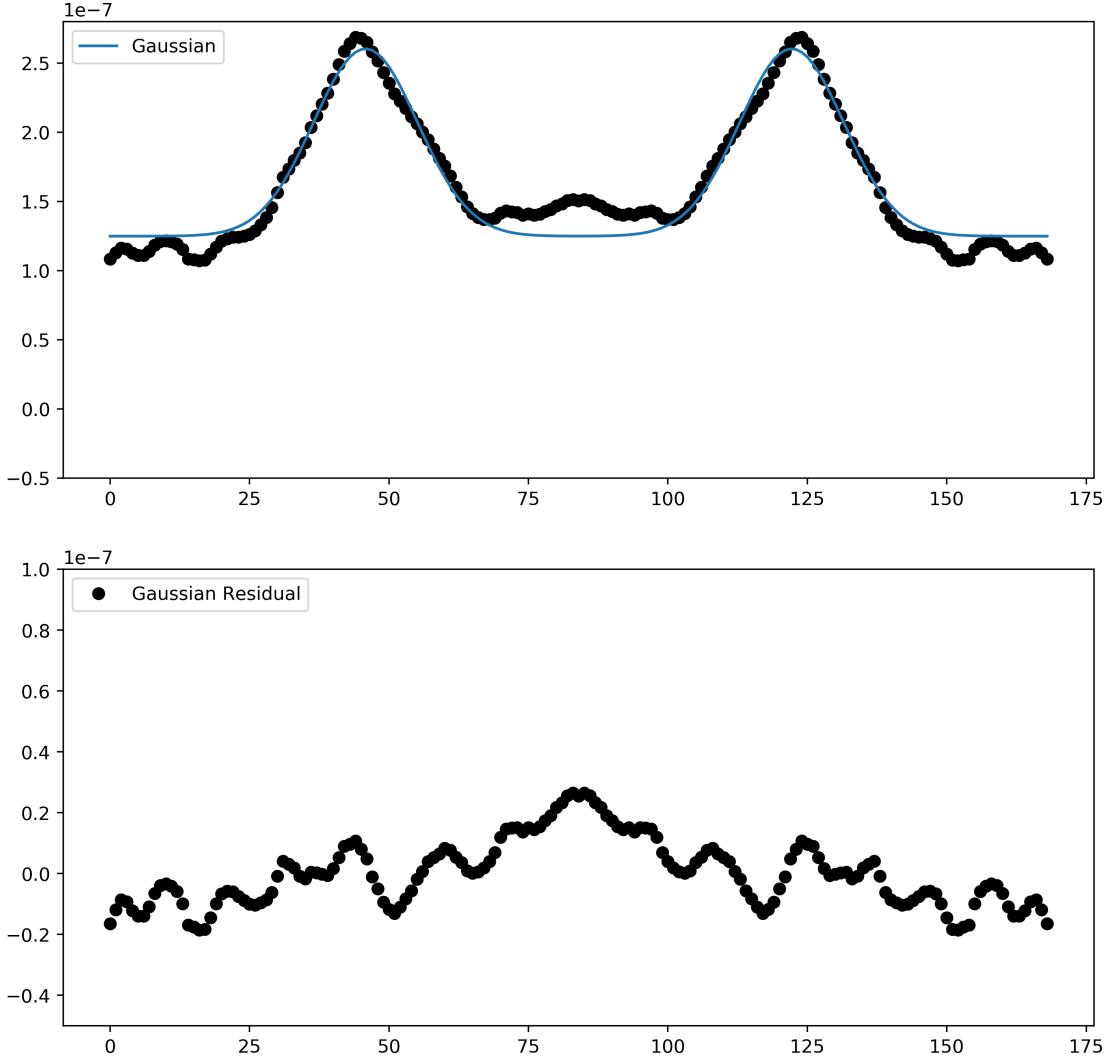


Figure 4.14: Gaussian Fit to Unphysical Pairs

#### 4.1.2 Random Stack

Performing a stack on a set of pseudo-randomly selected slices of the CMB, with the same galactic latitude as the physical pairs, we produced a stack like the one shown in Figure 4.15. This has no discernable structure, except for the small circular signal in the centre of the map. This is likely due to the rescaling effect still rescaling the CMB as if it were trying to align pairs. Because the angle that the pairs need to be rotated by should be evenly distributed, there should be some level of correlation in the stack, in a circular structure where we would otherwise expect a filament to be.

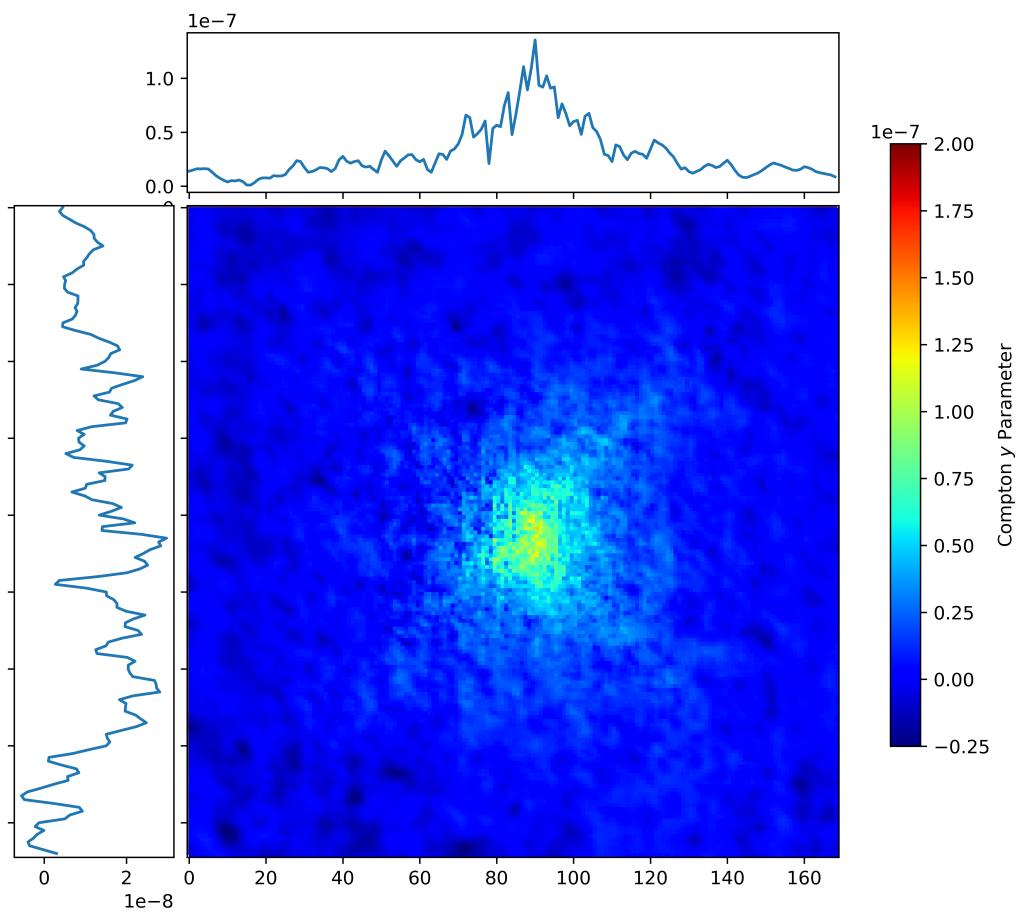


Figure 4.15: Stack of pseudo-random slices

# Chapter 5

## Discussion and Conclusion

### 5.1 Discussion

Our measure of the Compton  $y$  parameter in the filament of  $\bar{y} = 1.29 \times 10^{-8}$  at  $2.05\sigma$ , is comparable to similar work done by Tanimura et al. (2019) (referred to as T17), and by de Graaff et al. (2019) (referred to as G17).

T17 found 262,864 galaxy pairs at redshifts  $z < 0.4$  using the Sloan Digital Sky Survey luminous red galaxy catalogue, detecting a mean Compton parameter of  $y \approx 1 \times 10^{-8}$  in the *Planck*  $y$  map at a  $5.3\sigma$  confidence level.

G17 found 1,020,334 galaxy pairs in a redshift range  $0.43 < z < 0.75$  using the Sloan Digital Sky Survey CMASS catalogue, detecting a mean Compton parameter of  $y \approx 0.6 \times 10^{-8}$  in the *Planck*  $y$  map at the  $5.1\sigma$  confidence level.

The two catalogues used by T17 and G17 are both independent of each other, since they cover different redshift ranges, but the similar detection levels, and detection strengths lends strong evidence to the existence of filaments which can be detected by tracers in the CMB.

G17 directly calculated the baryon fraction from their residual Compton  $y$  parameter as being  $\sim 0.3\Omega_b$ . Our measurement of the Compton  $y$  parameter is much higher than theirs, but there are a number of factors which differentiate our measurement from previous work.

Previous calculations operated on the *Planck*  $y$  map only, which has an effective beam size of  $10'$ , which effectively acts to convolve any signal contained in the map. This introduces some error into the measurement, but mostly acts to smooth out the  $y$  map. The choice has to be made then, to not consider pairs that are too close to each other, because if they fall within the width of the beam, they will be functionally indistinguishable, and it will be impossible to separate the contributions from their halos and the filament. The  $y$  map we are using was produced primarily from the *SPT-SZ* detection maps, and has a significantly smaller effective beam size than the *Planck* map alone, of  $2'$ . This means that we can make a choice to include more close pairs than previous work, because they will be able to be effectively resolved in our  $y$  map.

There are also a number of significant differences between galaxy catalogues used by previous works, and this one. Both T17 and G17 made use of galaxy surveys with spectroscopic redshifts, as opposed to photometric redshifts. Spectroscopic redshifts are taken by directly measuring spectra for a given galaxy, and identifying known spectral features in it. This makes it very easy to get an accurate measure of the redshift of a galaxy, but the process of producing a spectrum takes a long

time, since it requires receiving enough photons to fill an entire spectrum. For this reason, many surveys make the choice to use photometric redshifts instead.

Photometric redshifts fundamentally measure the brightness of objects through various filters, and fits those to template objects with known redshifts. Often, this involves building a catalogue of spectroscopic objects from which a model can be constructed. Each new set of photometry is then applied to this model to determine the redshift of each object. This is a much quicker process, because it doesn't require long integration times in order to get a measure of an object's redshift. Because it requires fewer photons to make a direct detection, it also allows for the calculation of redshifts for much dimmer objects than for spectroscopic redshifts.

Because the algorithm assumes that the redshifts are accurate, this introduces some measure of spread in the bounds on the line of sight separations in the galaxy pairs. Ultimately, this doesn't have too large an effect on the measurement, because if a galaxy pair doesn't contain a filament, it will simply not contribute to the underlying signal, but will contribute to the halo signals.

It also seems that we haven't managed to successfully construct galaxy halos that take the form of gaussians from our list of galaxy pairs. This may be due to the uneven distribution of galaxy pair separations on the sky, so we could possibly correct for this by selecting an even distribution of transverse separations. The other possible correction that could be made here is to sample more pairs, over a larger area. This would drive the distribution of scales closer to sampling from a gaussian by the central limit theorem, and so hopefully drive the shape of the rescaled halos to the same distribution.

One possible method of mitigating this effect is to make better use of the null tests when analysing our physical pairs. This could be done by constructing sets of unphysical pairs which contain the same number as the physical set. We could then subtract the unphysical stack from the physical stack. Doing this for a number of different bundles of unphysical pairs would allow us to take an average, and hopefully model the overarching halo structure as subject to our analysis pipeline.

### 5.1.1 Conclusion

The significance of the signal detected is lower than that of previous work ( $\sim 2\sigma$  c.f.  $\sim 5\sigma$ ), due to the unreliability of our modelling, and the uncertainty introduced by uneven scaling factors. Despite this, the higher reading for our Compton  $y$  parameter is considerably higher than previous work, which we attribute to the higher precision of the  $y$  map we are using.

We performed our fit initially on the SPTpol footprint only, due to the expense in computation when constructing galaxy pairs. With less than a sixth of the original catalogue, over 700 000 pairs were produced. For the sake of edification, the pairing algorithm was run on the full SPT-SZ footprint, and it produced  $\sim 5.8$  million physical pairs, and over 15.5 million unphysical pairs. With this number of pairs, even the stacking algorithm becomes computationally expensive.

### Further Work

Further work can be done to strengthen the conclusions drawn by this work. Improvements in model construction, and refinement of the algorithm are possible, to eliminate possible systematics introduced by the mirroring process.

In order to do a direct computation of the overall baryon fraction from this  $y$ -parameter, we need to apply this algorithm to a set of simulations which occupy a similar comoving volume. Doing so could also give insight into sources of noise in the physical Compton- $y$  dataset, or those

introduced by the data analysis pipeline.

Further work can also apply this algorithm to other data sources. Any data source which acts as a tracer for large scale structure with known statistical information, can be used, since the algorithm doesn't discriminate based on input data type. This process was initially developed for  $\kappa$  convergence maps (Clampitt et al., 2016), but it can equally be applied to the other form of Sunyaev-Zel'dovich effect , Kinetic Sunyaev-Zel'dovich effect , which acts as a tracer for bulk velocities in the intervening structure.

All previous work has been done with single data sources, but performing it with multiple allows for calculations of cross correlations between independant data sources, which have some major advantages over single source detections. Many of the systematics between the data types will be uncorrelated, and so improve the quality of our signal, whilst driving the noise assosciated with the detection down.

As it stands, there is strong evidence for the existence of galactic filaments from several independant teams. This work is the first to perform it with a high resolution CMB dataset, and provides strong support for previous measurements.

# Bibliography

- Bristow, P. D. & Phillipps, S., 1994, On the Baryon Content of the Universe, *MNRAS* 267, 13
- Burles, S. & Tytler, D., 1998, Measurements of the deuterium abundance in quasar absorption systems, in Mezzacappa, A., editor, *Stellar Evolution, Stellar Explosions and Galactic Chemical Evolution*, p. 113
- Cen, R. & Fang, T., 2006, Where Are the Baryons? III. Nonequilibrium Effects and Observables, *ApJ* 650, 2, 573
- Cen, R. & Ostriker, J. P., 1999, Where Are the Baryons?, *ApJ* 514, 1, 1
- Cen, R. & Ostriker, J. P., 2006, Where Are the Baryons? II. Feedback Effects, *ApJ* 650, 2, 560
- Clampitt, J.; Miyatake, H.; Jain, B. & Takada, M., 2016, Detection of stacked filament lensing between SDSS luminous red galaxies, *MNRAS* 457, 3, 2391
- Colberg, J. M.; Krughoff, K. S. & Connolly, A. J., 2005, Intercluster filaments in a  $\Lambda$ CDM Universe, *MNRAS* 359, 1, 272
- Compton, A. H., 1923, A Quantum Theory of the Scattering of X-rays by Light Elements, *Physical Review* 21, 5, 483
- Crawford, T. M.; Chown, R.; Holder, G. P.; Aird, K. A.; Benson, B. A.; Bleem, L. E.; Carlstrom, J. E.; Chang, C. L.; Cho, H. M.; Crites, A. T. et al., 2016, Maps of the Magellanic Clouds from Combined South Pole Telescope and PLANCK Data, *ApJS* 227, 2, 23
- Danforth, C. W. & Shull, J. M., 2005, The Low-z Intergalactic Medium. I. O VI Baryon Census, *ApJ* 624, 2, 555
- Danforth, C. W. & Shull, J. M., 2008, The Low-z Intergalactic Medium. III. H I and Metal Absorbers at  $z < 0.4$ , *ApJ* 679, 1, 194
- Davé, R.; Cen, R.; Ostriker, J. P.; Bryan, G. L.; Hernquist, L.; Katz, N.; Weinberg, D. H.; Norman, M. L. & O'Shea, B., 2001, Baryons in the Warm-Hot Intergalactic Medium, *ApJ* 552, 2, 473
- de Graaff, A.; Cai, Y.-C.; Heymans, C. & Peacock, J. A., 2019, Probing the missing baryons with the Sunyaev-Zel'dovich effect from filaments, *A&A* 624, A48
- Epps, S. D. & Hudson, M. J., 2017, The Weak Lensing Masses of Filaments between Luminous Red Galaxies, *MNRAS* 468, 3, 2605
- Fukugita, M.; Hogan, C. J. & Peebles, P. J. E., 1998, The Cosmic Baryon Budget, *ApJ* 503, 2, 518
- Fukugita, M. & Peebles, P. J. E., 2004, The Cosmic Energy Inventory, *ApJ* 616, 2, 643
- Hill, J. C. & Spergel, D. N., 2014, Detection of thermal SZ-CMB lensing cross-correlation in Planck nominal mission data, *J. Cosmology Astropart. Phys.* 2014, 2, 030
- Hogg, D. W., 1999, Distance measures in cosmology, *arXiv e-prints* astro-ph/9905116
- Hojjati, A.; McCarthy, I. G.; Harnois-Deraps, J.; Ma, Y.-Z.; Van Waerbeke, L.; Hinshaw, G. & Le Brun, A. M. C., 2015, Dissecting the thermal Sunyaev-Zeldovich-gravitational lensing cross-correlation with hydrodynamical simulations, *J. Cosmology Astropart. Phys.* 2015, 10, 047
- Hubble, E., 1929, A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae,

*Proceedings of the National Academy of Science* 15, 3, 168

- Hurier, G.; Macías-Pérez, J. F. & Hildebrandt, S., 2013, MILCA, a modified internal linear combination algorithm to extract astrophysical emissions from multifrequency sky maps, *A&A* 558, A118
- Kompaneets, A. S., 1957, The Establishment of Thermal Equilibrium between Quanta and Electrons, *Soviet Journal of Experimental and Theoretical Physics* 4, 5, 730
- Levi, B. G., 1992, COBE measures anisotropy in cosmic microwave background radiation., *Physics Today* 45, 6, 17
- Ma, Y.-Z.; Van Waerbeke, L.; Hinshaw, G.; Hojjati, A.; Scott, D. & Zuntz, J., 2015, Probing the diffuse baryon distribution with the lensing-tSZ cross-correlation, *J. Cosmology Astropart. Phys.* 2015, 9, 046
- Peacock & Peacock, 1998, p. 66, Cambridge University Press
- Penton, S. V.; Shull, J. M. & Stocke, J. T., 2000, The Local Ly $\alpha$  Forest. II. Distribution of H I Absorbers, Doppler Widths, and Baryon Content, *ApJ* 544, 1, 150
- Penton, S. V.; Stocke, J. T. & Shull, J. M., 2004, The Local Ly $\alpha$  Forest. IV. Space Telescope Imaging Spectrograph G140M Spectra and Results on the Distribution and Baryon Content of H I Absorbers, *ApJS* 152, 1, 29
- Penzias, A. A. & Wilson, R. W., 1965, A Measurement of Excess Antenna Temperature at 4080 Mc/s., *ApJ* 142, 419
- Persic, M. & Salucci, P., 1992, The baryon content of the universe, *MNRAS* 258, 1, 14P
- Planck Collaboration; Aghanim, N.; Akrami, Y.; Ashdown, M.; Aumont, J.; Baccigalupi, C.; Ballardini, M.; Banday, A. J.; Barreiro, R. B. & Bartolo, N., 2018a, Planck 2018 results. VI. Cosmological parameters, *arXiv e-prints* arXiv:1807.06209
- Planck Collaboration; Akrami, Y.; Arroja, F.; Ashdown, M.; Aumont, J.; Baccigalupi, C.; Ballardini, M.; Banday, A. J.; Barreiro, R. B.; Bartolo, N. et al., 2018b, Planck 2018 results. I. Overview and the cosmological legacy of Planck, *arXiv e-prints* arXiv:1807.06205
- Rao, S. & Briggs, F., 1993, Neutral Hydrogen in Galaxies at the Present Epoch, *ApJ* 419, 515
- Remazeilles, M.; Delabrouille, J. & Cardoso, J.-F., 2011, CMB and SZ effect separation with constrained Internal Linear Combinations, *MNRAS* 410, 4, 2481
- Rephaeli, Y., 1995, Comptonization Of The Cosmic Microwave Background: The Sunyaev-Zeldovich Effect, *ARA&A* 33, 541
- Richter, P.; Savage, B. D.; Sembach, K. R. & Tripp, T. M., 2006, Tracing baryons in the warm-hot intergalactic medium with broad Ly  $\alpha$  absorption, *A&A* 445, 3, 827
- Rozo, E.; Rykoff, E. S.; Abate, A.; Bennett, C.; Crocce, M.; Davis, C.; Hoyle, B.; Leistedt, B.; Peiris, H. V.; Wechsler, R. H. et al., 2016, redMaGiC: selecting luminous red galaxies from the DES Science Verification data, *MNRAS* 461, 2, 1431
- Schaye, J., 2001, Model-independent Insights into the Nature of the Ly $\alpha$  Forest and the Distribution of Matter in the Universe, *ApJ* 559, 2, 507
- Shull, J. M.; Smith, B. D. & Danforth, C. W., 2012, The Baryon Census in a Multiphase Intergalactic Medium: 30% of the Baryons May Still be Missing, *ApJ* 759, 1, 23
- Smith, B. D.; Hallman, E. J.; Shull, J. M. & O'Shea, B. W., 2011, The Nature of the Warm/Hot Intergalactic Medium. I. Numerical Methods, Convergence, and O VI Absorption, *ApJ* 731, 1, 6
- Spergel, D. N.; Bean, R.; Doré, O.; Nolta, M. R.; Bennett, C. L.; Dunkley, J.; Hinshaw, G.; Jarosik, N.; Komatsu, E.; Page, L. et al., 2007, Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology, *ApJS* 170, 2, 377
- Steigman, G., 2007, Primordial Nucleosynthesis in the Precision Cosmology Era, *Annual Review of Nuclear and Particle Science* 57, 1, 463

- Tanimura, H.; Aghanim, N.; Douspis, M.; Beelen, A. & Bonjean, V., 2019, Detection of intercluster gas in superclusters using the thermal Sunyaev-Zel'dovich effect, *A&A* 625, A67
- Van Waerbeke, L.; Hinshaw, G. & Murray, N., 2014, Detection of warm and diffuse baryons in large scale structure from the cross correlation of gravitational lensing and the thermal Sunyaev-Zeldovich effect, *Phys. Rev. D* 89, 2, 023508
- Weinberg, D. H.; Miralda-Escudé, J.; Hernquist, L. & Katz, N., 1997, A Lower Bound on the Cosmic Baryon Density, *ApJ* 490, 2, 564
- Wright, E. L., 2004, Theoretical Overview of Cosmic Microwave Background Anisotropy, in Freedman, W. L., editor, *Measuring and Modeling the Universe*, p. 291
- Zwaan, M. A.; Staveley-Smith, L.; Koribalski, B. S.; Henning, P. A.; Kilborn, V. A.; Ryder, S. D.; Barnes, D. G.; Bhathal, R.; Boyce, P. J.; de Blok, W. J. G. et al., 2003, The 1000 Brightest HIPASS Galaxies: The H I Mass Function and  $\Omega_{HI}$ , *AJ* 125, 6, 2842