COSMOLOGY FROM THE BISPECTRUM

WILLIAM RICHARD COULTON

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
PHYSICS
ADVISER: PROFESSOR DAVID SPERGEL

SEPTEMBER 2018
Abstract

With a suite of current and up-coming experiments set to provide us the exquisite measurements of the sky in a wide range of frequencies, we are posed with the problem of how to maximally utilise these data sets. Historically two-point function measurements have been the tool of choice, and they have been highly effective. However, two-point functions are unable to extract all the available information, particularly on the small, non-linear scales. With the goal of providing complimentary measurements to the power spectrum, we use the bispectrum to explore a range of physical processes.

In the first half of this thesis, we discuss two applications of the bispectrum to probe large scale structure. Firstly, we measure the bispectrum from secondary anisotropies of the cosmic microwave background using data from the Atacama Cosmology telescope and the Planck Satellite. Secondly, we use a suite of simulations to explore the bispectrum of weak lensing shear maps. Using these simulations we explore how bispectrum measurements can be used to enhance our ability to measure cosmological parameters, including the sum of the masses of the neutrinos. We find that, for an experiment like the Large Synoptic Sky Survey, including information from the bispectrum will enhance parameter constraints by 30%.

In the second half of this thesis, we discuss two aspects regarding primordial gravitational waves. First, we discuss how measurements of the bispectrum can be used to probe primordial non-Gaussianity between tensor and scalar fields and present our preliminary constraints on interactions between these fields. Secondly, we discuss how bispectrum measurements of galactic foregrounds, including polarised dust and synchrotron emission, have the potential to help discriminate between the desired signal from primordial gravity waves and galactic contaminants.
Acknowledgements

Firstly and foremost I would like to thank my advisor, David Spergel, for his patience, support and guidance. I am continually inspired by his insight and vast knowledge. Often it seems as though David has a thousand different responsibilities and yet he always managed to make time to resolve my issues, for which I am very grateful. Secondly I would like thank Robert Lupton, who co-supervised my experimental project. Robert was incredibly patient with my questions and everything I do correctly in image analysis can be attributed to him. Beyond my supervisors, I have worked with, and learnt from, many incredible people. I would like to thank all the members of Gravity Group, Jadwin Hall and Peyton, whom have made my time so special to me. I would especially like to thank my many collaborators, in particular Robert Armstrong, Kendrick Smith, Jia Liu, Mathew Madhavacheril, Alexander van Engelen, and all of the ACTPol collaboration. I would like Suzanne Staggs, Lyman Page and Mariangela Lisanti for being on my pre-thesis and thesis committees and Jo Dunkley for reading all of my thesis. Without the efforts of these people this thesis would not have been possible. Thank you!

I am incredibly grateful to have had the opportunity to study at Princeton and I would like to thank everyone who has helped me on my path in Physics. This includes my family and the long list of inspiring and creative teachers, who have fostered my interest in Physics and taught me everything I know. I would also like to thank every one who had the patience to supervise me during my undergraduate research projects, particularly Stephen Wilkins, who supervised my first research experience in Astrophysics. This experience cultivated my interest in a research career and set me on my course towards a PhD in cosmology. Thank you!

My time at Princeton would not have been the same without the many friends I have made. It is a pleasure to thank Joaquin Turiaci, Debayan Mitra, Aaron Levy, Zachary Sethna, Shai Chester, Vladimir Kirilin and Aitor Lewkowycz for their friend-
ship; our conversations not only brightened my days but also enlightened my views on everything from physics to philosophy. I would like to thank the many, many members of Jadwin and Peyton hall whose company I have enjoyed throughout my years and who made Princeton feel like a home. Finally I would like thank Elena for her support, care and love during all these years. Thank you!
In memory of my mum
Contents

Abstract ................................................................. iii
Acknowledgements ....................................................... iv
List of Tables ............................................................ xi
List of Figures ............................................................ xiii

1 Introduction 1

2 Non-Gaussianity of Secondary Anisotropies from ACTPol and Planck 9

2.1 Introduction ......................................................... 9

2.2 Bispectrum estimators in the flat-sky approximation .................. 14

2.2.1 KSW estimator .................................................. 19

2.2.2 Estimator Normalization and Linear Term .................... 20

2.3 Bispectrum Templates ............................................. 22

2.3.1 Poisson point source templates ................................ 22

2.3.2 Thermal Sunyaev Zel’dovich, dusty star forming galaxy and radio galaxy templates .................................................. 24

2.3.3 Lensing cross secondary sources templates .................... 26

2.4 Data Sets ............................................................. 27

2.4.1 ACTPol data sets ................................................. 28

2.4.2 Planck Data sets ................................................... 28
3.2.4 Binned Bispectrum Estimator ........................................... 79
3.2.5 Interpolation ................................................................. 81
3.2.6 Likelihoods ................................................................. 83
3.3 The effect of neutrino mass on the power-spectrum .................... 85
3.4 The effect of neutrino mass on the bispectrum ......................... 89
3.5 Constraints ................................................................. 91
3.6 Conclusions ............................................................... 93

4 Constraining Primordial Tensor Non-Gaussianity with ACTPol and Planck 95
4.1 Introduction ............................................................... 95
4.2 Bispectrum estimator and Analysis Pipeline ............................ 98
  4.2.1 Flat-sky bispectrum estimator ....................................... 98
4.3 Bispectrum Templates .................................................... 100
  4.3.1 Primordial Templates in the flat sky regime ....................... 100
  4.3.2 Primordial Templates Scalar-Scalar-Scalar ....................... 103
  4.3.3 Primordial tensor-scalar-scalar non-Gaussianity .................. 106
  4.3.4 Primordial tensor-tensor-scalar non-Gaussianity .................. 108
  4.3.5 Primordial tensor-tensor-tensor non-Gaussianity ................. 110
  4.3.6 Foreground bispectra ................................................ 111
4.4 Data Sets and Pipeline .................................................. 113
  4.4.1 ACTPol Maps .......................................................... 113
  4.4.2 Planck Maps .......................................................... 113
  4.4.3 Analysis Pipeline .................................................... 114
4.5 Results ................................................................. 115
4.6 Conclusion ............................................................. 116
4.7 Appendix: Simulating non-Gaussianity in the Flat-Sky ............... 116
  4.7.1 Flat-sky simulation methods ........................................ 117
4.7.2 Odd Parity Estimator Validation  

5 Exploring Galactic foregrounds with the bispectrum  

5.1 Introduction  

5.2 Bispectrum Estimator  
5.2.1 Parity Even Estimator  
5.2.2 Parity Odd Estimator  
5.2.3 Linear Term  

5.3 Data Sets and Pipeline  
5.3.1 Data sets  
5.3.2 Pipeline  
5.3.3 Signal Inhomogeneity  

5.4 Foreground Bispectra  
5.4.1 353 GHz Bispectra  
5.4.2 Synchrotron Bispectra  
5.4.3 Dust - Synchrotron Cross Bispectra  
5.4.4 Cleaned Map Bispectra  

5.5 Discussion and Conclusions  

6 Conclusion  

Bibliography
## List of Tables

1.1 The values of the six $\Lambda$CDM parameters as reported in Planck Collaboration XIII (2016) .... 7

2.1 Template amplitudes obtained from a joint fit using 100 GHz and 217 GHz Planck maps with the ACTPol PA1 and PA2 148 GHz maps and $\ell_{\text{max}} = 7500$. Point sources below 269, 152, 30 mJy for frequencies 100, 217, 148 GHz respectively were masked and no masking of clusters was performed. The amplitudes, $A_i$, are overall scalings of our templates and are dimensionless. The Gaussian errors only indicate the significance of the non-Gaussianity in these data sets. The non-Gaussian error results include the one-halo contributions to the cosmic variance and these errors reflect the statistical uncertainty in the template amplitudes. ............... 33
2.2 A comparison between various different bispectrum measurements with all results given in $\mu K^3 sr^2$. The Komatsu et al. column contains the results from Komatsu et al. (2009), Calabrese et al. column contains the results from Calabrese et al. (2010) and Crawford et al. column is from Crawford et al. (2014). The final column shows what our model, with amplitudes fitted from the ACTPol and Planck data as presented here, would predict for these measurements. To compare our DSFG result to the Crawford et al. result we evaluate our model at $\ell = 3000$, 220 GHz and at their flux cut of 6.4 mJy. The factor of two difference between our results and those from Calabrese et al. (2010) is discussed in the text.

2.3 The expected signal-to-noise the kinetic Sunyaev Zel’dovich templates. We see that whilst the sensitivities for the individual templates is close to having constraining power, when the templates are jointly fit the similarity of the templates to our data sets leaves no constraining power.

2.4 A comparison between the expected and measured standard deviations of our estimator when applied to simulated Gaussian maps. The consistency of our estimator is a check of the optimality of our implementation, especially our approximate inverse covariance filtering.

3.1 The projected source counts per arcmin$^2$ used in this work.

4.1 Preliminary limits obtained from combining ACTPol 148GHz T+E with Planck SMICA B mode with $\ell_{\text{max}} = 5000$ and $\ell_{\text{min}} = 350$ and $\ell_{\text{min}} = 150$ for ACTPol and Planck respectively.
List of Figures

1.1 Figure from Perlmutter et al. (1999) showing the result of their supernovae measurements. ........................................ 3

1.2 Figures from the Planck experiment (Planck Collaboration XIII, 2016).

In the left figure are results from measurements of the temperature anisotropies and in right figure are results from measurements of the E polarization. The red curve in the right-hand plot is the predictions from the best fit parameters of the temperature results. .................... 5

1.3 Figure from Anderson et al. (2014) show the results from the BOSS experiment. The baryon acoustics oscillations seen here are signatures in the galaxy distribution of the same oscillations that are seen in CMB. 6
2.1 The one-halo bispectrum templates with $l_1 = l_2 = l_3$ for the tSZ, DSFGs and radio galaxies for three frequencies: 100, 148 and 217 GHz. The DSFG term contains the two and three galaxy terms but not the one galaxy term; the one galaxy term is a Poisson term with trivial scale dependence and so is excluded. Similarly, we exclude the radio point source term. The radio-tSZ is composed of two physical terms: the radio-tSZ-tSZ and the radio-radio-tSZ terms. The feature in the radio-tSZ template arises as the two contributions have different signs below 220 GHz and are dominant on different scales; the radio-tSZ-tSZ term is dominant on the largest scales and is positive at all frequencies and the radio-radio-tSZ term is dominant on the smallest scales and is negative at frequencies below 220 GHz. Similarly, the radio-DSFG term is composed of two physical terms: the radio-radio-DSFG and the radio-DSFG-DSFG terms. The scale is linear between $10^{-6}$ and $10^6$ and logarithmic outside this range. This leads to kinks in the curves that cross the log-linear transition.

2.2 The lensing cross secondary source bispectrum templates with $l_1 = l_2 = l_3$ at 148 GHz. We exclude the ISW-lensing template here as it has been well studied before. The scale is linear between $10^{-4}$ and $10^4$ and logarithmic otherwise.

2.3 The noise power spectra divided by the beam function squared for the four maps used in this analysis. The beams for the Planck 100 GHz and 217 GHz maps are treated as zero for $\ell > 2000$ and $\ell > 3000$ respectively.
2.4 The covariance matrix of template amplitudes only accounting for the Gaussian component to the six point function. The strong covariances between several templates highlights the difficult of disentangling all of the sources of non-Gaussianity. 34

2.5 The full one-halo covariance matrix for our measured estimators. The strong covariances present in this figure, but absent in Fig 2.4 arise due to the physical covariances of these templates. 35

2.6 A graphical representation of our joint fit results from table 2.1 excluding the lensing-ISW constraint in order to restrict the scales. The solid green line is the values predicted by our model and the dashed green line is the null value. 36

2.7 A set of graphs demonstrating how our pipelines can recover an input level of non-Gaussianity, input $A_i$, for a set of different templates with different max $\ell$. The amplitudes, $A_i$, are overall scalings of our templates and so are dimensionless. We compare our two methods of simulating non-Gaussian templates, the “direct” method and the “indirect” method, as described in section 2.10.2. Having verified that our estimators are unbiased on this set of templates we then use the “indirect” method to validate our full range of templates, which cannot easily be verified by direct simulation methods. 68

2.8 A map-by-map comparison of our estimators applied to maps with local type and Poisson noise non-Gaussianity simulated with the direct method. The input level of non-Gaussianity for local type this is $A_{\text{local}} = 2000$ and for Poisson noise this is $A_{\text{Poisson}} = 10.6$. It can be seen the two pipelines agree well on a map by map basis, increasing the confidence in the accuracy of our pipelines. 69
3.1 The difference in power of the $z = 1.0$ shear power spectrum between two cosmologies and the fiducial cosmology ($\sum m_\nu = 0.1, \Omega_m = 0.3, A_s = 2.1 \times 10^{-9}$) for 10 splits of the 10,000 simulations (dotted lines). The solid line is the difference in power from the RBF interpolation that uses the subset of cosmologies described in section 3.2.5.

3.2 A comparison of two methods of calculating the likelihood and how they behave as the number of simulations used to compute the likelihood is changed.

3.3 The shear power spectrum for three redshift configurations is plotted for the case of massless neutrinos (solid lines) and the case with $\sum m_\nu = 0.1$ eV (dotted lines). We compare the power-spectrum from simulations to power-spectra calculated using CAMB. The bottom plots show the fractional difference between the massive and massless neutrino power-spectra. The error bars show how well these configurations can be measured with an LSST-like experiment that observes half of the sky.

3.4 A comparison between constraints from a Fisher forecast with CAMB and constraints from the interpolated-likelihood analysis that is built using CAMB power-spectra. The contours are the 68% and 95% confidence levels.

3.5 A comparison between constraints from CAMB and our simulations. The contours are the 68% and 95% confidence levels.
3.6 In figures 3.1b and 3.6a the squeezed and equilateral slices for two cosmologies, \((\sum m = 0.1, \Omega = 0.3, A_s = 2.1 \times 10^{-9})\) and \((\sum m = 0.0, \Omega = 0.3, A_s = 2.1 \times 10^{-9})\) is plotted. We consider the bispectrum at four different redshift configurations, the \(z = 0.5\) auto-correlation bispectrum, \(z = 1.5\) auto-correlation bispectrum, the \(z = 2.5\) auto-correlation bispectrum and bispectrum from the cross correlation between \(z = 0.5, z = 1.5\) and \(z = 2.5\). The bottom panels show the fractional difference between the two cases. The error bars show how well these configurations can be measured with an LSST-like experiment that observes half of the sky.

3.7 68% parameter constraints from the power-spectrum, bispectrum and joint analysis.

3.8 A comparison between power-spectrum and bispectrum constraints on \(\Omega_M\) and \(\sum m\) from a single redshift measurement at \(z = 1.0\) against tomographic measurements from five redshift bins, \(z \in \{0, 0.5, 1, 1.5, 2, 2.5\}\). The total number of galaxies per arcmin\(^2\) is the same for the two cases.

4.1 A preliminary test of the new non-Gaussianity simulation method, as described in 4.7.1, to simulated tensor-scalar-scalar BTT non-Gaussianity. We see that there is an unknown issue that causes a discrepancy between the simulated and recovered value. We also simulate non-Gaussianity via the method described in Smith & Zaldarriaga (2011).

5.1 The average 1-D bispectrum of 100 realizations of a Poisson random field compared with the theoretical value.
5.2 The average 1-D chi-squared of 100 realizations of a Poisson noise and
100 realizations of a Gaussian CMB. ........................................ 134

5.3 The 353 GHz TT, EE and BB pseudo-Cl power-spectrum obtained
from 70% of the sky. ............................................................. 135

5.4 The parity even 1-D bispectrum and 1-D chi-squared for the 353 GHz
map. We used 70% of the sky for these measurements. The dotted
lines represent negative values of the bispectrum. ...................... 135

5.5 1-D chi-squared from half-ring difference maps for 70% of the sky. . 136

5.6 EET (solid lines) and BBT (dotted lines) 1-D 353 GHz bispectra for
different fractions of the sky. ................................................. 136

5.7 Equilateral slice of 353 GHz bispectrum 70% of the sky. .......... 136

5.8 The parity odd 1-D chi-squared for the 353 GHz maps. 70% of the sky
was used for these measurements. .......................................... 136

5.9 The parity even 1-D bispectrum and 1-D chi-squared for the command-
der synchrotron map. We used 70% of the sky for these measurements.
The dotted lines represent negative values of the bispectrum. ...... 140

5.10 The 1-D chi-squared for 353 GHz -synchrotron B mode cross correla-
tions. The solid lines are for the parity even estimator and the dotted
lines are for the parity odd case. ............................................. 140

5.11 The 1-D chi-squared for the cross bispectra between the 353 GHz maps
and the Planck 100,143 and 217 GHz B mode map for 70% of the sky.
The solid lines are the signal and the dashed lines are noise maps
constructed from half-ring maps. .......................................... 141

5.12 The 1-D chi-squared for the cross bispectra between the 353 GHz T,
E and B map and the SMICA B mode map for 70% of the sky ....... 141
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.13</td>
<td>The 1-D chi-squared for the cross bispectra between the 353 GHz T, E and B map and the B mode map from the different cleaning maps for 70% of the sky</td>
<td>142</td>
</tr>
<tr>
<td>5.14</td>
<td>The 1-D chi-squared for parity even cross bispectra between 353 GHz T, E and B maps and SMICA T and E maps on 70% of the sky</td>
<td>142</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Over the last 100 years our understanding of the universe around us has been transformed. Astoundingly we now have answers to many of the big questions which occupied the scientists’ minds during the 20th century such as what is the history of our universe? Has it always existed in this state? If not what is its age and history? Where did galaxies come from? Where did elements come from? What is the fate of our universe? This transformation has been driven jointly by theoretical advancements and ground breaking experiments.

On the theoretical side, this revolution began with Einstein’s theory of general relativity in 1915 [Einstein 1915]. Shortly after Einstein published his equations Friedmann, building on the assumptions of isotropy and homogeneity, derived a solution that described how the coarse scale properties of the universe evolved [Friedman 1922, Friedmann 1924]. His work was independently rediscovered by Lemaître and further developed by Roberston and Walker [Lemaître 1931, Walker 1935, Robertson 1935]. The Friedmann-Lemaître-Roberston-Walker (FLRW) metric is now the backbone of theoretical cosmology and is used to describe the evolution of the universe.
On the experimental side, in 1927 Edwin Hubble found that distant objects were receding faster than nearby objects \cite{Hubble1929}. Before Hubble’s discovery the majority of the community believed in the steady state universe. Einstein modified his field equations with the addition of a cosmological constant (known as \( \Lambda \)) in an attempt to build a consistent theoretical model of the steady state universe \cite{Einstein1917}. Hubble’s work, paired with the theoretical advancements, began a decades long discussion as to whether the universe existed in a steady state or evolved from a big bang. In the 1950s the steady state universe was further developed by Bondi, Hoyle and Gold \cite{Hoyle1948, BondiGold1948}. Their work explained Hubble’s results by imposing a specific rate of matter creation. Meanwhile, work on the big bang model lead to the realization that, unlike the steady universe, if the universe began in a hot dense phase then it should be filled with relic radiation, with a temperature of \( \sim 5 \text{ K} \) \cite{AlpherHerman1948}.

Whilst radio galaxy number counts provided evidence in favour of the big bang model \cite{RyleClarke1961}, it was not until the cosmic microwave background (hereafter CMB) was discovered by \cite{PenziasWilson1965}, and identified by \cite{Dickeetal1965} as the relic radiation, that the big bang model was accepted. The big bang model, combined with models of nucleosynthesis first proposed by \cite{Alpheretal1948}, provides the basis of our current understanding of the history of the universe. During the subsequent years, many of the details of the history of our universe have been fleshed out. In 1967 Sachs and Wolfe predicted that the seeds of galaxies and clusters should be imprinted on the CMB \cite{SachsWolfe1967}. The details of the evolution of these seeds were calculated and the spectrum of these initial seeds was hypothesised to be scale invariant \cite{Peebles1965, Silk1968, PeeblesYu1970, Harrison1970, Zeldovich1972}.

Concurrently to the work on the CMB, the history and formation of galaxies and clusters was developed. In 1962 \cite{Eggetal1962} proposed that galaxies form from
the collapse of monolithic gas clouds. In 1970 Zel’dovich (1970) developed the model of pancake collapse and in 1974 Press and Schechter developed their model which predicted the number density of collapsed objects (Press & Schechter, 1974). Then in 1977-78 White & Rees (1978) proposed the bottom-up scheme for galaxy formation, where large galaxies and clusters are formed by the merger of smaller objects.

At this point in my highly selective summary of the history of cosmology, it almost seems as though the field of cosmology is approaching the end. However, the reality of the case is very different. As measurements of the CMB accrued, the CMB was found to be very homogeneous and this posed a problem. In the simple big bang model there is no reason for opposite parts of the sky, which are not casually connected, to have the same temperature. This problem is known as the horizon problem and, after being theorised by Rindler (1956), was discovered in our universe. Further, studies of galaxies by Rubin & Ford (1970) rediscovered the conclusions of Zwicky (1937) that visible matter alone was not sufficient to explain the rotation curves of galaxies. Thus, either Einstein’s equations require modification or there is an additional component of matter that has been previously unaccounted for. A final
paradigm shift occurred in 1998 when two teams (the High-z Supernova Search Team and the Supernova Cosmology Project) independently discovered that the rate of expansion of the universe was increasing rather the decreasing (as is expected for a universe filled only with radiation, matter and dark matter) (Perlmutter et al., 1999; Riess et al., 1998). This discovery requires the addition of dark energy (possibly in the form of a cosmological constant as proposed by Einstein) to the cosmological model as can be seen in figure 1.

Since the discovery of each of these issues there has been huge theoretical progress on each of them. In the 1970s Guth, Linde, Steinhard, Starobinsky and Albrecht proposed and developed the inflationary theory, which can explain the horizon problem (Guth, 1981; Linde, 1982; Albrecht & Steinhardt, 1982). In this theory the universe underwent a period of exponential expansion. In this theory, regions of space that are initially causally connected, become separated by the exponential expansion of space-time. This explains why the CMB can be so homogeneous, as disconnect parts of the universe today were in causal connection before inflation began. This theory has been hugely successful and is widely accepted in the field. However we are lacking direct evidence for inflation and, as experiments start to rule out some of the traditional inflationary models (Planck Collaboration XX, 2016), there is an increasingly heated discussion as to whether inflation universally has fatal theoretical issues (Ijjas et al., 2014; Ijjas & Steinhardt, 2016). In recent years an alternative family of solutions, the ekpyrotic models, have been put forward that can equally well describe the universe, though they also have their own theoretical issues (Khoury et al., 2002; Linde, 2014). In regards to dark matter, there is a wide range of possible theoretical candidates ranging from primordial black holes to massive weakly interacting particles to axions (see e.g. Garrett & Dīda, 2011; Lisanti, 2017, for a review). Similarly, there is now a plethora of theoretical models for dark energy ranging from Einstein’s cosmolog-
Figure 1.2: Figures from the Planck experiment [Planck Collaboration XIII, 2016]. In the left figure are results from measurements of the temperature anisotropies and in right figure are results from measurements of the E polarization. The red curve in the right-hand plot is the predictions from the best fit parameters of the temperature results.

physical constant, to modifications of gravity to scalar fields (see e.g. Tsujikawa, 2013, Mortonson et al., 2014 for a review).

These theoretical advancements have been driven and constrained by a suite of experiments. In 1994 COBE measured the spectrum of the CMB and found it to be a blackbody consistent with the big bang model (Mather et al., 1994). It also discovered small fluctuations in CMB temperature, of order 1 part in $10^4$, in CMB (Smoot et al., 1992). These fluctuations have now been studied by a wide array of different experiments (Bennett et al., 2013, Schaffer et al., 2011, Dunkley et al., 2011), and are the seeds of future structure. The imprint of dark matter can be seen in the CMB and is amongst the strongest evidence for dark matter, as it is generally difficult to explain this signature with dark matter. In addition, the polarized component of the CMB fluctuations has been discovered and these polarized fluctuations have been studied (Kovac et al., 2002, Naess et al., 2014, Austermann et al., 2012). The results from the most recent satellite experiment, the Planck satellite, are shown in figure 1.2a. It can been seen that the temperature fluctuations are measured to cosmic variance accuracy over scales ranging from 100 degrees down to arc-minute
Figure 1.3: Figure from Anderson et al. (2014) show the results from the BOSS experiment. The baryon acoustics oscillations seen here are signatures in the galaxy distribution of the same oscillations that are seen in CMB.

scales! Measurements of the polarised component, which are shown in figure 1.2b, are rapidly approaching the precision of the temperature measurements. The polarized component has allowed a precise test of our theoretical model as the line in figure 1.2b was obtained not by a fit to the polarised data, but instead as a prediction from the temperature data. The outstanding agreement is a testament to the accuracy of our theoretical model.

Concurrently to the microwave measurements of the CMB there has been impressive results from optical experiments. Most notably the Two-degree-Field Galaxy Redshift Survey (2dF Galaxy Redshift Survey), which performed a photometric survey of ~ 1500 deg\(^2\) of the sky (Peacock et al., 2001), and the Sloan Digital Sky Survey (SDSS), which imaged around 35% of the sky with a photometric and a spectroscopic survey (Gunn et al., 2006). The SDSS survey has provided maps of the structure out to redshift ~ 0.2. Additionally, through both of these surveys we have measured the baryon acoustic oscillation (BAO) feature from the galaxy distribution and the latest BAO measurements from Anderson et al. (2014) are shown in 1. The BAO feature is the signature of the same oscillations seen in the CMB. It is remarkable that we have evidence of the same physics, but separated by billions of years! It should be empha-
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Best Fit Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega_b h^2$</td>
<td>0.02230 ± 0.00014</td>
</tr>
<tr>
<td>$\Omega_c h^2$</td>
<td>0.1188 ± 0.0010</td>
</tr>
<tr>
<td>$t_0$</td>
<td>13.799 ± 0.021 $\times 10^9$ years</td>
</tr>
<tr>
<td>$n_s$</td>
<td>0.9667 ± 0.0040</td>
</tr>
<tr>
<td>$A_s \times 10^9$</td>
<td>2.441$^{+0.088}_{-0.092}$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.066 ± 0.012</td>
</tr>
</tbody>
</table>

Table 1.1: The values of the six $\Lambda$CDM parameters as reported in Planck Collaboration XIII (2016).

sized that there are numerous other experiments from radio waves, to X-rays to dark matter detectors, whose results have made major contributions to the development of cosmology that, for brevity, have not been mentioned here.

The current state of cosmology can be neatly summarised by considering the current theoretical model of our universe, $\Lambda$CDM. The $\Lambda$CDM model is characterized by six parameters: $\Omega_b h^2$, the energy density in baryons; $\Omega_{CDM} h^2$, the energy density in dark matter; $t_0$, the age of the universe; $n_s$ the tilt of the primordial power spectrum; $A_s$ the amplitude of primordial scalar fluctuations; and $\tau$, the reionization optical depth. By combining the optical measurements with CMB and supernova measurements we are able to accurately constrain these parameters and the current best constraints on these parameters are shown in table 1.1. It is remarkably that this simple, six parameter model can explain almost all of the diverse range of measurements performed to date, from the experiments described above to cluster counts to weak lensing correlation functions (e.g. DES Collaboration et al., 2017; Planck Collaboration XXIV, 2016).

Whilst we are now in the era of precision sciences, where we are making percentage accuracy measurements of the properties of the universe, answers to the fundamental questions from the second half of the 20th century have remained elusive. These open questions stand in stark contrast to the precision measurements of our era and are at the heart of cosmology. The major focus of many on-going and up-coming surveys
is the attempt to resolve of one or more of these questions (Abazajian et al., 2016a; LSST Science Collaboration et al., 2009a).

As the title suggests this thesis will be focused on the bispectrum, which is the harmonic equivalent of the three point function i.e. it measures how correlate three Fourier (or spherical harmonic) coefficients are. This immediately raises the questions of why the bispectrum? Why not the power spectrum or equivalently the two point function? Two point statistics are the work horse of modern cosmology and have been used extensively to study the CMB and large scale structure. The fluctuations in the CMB are described to a high accuracy by Gaussian statistics and as such can be characterised solely with two point statistics. However, if we wish to assess the assumption of Gaussianity, we need to use higher point statistics and the non-linear evolution of the CMB fluctuations into galaxies transfers information from the two point statistics to higher point correlation functions, and thus to extract all of the information available in surveys of late-time structure we need to use high point statistics.

In chapter 2 we develop models for, and constrain the bispectrum of the CMB’s secondary anisotropies. In chapter 3 we study the bispectrum of weak lensing shear maps and discuss how bispectrum measurements can be used to tighten constraints on the sum of the masses of the neutrinos. In chapter 4 we study the primordial bispectrum arising from tensor-scalar interactions and use data from ACTPol and Planck to constrain this type of non-Gaussianity. In chapter 5 we used bispectrum measurements to characterise polarized galactic dust and synchrotron and discuss how bispectrum measurements can be used in primordial gravitational wave searches. Finally in chapter 6 we present our conclusions.
Chapter 2

Non-Gaussianity of Secondary Anisotropies from ACTPol and Planck

This chapter is adapted from Coulton et al. (2017). The project was conducted under the supervision of Professor David Spergel and received valuable input from many members of the Atacama Cosmology Telescope collaboration. The results of this chapter have been presented at the following conferences and meetings: COSMO-21, Advances in Theoretical Cosmology in Light of Data, Rencontres du Vietnam, ACT collaboration meetings (2016 and 2017) and the SO collaboration meeting (2017). It has been submitted for publication with the Journal of Cosmology and Astroparticle Physics.

2.1 Introduction

Secondary sources of cosmic microwave background (CMB) anisotropy occur as photons propagate from the surface of last scattering to us. Lensing distortions arise as the paths of CMB photons are deflected by intervening matter.
der, 1987). The thermal Sunyaev Zel’dovich (tSZ) effect is induced by the inverse
Compton scattering of CMB photons as they travel through ionized gas, distorting
the frequency spectrum of the photons (Sunyaev & Zeldovich, 1970). The additional
kinematic Sunyaev-Zel’dovich (kSZ) effect produces a near-blackbody temperature
shift if the ionized gas possesses a bulk velocity along the line of sight (Sunyaev &
Zeldovich, 1972). Radio galaxies and dusty star-forming galaxies (DSFGs) are the
two dominant populations of extragalactic emissive sources at millimeter wavelengths.
Radio galaxies have large levels of synchrotron emission, primarily from active galac-
tic nuclei, and in dusty star-forming galaxies the radiation arises from thermal dust
emission.

Many of these secondary sources have been intensely studied. The power spectrum
of the tSZ effect was measured by Planck Collaboration XXII (2016) and its ampli-
tude on small angular scales has been measured (e.g. Dunkley et al., 2013; George
et al., 2015; Planck Collaboration XI, 2016). The thermal SZ effect in galaxy clus-
ters has been studied through cluster number counts (e.g. Staniszewski et al., 2009;
Hasselfield et al., 2013; de Haan et al., 2016; Planck Collaboration XXIV, 2016) and
through cross-correlations with numerous other tracers (e.g. Planck Collaboration
The gas properties in these clusters have been studied at low redshifts through X-
ray measurements (Arnaud et al., 2010). The dusty star-forming galaxies have been
measured through their number counts with Herschel and Spitzer (e.g. Glenn et al.,
2010; Berta et al., 2011b; Béthermin et al., 2010) and power spectrum (Planck Col-
laboration XXX, 2014) as well as through cross-correlations with lensing and other
tracers (e.g. Planck Collaboration XVIII, 2014; Planck Collaboration XXIII, 2016;
van Engelen et al., 2015). Similarly the radio galaxies’ number counts have been
measured (de Zotti et al., 2010) and hints of their correlation with the tSZ effect have
been seen (Gralla et al., 2014; Gupta et al., 2017). Finally the kSZ effect has been
studied through a wide variety of different methods and different data sets (e.g. Hand et al., 2012; De Bernardis et al., 2017; Planck Collaboration Int. XXXVII, 2016; Soergel et al., 2016; Planck Collaboration LIII, 2017; Schaan et al., 2016; Hill et al., 2016). All of these measurements have provided information about the astrophysics of these sources, but the non-Gaussian nature of these sources means that we can gain complementary information by analyzing the higher order correlation functions (e.g. Wilson et al., 2012; Hill & Sherwin, 2013; Crawford et al., 2014; Hill et al., 2014; Cooray et al., 2000; Timmons et al., 2017).

Here we focus on the bispectrum from secondary sources. The bispectrum is the harmonic-space transform of the spatial three-point correlation function, which is zero for a purely Gaussian random field. Numerous observed and hypothesized sources give a nonzero bispectrum contribution in the microwave sky. These fall into four categories: primordial sources, secondary sources, second-order gravitational terms and galactic foregrounds. Primordial non-Gaussianity offers another handle on constraining the physics of the early universe and could provide the ability to differentiate between many classes of inflationary theories, though so far it has been found to be consistent with zero (Komatsu et al., 2011; Planck Collaboration XXIV, 2014; Planck Collaboration XVII, 2016); see Liguori et al. (2010); Yadav & Wandelt (2010); Chen (2010) for reviews of primordial non-Gaussianity. Second-order gravitational terms arise from non-linearities in the transfer function from the initial perturbations to the late time temperature fluctuations in CMB (Pettinari et al., 2013; Bartolo et al., 2006, 2007; Pitrou et al., 2010). We do not discuss these here as they are below the sensitivity of our experiments but they will be important for future, more sensitive measurements (Pettinari et al., 2014). Galactic foreground sources include thermal dust emission and synchrotron emission; these are typically most important on the largest scales (degree scales). As we restrict our analysis to small angular scales
(ℓ > 200), these sources of non-Gaussianity are not considered in this work (Gold et al., 2011). We focus on investigating the bispectrum from CMB secondary sources.

The large scale DSFG bispectrum has been measured in Planck Collaboration XXX (2014) as has the large scale bispectrum from the tSZ effect (Planck Collaboration XXII, 2016) and the bispectrum from correlations between the Integrated Sachs-Wolfe and gravitational lensing (Planck Collaboration XVII, 2016; Planck Collaboration XXI, 2014). The skewness of the tSZ effect was examined in Wilson et al. (2012) and Hill & Sherwin (2013) and the one point probability density function (PDF) of the tSZ effect has been studied in Hill et al. (2014). The bispectrum from the kSZ effect and galaxies has been probed by pairwise estimators (Hand et al., 2012; De Bernardis et al., 2017; Soergel et al., 2016; Planck Collaboration Int. XXXVII, 2016), by combining stacked clusters with their reconstructed velocities (Schaan et al., 2016), and through projected field estimators (Hill et al., 2016). Crawford et al. (2014) explored the bispectrum from secondary sources constraining the small-scale tSZ effect, DSFG and Poisson contributions to the bispectrum. Our work builds on aspects of these previous works; in particular we extend these analyses to include the one-halo contributions to the cross-correlations among the tSZ effect, DSFGs and radio galaxies, such as tSZ-tSZ-DSFG and radio-DSFG-tSZ bispectra. Our sensitivity to these terms is enhanced compared to Crawford et al. (2014) due to our inclusion of cross frequency bispectra. By jointly constraining these sources we hope to clearly identify the contributions arising from the differing components.

Bispectrum measurements of secondary sources could provide a wealth of astrophysical information. One halo tSZ bispectrum measurements probe gas pressure profiles on small scales, which provide insight into galaxy cluster processes such as AGN feedback (Battaglia et al., 2012a). Hurier & Lacasa (2017) have used measurements of the tSZ bispectrum, power spectrum and cluster counts to constrain the hydrostatic mass bias, the matter density and the amplitude of fluctuations. Cross-
correlations among the tSZ effect, radio galaxies and DSFGs can disentangle their contributions to the temperature power spectrum measurements, which could lead to a better understanding of the observed deficit of small-scale tSZ power (Planck Collaboration XXII, 2016). The cross-bispectra from the tSZ and optical weak-lensing convergence maps can help constrain hydrostatic mass bias (Nelson et al., 2014; Rasia et al., 2006), a limiting systematic in tSZ cluster cosmology, as well as constraining cosmological parameters (Bhattacharya et al., 2012; Crawford et al., 2014; Wilson et al., 2012). Beyond the tSZ effect, the small-scale DSFG bispectrum can constrain the spatial distribution of DSFGs and the cross-bispectrum from the tSZ and dusty galaxies or radio galaxies probes the masses of haloes these galaxies occupy. The bispectrum provides a flux-weighted count of the total number of these galaxies which can be combined with other measures of galaxy counts to constrain their population properties. DSFGs are a tracer of star formation (Elbaz et al., 2007) and bispectrum measurements can elucidate the interplay between cluster physics and star formation. Bispectrum measurements can help constrain the power spectrum of the kSZ effect (Reichardt et al., 2012), and joint measurements of the kSZ bispectrum and power spectrum can probe cluster thermodynamics and the distribution of baryons (Battaglia et al., 2017; Schaan et al., 2016). Finally, bispectrum measurements can be used to characterize extragalactic biases to CMB lensing power spectra (van Engelen et al., 2014; Osborne et al., 2014).

In this work we analyze microwave intensity data from the Atacama Cosmology Telescope Polarimeter (ACTPol) at 148 GHz over a sky area of $\sim 550$ deg$^2$ (Thornton et al., 2016; Louis et al., 2017), combined with Planck 100 GHz and 217 GHz intensity data (Planck Collaboration VI, 2014) in the same region. The ACTPol data have an effective angular resolution of 1.4 arcmin, compared to Planck’s effective resolution of 9.7 arcmin (100 GHz) and 5.0 arcmin (217 GHz). We use the Planck 2015 cosmology (Planck Collaboration XIII, 2016).
In section 4.2 of this paper we review the Komatsu, Spergel and Wandelt (KSW) estimator (Komatsu et al., 2005) and its flat-sky limit and section 4.3 provides an overview of our secondary CMB anisotropy templates. Sections 4.4 and 2.5 briefly describe the data sets used here and the analysis pipeline applied to the data. We present our results and conclusions in sections 4.5 and 2.7. Appendix 2.8 derives the correspondence between the full-sky and flat-sky bispectrum estimators, while Appendix 2.9 contains a detailed derivation of the bispectrum templates for the major non-Gaussian microwave components. The details of our pipeline validation tests are described in Appendix 2.10. Appendix 2.11 contains the calculation of the non-Gaussian contribution to our estimator errors and Appendix 2.12 describes how to calculate the DSFG n point functions used in calculating the non-Gaussian errors.

2.2 Bispectrum estimators in the flat-sky approximation

The ACTPol and Planck experiments measure fluctuations in the specific intensity, \( \Delta I_\nu(n) \), of CMB across the sky at a frequency \( \nu \). These intensity fluctuations are then related to a temperature fluctuations \( \Delta T(n) \) via

\[
\Delta T(n) = k_B T^2 c^2 \frac{e^{hv/(k_BT)} - 1}{2h^2\nu^4} \Delta I_\nu(n),
\]

where \( k_B \) is the Boltzmann’s constant, \( T \) is the CMB temperature, \( h \) is the Planck’s constant and \( c \) is the speed of light. In this work we focus on small regions of sky, so a flat-sky approximation is valid. In this regime the temperature anisotropies are decomposed as Fourier modes:

\[
\frac{\Delta T(n)}{T} = \int \frac{d\ell^2}{4\pi^2} a_\ell e^{i n \cdot \ell}.
\]
This is analogous to full-sky analyses where the temperature fluctuations are expanded into spherical harmonics, $Y_{\ell,m}$:

$$\frac{\Delta T(n)}{T} = \sum_{\ell} \sum_{-\ell < m < \ell} Y_{\ell,m}(n)a_{\ell,m}.$$  \hfill (2.3)

The flat-sky approximation is accurate to better than 1% for $\ell > 200$ (Loverde & Afshordi, 2008). We will focus our analysis on the bispectrum, which is equal to the ensemble average of three $a^X_\ell$ (where the superscript $X$ denotes maps from different frequencies or telescope arrays):

$$B^{(X_1,X_2,X_3)}(\ell_1, \ell_2, \ell_3) = \left\langle a_{\ell_1}^{X_1} a_{\ell_2}^{X_2} a_{\ell_3}^{X_3} \right\rangle.$$  \hfill (2.4)

Under the assumption of rotational invariance, the flat-sky bispectrum can be expressed as (Spergel & Goldberg, 1999a; Hu, 2000a)

$$B^{X_1,X_2,X_3}(\ell_1, \ell_2, \ell_3) = 4\pi^2 \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) b^{X_1,X_2,X_3}_{\ell_1\ell_2\ell_3}.$$  \hfill (2.5)

where $b^{X_1,X_2,X_3}_{\ell_1\ell_2\ell_3}$ is the reduced bispectrum. The assumption of rotational invariance is justified as the bispectra considered in this work are from extragalactic sources. The full sky bispectrum has a corresponding form for rotational invariance; see Appendix 2.8 for a more detailed discussion of the correspondence.

In the case of weak non-Gaussianity, we can obtain an optimal (minimum-variance) estimator for the amplitude $A$ of a non-Gaussian component with a given reduced bispectrum, $b^{\text{th}}_{\ell_1\ell_2\ell_3}$. In full-sky analyses the optimal estimator can be written as
where $X$ parameterizes the different sky maps, $(C^{-1})^{X_i X_j}$ is the inverse of the covariance matrix $C^{X_i X_j} = \langle a_{\ell_1 m_1}^{X_i} a_{\ell_2 m_2}^{X_j} \rangle$, $N$ is the normalization and $G_{\ell_1 \ell_2 \ell_3}^{m_1 m_2 m_3}$ is the Gaunt integral over the product of three spherical harmonics (see equation 2.39 for the Gaunt integral definition). The normalization is chosen such that a unity amplitude is returned for an input map with a reduced bispectrum equal to the theoretical reduced bispectrum. The covariance matrix includes two components:

$$C_{\ell,m,\ell',m'} = w_{\ell,m} w_{\ell',m'} C_{\text{Signal}}^{\ell,m,\ell',m'} + N_{\ell,m,\ell',m'}.$$  \hspace{1cm} (2.7)

The $w_{\ell,m}$ carries the information of the beam profile and the pixelization function, $C_{\text{Signal}}^{\ell,m,\ell',m'}$ is the covariance of the sky signal and $N_{\ell,m,\ell',m'}$ is the noise covariance.

For the flat-sky limit, we start from the distribution for Gaussian fluctuations:

$$P^{(G)}(a^{\nu_1}, a^{\nu_2}, ...) \propto \exp \left[ -\frac{1}{2} \sum_{X_1, X_2} \int d^2 \ell_1 d^2 \ell_2 a^{X_1}_{\ell_1} (C^{-1}_{\ell_1,\ell_2})^{X_1 X_2} a^{X_2}_{\ell_2} \right],$$  \hspace{1cm} (2.8)

where the $X_i \in \{\nu_1, \nu_2, \ldots\}$, and analogously

$$C_{\ell,\ell'} = w_{\ell} w_{\ell'} C_{\text{Signal}}^{\ell,\ell'} + N_{\ell,\ell'},$$  \hspace{1cm} (2.9)

hereafter $C = C^{\text{Total}}$. Assuming that the amount of non-Gaussianity is small, we can then perform an Edgeworth expansion (Taylor & Watts, 2001; Juszkiewicz et al., 16).
notation to denote inverse covariance filtered maps: $\hat{C}^{-1}_X a$. The linear term (Yadav et al., 2008; Creminelli et al., 2006). We calculate the linear term

$$P(a_{\nu_1}, a_{\nu_2}, \ldots) \propto \left(1 - \frac{1}{6} \sum_{\ell} \int \prod_i d^2 \ell_i \left(a_{\ell_1} a_{\ell_2} a_{\ell_3}\right) \frac{\partial}{\partial a_{\ell_1}} \frac{\partial}{\partial a_{\ell_2}} \frac{\partial}{\partial a_{\ell_3}} + \ldots\right) \times P^{(G)}(a_{\nu_1}, a_{\nu_2}, \ldots) \propto \left(1 - \frac{1}{6} \sum_{\ell} \sum_{\ell'} \int \prod_i d^2 \ell_i \left(a_{\ell_1} a_{\ell'_1} a_{\ell'_3}\right) \left((C_{\nu_1, \ell_1}^{-1})^{X_1X'_1 a_{\ell'_1}} a_{\ell'_1}\right) \frac{\partial}{\partial a_{\ell'_1}} \frac{\partial}{\partial a_{\ell'_3}} \frac{\partial}{\partial a_{\ell'_3}} + \ldots \right) \times P^{(G)}(a_{\nu_1}, a_{\nu_2}, \ldots).$$

Then we impose rotational invariance on the bispectrum (Hu, 2000a)

$$\left(a_{\ell_1} a_{\ell_2} a_{\ell_3}\right) = A 4 \pi^2 b_{\ell_1\ell_2\ell_3}^{X_1X_2X_3, \text{th}} \delta^{(2)}(\ell_1 + \ell_2 + \ell_3),$$

where $\ell = |\ell|$ and obtain the optimal estimator as

$$\hat{A} = \frac{1}{N} \bar{A} = \frac{1}{N} \sum_{\ell} \sum_{\ell'} \int \prod_i d^2 \ell_i \left(4 \pi^2 b_{\ell_1\ell_2\ell_3}^{X_1X_2X_3, \text{th}} \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) \times \left((C_{\nu_1, \ell'_1}^{-1})^{X_1X'_1 a_{\ell'_1}} a_{\ell'_1}\right) \frac{\partial}{\partial a_{\ell'_1}} \frac{\partial}{\partial a_{\ell'_3}} \frac{\partial}{\partial a_{\ell'_3}} + \ldots \right),$$

where $N$ is the normalization term as before, $\bar{A}$ is the unnormalized template amplitude, and $\hat{A}$ is the estimated template amplitude. For simplicity, we use the following notation to denote inverse covariance filtered maps:

$$C_{\ell}^{-1} X a \equiv \sum_{X'} \int d^2 \ell' (C_{\ell, \ell'}^{-1})^{X' X a_{\ell'}}.$$

The last term in equation 4.7 which depends only on one $a_{\ell}^X$, is known as the linear term (Yadav et al., 2008; Creminelli et al., 2006). We calculate the linear term
and the estimator normalization by ensemble averages, which is described in section 2.2.2. In the optimal case the variance of this estimator is related to the normalization by:

\[ \text{Var}(\hat{A}) = \frac{1}{N}. \]  

(2.14)

In our analysis we apply a real space mask, \( M(n) \), and this alters this formula to

\[ \text{Var}(\hat{A}) = \frac{f^{(6)}_{\text{sky}}}{f^{(3)^2}_{\text{sky}} N_{\text{NoMask}}} \]  

(2.15)

where the proportionality constant depends on \( f^{(n)}_{\text{sky}} = 1/(N_{\text{pix}}) \int d^n M(n)^n \) and \( N_{\text{pix}} \) is the total number of pixels. We will fit multiple templates simultaneously so we will use an extended version of this estimator:

\[ \hat{A}^{\text{joint}}_i = \sum_j N_{i,j}^{-1} \hat{A}_j \]  

(2.16)

where \( N_{i,j} \) is a generalized normalization constant and the covariance of the estimators is given by:

\[ \text{Cov}(\hat{A}^i, \hat{A}^j) = \frac{f^{(6)}_{\text{sky}}}{f^{(3)^2}_{\text{sky}}} N_{i,j}^{-1}. \]  

(2.17)

Whilst this estimator is derived in the limit of weak non-Gaussianity, it is unbiased for large non-Gaussianity but not optimal. For large non-Gaussianity there could be additional contributions to the template covariances and these are discussed in Appendix 2.11. Hereafter the covariance matrix in 2.17 is called the ‘Gaussian’ covariance and the covariance containing the contributions described in Appendix 2.11 is called the ‘non-Gaussian’ covariance.
2.2.1 KSW estimator

The Komatsu, Spergel and Wandelt (KSW) estimator (Komatsu et al., 2005) offers the ability to compute the amplitude of separable or nearly separable bispectra in an efficient manner. The technique uses the separability to significantly reduce the computational cost. Consider the following reduced bispectrum:

\[
b_{\ell_1,\ell_2,\ell_3} = w_{\ell_1}w_{\ell_2}w_{\ell_3} \int dr \left( \alpha_{\ell_1}(r)\beta_{\ell_2}(r)\gamma_{\ell_3}(r) + \alpha_{\ell_2}(r)\beta_{\ell_3}(r)\gamma_{\ell_1}(r) \right) + 4 \text{ permutations of } \ell_1, \ell_2 \text{ and } \ell_3
\]  

(2.18)

where \(w_{\ell}\) is again the window function with pixelation window and the integral over \(\alpha, \beta\) and \(\gamma\) is the theoretical signal bispectrum. Many types of primordial bispectra have reduced bispectra with this structure. We have assumed that we have only one map so that, for simplicity, we can drop the \(X_i\) superscripts. Inserting this form into our estimator and rearranging we attain the following:

\[
\hat{A} \propto \int dr \int d^2\text{n} \times 
\left( \int d^2\ell_1 w_{\ell_1} \alpha_{\ell_1}(r)C_{\ell_1}^{-1}ae^{-in\cdot\ell_1} \int d^2\ell_2 w_{\ell_2}\beta_{\ell_2}(r)C_{\ell_2}^{-1}ae^{-in\cdot\ell_2} \int d^2\ell_3 w_{\ell_3}\gamma_{\ell_3}(r)C_{\ell_3}^{-1}ae^{-in\cdot\ell_3} \right) 
\left( \int d^2\ell_1 d^2\ell_2 w_{\ell_1} \alpha_{\ell_1}(r)C_{\ell_1}^{-1}ae^{-in\cdot\ell_1}w_{\ell_2}\beta_{\ell_2}(r)C_{\ell_2}^{-1}ae^{-in\cdot\ell_2} \int d^2\ell_3 w_{\ell_3}\gamma_{\ell_3}(r)C_{\ell_3}^{-1}ae^{-in\cdot\ell_3} \right) 
\right)
\]  

(2.19)

where \(\ell \equiv |\ell|\). Note that we also used the relation \(C_{\ell,\ell'}^{-1} = \langle C_{\ell}^{-1}aC_{\ell'}^{-1}a \rangle\) to reformulate the linear term. The separation of the integrals allows an efficient calculation of the estimator, and this method can be extended for templates with double integrals as we will describe in the next section.

When applied to real data we discretise the integrals over \(\ell\)'s and the resulting sums are efficiently calculated via fast Fourier transforms, for which we use the FFTW package (Frigo et al., 2005).
2.2.2 Estimator Normalization and Linear Term

We compute the estimator’s normalization and linear term using the ensemble average method described in Smith & Zaldarriaga (2011). We find the linear term is negligible for this work and this is discussed below. We outline this method here and refer the reader to the original paper for more details. First we generate sets of Gaussian simulations with power spectrum $C_{\ell}^{\text{Total}}$. To achieve this we use the CAMB (Lewis et al., 2000) code to generate the underlying CMB power spectrum and apply the beam and pixel window function, as has been presented in Hasselfield et al. (2013) for ACTPol and in Planck Collaboration VII (2014) for Planck. We add secondary sources to our CAMB power spectrum with amplitudes as determined in Dunkley et al. (2013). For Planck, the noise power is estimated from simulations (Planck Collaboration XII, 2016). For the ACTPol maps we have four splits of the data with uncorrelated noise, created by using data taken on every fourth night (see also Section 4.4). From these, we estimate the noise power, assuming it is diagonal, via the following equation:

$$N_{\ell} \approx \frac{1}{4} \sum_i a_i^* a_i - \frac{1}{12} \sum_{i \neq j} a_i^* a_j^*.$$  \hspace{1cm} (2.20)

where $a_i^*$ is the power in the $i^{th}$ split of the data. We assume that the noise is uncorrelated between the different arrays. We apply the same masks as used in the data to our simulations, when masking is required. The real noise is anisotropic and we model the anisotropy in our simulations by weighting the noise simulations in real space by the square root of the hits map.
For each Gaussian simulation, we then calculate the following quantity:

\[
\nabla_i T(a^X) = \frac{1}{2} \sum_{X_1, X_2} \int \int d^2 \ell_1 d^2 \ell_2 4\pi^2 \delta^{(2)}(\ell + \ell_1 + \ell_2) b^{i, X_1, X_2}(\ell_1, \ell_2, \ell_3) a_{\ell_1}^X a_{\ell_2}^X, \tag{2.21}
\]

where the index \( i \) labels the type of non-Gaussianity. We then use this quantity to calculate the normalization:

\[
N_{i,j} \propto \int \int d^2 \ell_1 d^2 \ell_2 \left( \frac{1}{3} \left\langle \nabla_i T(C^{-1} a^X_1) C_{\ell_1, \ell_2}^{X_1, X_2} (\nabla_j T(C^{-1} a^X_2))^* \right\rangle - \frac{1}{3} \left\langle \nabla_i T(C^{-1} a^X_1) \right\rangle C_{\ell_1, \ell_2}^{X_1, X_2} \left( (\nabla_j T(C^{-1} a^X_2))^* \right) \right) \tag{2.22}
\]

and the linear term:

\[
\int \prod_{i} d^2 \ell_i 4\pi^2 \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) b^{i, X_1, X_2, X_3}(\ell_1, \ell_2, \ell_3) C_{-\ell_1, \ell_2, \ell_3}^{-1, X_1, X_2, X_3} a \,
\]

\[
= \int d^2 \ell_3 \left\langle \nabla_i T(C^{-1} a^X_3) \right\rangle C^{-1, X_3}_{\ell_3} a. \tag{2.23}
\]

The normalization proportionality constant depends on \( f_{\text{sky}}^{(n)} \). We use this method to calculate the normalization as it was found to be more efficient than a more direct approach.

The linear term of the estimator is only important when the covariance matrix has large off diagonal terms. This can be seen by considering the case when \( C_{\ell, \ell'}^{-1, X, X'} \) is diagonal (\( \ell = \ell' \)). In this case the contribution from the linear term vanishes.

In our pipeline (described in section 2.5) we take steps to reduce the impact of mode coupling and inhomogeneous noise, which are the main sources of off-diagonal covariance terms. These steps mean that the covariance matrix only has significant off-diagonal contributions when \( \ell \sim \ell' \) and thus the main contributions to the linear
term will be from squeezed configurations (configurations with two large $\ell$ and one small $\ell$). The configuration dependence of the linear term is discussed in [Creminelli et al., 2006] and Bucher et al. (2016) demonstrate that the linear term is a minor effect for non-squeezed configurations. As discussed in section 2.5 we mask the lowest $\ell$ modes to avoid ground contamination and in doing so we strongly suppress squeezed configurations and there the linear term. For the templates used here we found that the linear term contributed negligibly to our results and that when it was included we required many more simulations for the ensemble averages to converge than if we neglected the linear term. After verifying it was negligible for all the templates we proceeded to neglect its contribution.

2.3 Bispectrum Templates

In this work we consider non-Gaussianity from lensing-induced anisotropies, the tSZ effect, DSFGs and radio galaxies. The kSZ effect is too small to be constrained with the data used in this work; the form of its bispectrum template is described in Appendix 2.9.6 for completeness. In this section we present a brief overview of the structure of the bispectrum templates that are used in this analysis and provide the details of these calculations in Appendix 2.9.

2.3.1 Poisson point source templates

Unclustered point sources in our maps follow Poisson statistics. Their reduced bispectrum is a flux weighted count of the number of point sources and has the form

$$b^{\text{Point Sources}}_{\ell_1, \ell_2, \ell_3} = k_\nu^3 \int_0^{S_c} dS_\nu S_\nu^3 \frac{d n}{d S d \Omega},$$

where $S_\nu$ is the flux density, $d n/d S d \Omega$ is the number count per steradian per flux interval, $S_c$ is the flux cut and $k_\nu$ is the conversion from flux to map temperature.
Figure 2.1: The one-halo bispectrum templates with $l_1 = l_2 = l_3$ for the tSZ, DSFGs and radio galaxies for three frequencies: 100, 148 and 217 GHz. The DSFG term contains the two and three galaxy terms but not the one galaxy term; the one galaxy term is a Poisson term with trivial scale dependence and so is excluded. Similarly, we exclude the radio point source term. The radio-tSZ is composed of two physical terms: the radio-tSZ-tSZ and the radio-radio-tSZ terms. The feature in the radio-tSZ template arises as the two contributions have different signs below 220 GHz and are dominant on different scales; the radio-tSZ-tSZ term is dominant on the largest scales and is positive at all frequencies and the radio-radio-tSZ term is dominant on the smallest scales and is negative at frequencies below 220 GHz. Similarly, the radio-DSFG term is composed of two physical terms: the radio-radio-DSFG and the radio-DSFG-DSFG terms. The scale is linear between $10^{-6}$ and $10^6$ and logarithmic outside this range. This leads to kinks in the curves that cross the log-linear transition.
Figure 2.2: The lensing cross secondary source bispectrum templates with \( l_1 = l_2 = l_3 \) at 148 GHz. We exclude the ISW-lensing template here as it has been well studied before. The scale is linear between \( 10^{-4} \) and \( 10^4 \) and logarithmic otherwise.

The flux cut is uniform across the survey area and different cuts are applied for the different frequency map as described in section 4.4. Radio galaxies and dusty star forming galaxies (DSFGs) both have a Poisson point source contribution and the details of their source counts is given in sections 2.9.1 and 2.9.3.

### 2.3.2 Thermal Sunyaev Zel’dovich, dusty star forming galaxy and radio galaxy templates

We construct templates for non-Gaussianity from auto and cross correlations from the tSZ effect, DSFG and radio galaxies with the halo model. Our analysis is focused on the smallest scales and so we only include the one halo term. Here we briefly overview the general prescription for calculating one halo bispectrum templates. Our discussion is based on that of Hill & Pajer (2013) and Lacasa et al. (2014).

In the halo model it is assumed that all the matter is contained in dark matter halos. We then assume that each quantity of interest is distributed within each cluster with profile \( X^i(x, M) \), where \( X^i \) could be the pressure of the halo gas or the distribution of radio galaxies and \( M \) is the cluster virial mass. Note that we use the index \( i \) to label different physical effects and not different halos, as is common in
the halo model literature. We also make the assumption that these profiles are all spherically symmetric. The reduced bispectrum is given by

\[ b_{\ell_1,\ell_2,\ell_3}^{ij,k} = \int d\chi_i d\chi_j d\chi_k a(\chi_i) a(\chi_j) a(\chi_k) \int dxx^2 \prod_a dk_a k_a J_{\ell_a} (k_a \chi_a) J_{\ell_a + \frac{1}{2}} (k_a \chi) \]

\[ \times \frac{1}{k_a \sqrt{\chi \chi_a}} \int d \ln M \frac{dn}{d \ln M} \tilde{X}^i(k_i, M) \tilde{X}^j(k_j, M) \tilde{X}^k(k_k, M), \]

(2.25)

where \( k_i = |k_i| \), \( \chi \) is the comoving distance, \( a \) is the scale factor, \( dn/d \ln M \) is the mass function and \( \tilde{X}(k, M) \) is the Fourier transform of the line of sight projected profile \( X(x, M) \). Using the Limber approximation we have the final form (see Appendix 2.8.1 for the details)

\[ b_{\ell_1,\ell_2,\ell_3}^{ij,k} = \int d\chi \frac{a(\chi)a(\chi)a(\chi)}{\chi^4} \int d \ln M \frac{dn}{d \ln M} \tilde{X}^i(k_i, M) \tilde{X}^j(k_j, M) \tilde{X}^k(k_k, M) \]

\[ = \int dz \frac{d^2 V}{dz d\Omega} \frac{a(\chi)^3}{\chi^6} \int d \ln M \frac{dn}{d \ln M} \tilde{X}^i \left( \frac{\ell_1 + \frac{1}{2}}{\chi(z)}, M \right) \tilde{X}^j \left( \frac{\ell_2 + \frac{1}{2}}{\chi(z)}, M \right) \tilde{X}^k \left( \frac{\ell_3 + \frac{1}{2}}{\chi(z)}, M \right). \]

(2.26)

With this formalism we construct eight templates to characterize the non-Gaussianity from auto and cross correlations among the tSZ effect, DSFG and radio galaxies. The equatorial slice for these bispectra is plotted for three different frequencies in figure 2.1. The tSZ-tSZ-tSZ template describes non-Gaussianity from the tSZ effect only; the DSFG template describes the Poisson and clustered components of the DSFGs. The tSZ-tSZ-DSFG and tSZ-DSFG-DSFG templates model the cross correlations between the tSZ effect and DSFGs. The radio galaxy template describes the Poisson contributions of the radio galaxies. The radio-tSZ template contains two terms, the radio-radio-tSZ and radio-tSZ-tSZ terms, and arises from cross correlations between radio galaxies and the tSZ effect. The radio-DSFG template, which similarly contains both the radio-radio-DSFG and radio-DSFG-DSFG terms, describes the cross correlations between the DSFGs and radio galaxies. Finally the
radio-DSFG-tSZ template describes the cross correlations among the three effects. Extending the results shown in [Lacasa et al. (2014)] we find that all the one-halo bispectrum terms are largely insensitive to the configuration but display strong scale dependence. Motivated by this we plot only equatorial slices, where $\ell_1 = \ell_2 = \ell_3$, of our templates to show the scale dependence.

2.3.3 Lensing cross secondary sources templates

Bispectra from lensing-induced anisotropies have been well studied [Spergel & Goldberg 1999a, Goldberg & Spergel 1999, Lewis et al. 2011]. We briefly summarize the origin of this non-Gaussianity here. We can decompose the anisotropies into contributions from the early universe and late-time sources:

$$\Delta T(n) = \Delta T^P(n + \nabla \phi) + \Delta T^S(n), \quad (2.27)$$

where $\Delta T^P$ is the unlensed CMB fluctuations, $\phi$ is the lensing potential, and $\Delta T^S$ encodes the contributions from late time sources such as radio galaxies or clusters via the tSZ effect. If we expand in the lensing potential to first order we find

$$\Delta T(n) = \Delta T^P(n) + \nabla \phi \cdot \nabla \Delta T^P(n) + \Delta T^S(n). \quad (2.28)$$

From this we see that we can get non-vanishing bispectra arising from the terms $\left(\Delta T^P(n) \nabla \phi \cdot \nabla \Delta T^P(n) \Delta T^S(n)\right)$ as the secondary sources are correlated with the structures that cause the lensing. More precisely these lead to the following reduced bispectrum

$$b_{\ell_1, \ell_2, \ell_3}^{\text{lensing-sec.}} = -\ell_1 \cdot \ell_2 C_{\ell_1}^{TT} C_{\ell_2}^{\phi S} + 5\text{ permutations}, \quad (2.29)$$
where $C_{\phi S}$ is the cross-correlation between the lensing potential and the secondary source. We consider four lensing bispectrum templates arising from cross-correlations with the tSZ effect, radio galaxies, DSFG and the integrated Sachs Wolfe effect. In figure 2.2 we show the equatorial slice through the bispectrum for several of the lensing templates considered in this work. The shape of the lensing templates is largely insensitive to frequency so we display only a single frequency here. The shape of the ISW-lensing template has been well studied before (Lewis et al., 2011) and so is not shown here. The other lensing templates have a similar spatial dependence so we plot only the equatorial configuration here.

2.4 Data Sets

The analysis in this paper is focused on using the small-scale information from the ACTPol experiment. Several non-Gaussian sources, such as the Poisson contributions from DSFG and radio galaxies, have similar template shapes and cannot be distinguished with the current 148GHz ACTPol maps. The different spectral behavior of
the sources enables their differentiation with multi-frequency data. For this purpose we use data from the \textit{Planck Experiment}. Through combining ACTPol and \textit{Planck} 100 GHz and 217 GHz data we use the small-scale information from ACTPol and the multifrequency data from \textit{Planck}. Most of our constraining power will come from bispectrum configurations involving one \textit{Planck} map and two ACTPol maps. Figure \ref{fig:2.3} shows that our estimator will be most sensitive to scales of $1000 < \ell < 2000$ for the \textit{Planck} maps and $2000 < \ell < 5000$ for the ACTPol maps.

\subsection*{2.4.1 ACTPol data sets}

We use data in the ‘D56’ field, a patch of sky on the equator with coordinates $-7.2^\circ < \text{dec} < 4^\circ$ and $352^\circ < \text{RA} < 41^\circ$. This is part of the data described in \cite{Louis2017}; in particular we use only the wide field and not the deep fields. In this work we use the data from the two arrays (called PA1 and PA2 hereafter) which observed the sky at 148 GHz. The ACT experiment has a full width at half maximum (FWHM) of $1.4'$ at 148 GHz. As was shown in \cite{Louis2017} non-white atmospheric noise, from atmospheric temperature brightness fluctuations, dominates the largest scales. On the smallest scales the noise is approximately white at $31 \mu K$-arcmin and $25 \mu K$-arcmin for PA1 and PA2 respectively. We mask all the point sources whose fluxes were measured to be above $30 \text{ mJy}$ with discs of radius $5'$ and perform no masking of clusters or of galactic dust.

\subsection*{2.4.2 Planck Data sets}

We use only the \textit{Planck} 100 and 217 GHz maps as the lower frequencies have too limited sensitivity to the small-scales (due to the $> 5'$ beam) and the higher frequencies are obscured by dust in this region. Instead of using the full information of the Planck maps we remain in the flat-sky regime and use only the portion of the \textit{Planck} data that overlaps with the ACTPol D56 region. The ACTPol and Planck
maps we used as input do not come with the same pixelization or coordinate system. ACTPol uses equatorial coordinates in equal-area coordinates pixelization (CEA), while Planck uses HEALPix-pixelized galactic coordinates. Before cross-correlating the maps, we reprojected Planck onto the ACTPol pixels by first expanding it in spherical harmonic coefficients and rotating these coefficients from galactic to equatorial coordinates using healpy, and then evaluating these coefficients on the ACTPol pixels using libsharp (Reinecke & Seljebotn, 2013).

We mask point sources that were detected in the ACTPol 148 GHz map with fluxes above 269 mJy and 152 mJy for the 100 GHz and 217 GHz maps respectively with discs of radii 12.5′ and 6.5′. These masking levels are Planck’s 90% completeness limit for these frequencies (Planck Collaboration XXVI, 2016). We chose a point source mask based on source detected in the ACT maps (rather than using the Planck point source mask) to ensure we had a constant flux cut across the map. We perform no masking of clusters or of galactic dust. The Planck experiment’s beam FWHM are 9.7′, and 5.0′ at 100 GHz and 217 GHz respectively. The small scale Planck noise is white with levels of 77.4 µK-arcmin and 46.8 µK-arcmin at 100 GHz and 217 GHz respectively (Planck Collaboration VIII, 2016). To avoid issues of nearly singular matrices during our inverse covariance filtering we add uncorrelated noise to the Planck maps with a power spectrum equal to 10% of the primary CMB (with the appropriate beam and window functions). Without this the largest scale modes of the Planck maps would be very highly correlated and this can lead to numerical instabilities when performing inverse covariance operations.

2.5 Analysis Pipeline

Full inverse-variance weighting of these maps is a computationally costly process and whilst methods exists for performing this (Elsner & Wandelt, 2013; Smith et al.,
2009), the Planck team has shown (Planck Collaboration XXIV, 2014) that near optimality can be obtained using approximate methods that assume the covariance matrix is diagonal. In this work we will use the approximate method described in Planck Collaboration XXIV (2014).

First we apply a point source mask that masks the brightest point sources as stated in sections 2.4.1 and 2.4.2. We then use a similar method to Planck Collaboration XXIV (2014), in which we fill in masked point source pixels with the average of their neighbors and iterate until the solution converges. Convergence is typically attained within 1000 iterations and 2000 iterations are used for our final analysis. This infilling procedure reduces mode coupling and allows the approximation of a diagonal covariance matrix without loss of optimality.

Next we apply our real space mask. The real space mask is constructed from two components. The first is the square root of the smoothed hit counts map ($H_s(n)$). The hits map was smoothed with a Gaussian of FWHM 9.5' to limit leakage. The second component is a smoothed top hat mask $F(n)$. To obtain $F(n)$, we mask all pixels with fewer than 5500 counts in either of the ACTPol arrays and then convolve this mask with a cosine squared of width 25'. This mask serves two purposes: it down-weights noisy and infrequently observed regions of the sky, thereby reducing the influence of inhomogeneous noise, and it serves as an edge taper that reduces mode coupling. The total mask $M(n)$ is given by

$$ M(n) \propto \frac{H_s(n)}{H_s(n) + 15000} F(n). $$

On the scale of our smoothed hits-map mask, the power in the map for heavily observed regions (hit counts $\geq 15,000$) is dominated by the CMB. In these regions of the maps our final mask is constructed to be approximately uniform to prevent over-weighting these regions.
The FFT methods we use assume periodic boundary conditions. To prevent spurious correlations we zero-pad our maps by a factor of two before performing a Fourier transform on the maps. Finally, we Fourier transform our maps and then apply a \( k \)-space mask to our maps. For the ACTPol maps, we mask modes with \( \ell < 100 \) and apply a sine squared filter to modes with \( 100 < \ell < 500 \); these modes are dominated by atmospheric \( 1/f \) noise. This filter reduces the leakage of power from these modes to higher \( \ell \) modes, which have less power. To avoid complications of the sine squared mask, we apply a mask to our theoretical templates for the ACTPol maps that masks all modes with \( \ell < 500 \). We then remove all modes with \( |\ell_x| < 90 \) and \( |\ell_y| < 90 \), as these are dominated by ground contamination (Das et al., 2011). For the Planck maps, we filter modes with \( \ell < 200 \) by applying a sine squared filter to modes with \( 0 < \ell < 200 \) and then applying a \( \ell < 200 \) cut to our templates. This is done for two reasons; first the flat-sky approximation becomes less accurate at low \( \ell \) and second the two and three halo terms cannot be ignored if the lowest \( \ell \) modes are used.

Once processed as above, we use the maps to estimate the amplitudes of the theoretical templates described in section 4.3. We use the KSW implementation of equation 4.7 to calculate the template amplitudes and equation 2.16 to jointly fit our templates. The Gaussian errors are calculated using equation 2.17. Our estimator jointly fits the amplitudes of all the non-Gaussian templates and uses all the possible auto and cross bispectra from the four maps. We use an \( \ell_{\text{max}} = 7500 \). This cut off was chosen as there is little signal at small scales (higher \( \ell \)) and it is computational expensive to include the smaller scales. The validation of the analysis pipeline is described in Appendix C.
2.6 Template measurements with ACTPol and Planck

In table 2.1 we present the measured template amplitudes and the Gaussian errors for these templates. The “joint fit” results come from simultaneously fitting all of the templates. The Gaussian error characterizes how well we can measure the different types of non-Gaussianity given the finite number of noisy modes in our data set. Our measured amplitudes with the Gaussian error describe the level of non-Gaussianity and its significance in this data set. We find significant amplitudes for the tSZ-tSZ-tSZ, tSZ-tSZ-DSFG, radio galaxy, and radio-DSFG-tSZ templates. When simultaneously measuring multiple templates, the Gaussian covariance represents how well the estimator can differentiate between the types of non-Gaussianity. In figure 2.4 we present the Gaussian covariance matrix for our measurements; it can be seen that, despite the use of multi-frequency information, we still have several highly correlated terms. In particular, we see strong covariances between the radio-DSFG term and the DSFG only term; between the radio-tSZ term and the tSZ only term; and between the tSZ-DSFG-DSFG term and the radio galaxy term. These strong covariances, as well as the others in the dataset, make our measurements particularly sensitive to our models, since inaccurate theoretical models can ‘leak’ the signal of one template into another, tightly correlated template.

As we have strong detections of these templates it is important to verify whether the Gaussian errors (that are valid in the limit of weak non-Gaussianity) are still dominant and, if not, to include the non-Gaussian contribution to the cosmic variance. Our method for calculating the six point function is described in Appendix 2.11. In table 2.1 we present our constraints on the amplitudes with the non-Gaussian cosmic variance errors; the measured amplitudes with the full non-Gaussian errors represents our constraints on the population amplitudes of these templates. It can be seen that
<table>
<thead>
<tr>
<th>Type</th>
<th>Measured $A_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gaussian Errors</td>
</tr>
<tr>
<td>lensing x radio</td>
<td>$-0.31 \pm 6.26$</td>
</tr>
<tr>
<td>lensing x tSZ</td>
<td>$1.74 \pm 1.55$</td>
</tr>
<tr>
<td>lensing x DSFG</td>
<td>$0.43 \pm 0.41$</td>
</tr>
<tr>
<td>lensing x ISW</td>
<td>$47.86 \pm 29.22$</td>
</tr>
<tr>
<td>tSZ-tSZ-tSZ</td>
<td>$0.80 \pm 0.25$</td>
</tr>
<tr>
<td>tSZ-tSZ-DSFG</td>
<td>$1.20 \pm 0.29$</td>
</tr>
<tr>
<td>tSZ-DSFG-DSFG</td>
<td>$-0.96 \pm 0.47$</td>
</tr>
<tr>
<td>radio-DSFG-tSZ</td>
<td>$6.03 \pm 1.31$</td>
</tr>
<tr>
<td>DSFG-DSFG-DSFG</td>
<td>$1.65 \pm 0.44$</td>
</tr>
<tr>
<td>radio-tSZ</td>
<td>$1.20 \pm 0.82$</td>
</tr>
<tr>
<td>radio-DSFG</td>
<td>$-0.45 \pm 1.40$</td>
</tr>
<tr>
<td>radio-radio-radio</td>
<td>$0.99 \pm 0.08$</td>
</tr>
</tbody>
</table>

Table 2.1: Template amplitudes obtained from a joint fit using 100 GHz and 217 GHz Planck maps with the ACTPol PA1 and PA2 148 GHz maps and $\ell_{\text{max}} = 7500$. Point sources below 269, 152, 30 mJy for frequencies 100, 217, 148 GHz respectively were masked and no masking of clusters was performed. The amplitudes, $A_i$, are overall scalings of our templates and are dimensionless. The Gaussian errors only indicate the significance of the non-Gaussianity in these data sets. The non-Gaussian error results include the one-halo contributions to the cosmic variance and these errors reflect the statistical uncertainty in the template amplitudes.

the errors for some of the terms are effectively doubled by the inclusion of the six point terms. Figure 2.6 shows our results with the Gaussian only errors and with the full non-Gaussian errors. In figure 2.5 we present the covariance matrix of the templates when we include the full six point function. The full covariance deviates significantly from the Gaussian-only covariance matrix further indicating the necessity of including the non-Gaussian components. The source of the non-Gaussian covariances differs fundamentally from those in the Gaussian case: the Gaussian covariances arise due to the similarity of the templates and the difficulty of differentiating between them with our data sets, whereas the non-Gaussian covariances arise from the physical correlations among the sources of these templates. For example, the large covariance that exists between the tSZ-tSZ-tSZ and tSZ-tSZ-DSFG template primarily arises from the Poissonian fluctuations of the number of dark matter halos, as an increase (or decrease) in the number of dark matter halos will increase (decrease) the tSZ-
Figure 2.4: The covariance matrix of template amplitudes only accounting for the Gaussian component to the six point function. The strong covariances between several templates highlights the difficult of disentangling all of the sources of non-Gaussianity.

tSZ-tSZ bispectrum and tSZ-tSZ-DSFG bispectrum in an identical and correlated manner.

2.7 Discussion, conclusions and future directions

We have applied the KSW estimator to simultaneously measure a wide range of sources of non-Gaussianity in the CMB. These first measurements with ACTPol extend the similar work done by Crawford et al. (2014) with the SPT experiment by examining cross-correlations between the sources. In future work we will have the sensitivity to constrain the bispectrum of the kinetic Sunyaev Zel’dovich effect (see
Figure 2.5: The full one-halo covariance matrix for our measured estimators. The strong covariances present in this figure, but absent in Fig 2.4 arise due to the physical covariances of these templates.

Appendix 2.9.6 for current sensitivities) and thereby fully constrain the bispectrum of known CMB secondary sources.

There have been several previous measurements of non-Gaussianity from secondary anisotropies. In table 2.2 we show the measurements of the bispectrum from the DSFG and radio galaxies from previous work. We compare our results to measurements of the bispectrum amplitude of radio sources from Komatsu et al. (2009) and Calabrese et al. (2010) for the WMAP Q, V and W bands and to measurements of the DSFGs bispectrum from Crawford et al. (2014). We also show the bispectrum contribution that we would predict from our model as constrained by the measurements. For the WMAP results we assume that the only contribution is from the radio galaxies and use a flux cut of 0.5 Jy at 22 GHz. In both of the analyses of the WMAP
Figure 2.6: A graphical representation of our joint fit results from table 2.1 excluding the lensing-ISW constraint in order to restrict the scales. The solid green line is the values predicted by our model and the dashed green line is the null value.

Table 2.2: A comparison between various different bispectrum measurements with all results given in $\mu K^3 \text{sr}^2$. The Komatsu et al. column contains the results from Komatsu et al. (2009), Calabrese et al. column contains the results from Calabrese et al. (2010) and Crawford et al. column is from Crawford et al. (2014). The final column shows what our model, with amplitudes fitted from the ACTPol and Planck data as presented here, would predict for these measurements. To compare our DSFG result to the Crawford et al. result we evaluate our model at $\ell = 3000$, 220 GHz and at their flux cut of 6.4 mJy. The factor of two difference between our results and those from Calabrese et al. (2010) is discussed in the text.
data, the authors mask all the point sources detected in Wright et al. (2009) as well as radio galaxies identified in external surveys with fluxes greater than 0.5 Jy and 22 GHz (as described in Hinshaw et al., 2007). The comparison to the WMAP Q and V bands is done by extrapolating our model beyond the frequencies measured in this work. To compare our measurements of DSFGs with Crawford et al. (2014) we have combined their measurements of the Poisson and clustered components of the DSFGs and compared this to our model’s prediction for the DSFG contribution at $\ell = 2000$ and 220 GHz. The radio Poisson contribution at this frequency with a 6.4 mJy mask is negligible. We see that our model has good agreement with the measurements in Komatsu et al. (2009), but differ from Calabrese et al. (2010) by a factor of two. A similar factor of two between the Komatsu et al. (2009) and Calabrese et al. (2010) results was discussed in Calabrese et al. (2010). The origin of this disparity is not understood. We see good agreement with the DSFG measurements, though the errors are large; a more interesting comparison will be possible with future data sets. Our measurements of the tSZ bispectrum and lensing-ISW bispectrum are consistent with previous work (Crawford et al., 2014; Planck Collaboration XXI, 2016; Wilson et al., 2012; Calabrese et al., 2010), but with large errors.

From our tSZ bispectrum template amplitude we can compute a constraint on $\sigma_8$, the amplitude of matter fluctuations smoothed on $8h^{-1}$ Mpc scales. Our analysis follows that of Crawford et al. (2014). Using priors from Planck Collaboration XIII (2016), we marginalize over the following cosmological parameters: the baryon density, $\Omega_b$; the matter density, $\Omega_m$; the Hubble constant, $H_0$; and the scalar spectral index with $n_s$. Finally we assume an astrophysical modeling uncertainty of 35% (Bhattacharya et al., 2012; Hill & Sherwin, 2013). The resulting constraint on $\sigma_8$ is

$$\sigma_8 = 0.79 \pm 0.07.$$ (2.31)
We jointly fit the tSZ amplitude with the other secondary templates which provides some robustness to possible contamination. However, large degeneracies mean that our results are sensitive to the theoretical modeling; this is discussed further in the paragraph below. We do not include any uncertainty in our $\sigma_8$ constraint from the joint modeling.

We now consider the ability of the data set used in this work to discriminate among the dozen bispectrum templates in our model. The model is not a good fit to the data, with $\chi^2 = 43.6$ for 12 degrees of freedom. This poor fit is likely due to the strong degeneracies seen in the Gaussian covariance matrix (figure 2.4) and deficiencies in the halo model templates. These degeneracies mean that small changes to the covariance matrix lead to large shifts in our results. Physical fluctuations, for example those caused by fluctuations in the number of massive dark matter halos in the patch, or deficiencies in our modeling could mean that our model covariance differs from the true Gaussian covariance of our data and could cause leakage of one template into another. The interpretation of the chi-squared is further complicated as the distribution of the template amplitudes is non-Gaussian. The full non-Gaussian probability distribution for the template amplitudes is difficult to calculate and will be the subject of future work. This suggests that the theoretical uncertainty of our results could be significant and is likely the limiting factor in our analysis; as such, we stress that our results are a proof of principle for this approach, rather than accurate measurements of these secondary anisotropies.

The issues raised above will be addressed with future measurements. Multi-frequency small-scale information, particularly at low frequencies, which will be available from Advanced ACT, SPT3G, Simons Observatory and CMB-S4 (Abazajian et al. 2016b; Henderson et al. 2016; Benson et al. 2014), should break the strong degeneracies. This will significantly reduce the dependence on the modeling and allow the use of advanced statistics, such as Skew-C$_\ell$s (Munshi & Heavens 2010), to assess
the accuracy of our templates, and to study the astrophysics of these sources. In particular, it will be interesting to see if the high amplitudes of the radio-tSZ-DSFG terms are caused by template degeneracies or are due to radio galaxies occupying lower mass clusters at redshift \( \sim 1 \) than is assumed in our model. If it is the later then it suggests that radio galaxies, DSFGs and the tSZ effect can be found in the same halos and understanding the correlations among these terms will be important for searches for the kSZ signal.

The calculation of the six point function indicated that non-Gaussian contributions for such measurements are already important and cannot be neglected in future more sensitive measurements. The contribution of the non-Gaussian component to the variance has the potential to limit the power of the bispectrum to measure astrophysics and cosmology through its reduction on the information. However there are many measures that can be employed to reduce this error. For example, masking the largest clusters will significantly reduce the importance of these terms as the six point function receives the largest contribution from low redshift and massive clusters (even more so than the bispectrum itself). An even better approach, which was done in \cite{Hurier & Lacasa 2017}, would be to jointly analyze the bispectrum and cluster counts; such an approach would allow a reduction of the errors without loss of the information from the largest clusters.

2.8 Appendix: The flat-sky -full-sky bispectrum correspondence

In this section we follow the power spectrum derivation of \cite{Loverde & Afshordi 2008} and derive a relation between the full-sky and flat-sky bispectrum.
2.8.1 Full-sky

The 2-D three-point function of three fields is

\[
\langle A(n_1)B(n_2)C(n_3) \rangle = \left\langle \int dr_1dr_2dr_3 F_A(r_1)A^{3D}(r_1,n_1)F_B(r_2)B^{3D}(r_2,n_2)F_C(r_3)C^{3D}(r_3,n_3) \right\rangle,
\]

(2.32)

where \(F_X(x)\) is the projection kernel and in harmonic space we have

\[
\langle A_{\ell_1,m_1}B_{\ell_2,m_2}C_{\ell_3,m_3} \rangle = \int d\Omega_1 d\Omega_2 d\Omega_3 \left\langle A(n_1)B(n_2)C(n_3)Y^*_{\ell_1,m_1}(n_1)Y^*_{\ell_2,m_2}(n_2)Y^*_{\ell_3,m_3}(n_3) \right\rangle.
\]

(2.33)

To proceed we decompose each 3D field into its Fourier components and then use the plane wave expansion:

\[
e^{i k \cdot x} = 4\pi \sum_{\ell,m} i^\ell j_\ell(kr)Y^*_{\ell,m}(\hat{k}) Y^*_{\ell,m}(\hat{n}).
\]

(2.34)

Performing these steps and evaluating the angular integrals leaves

\[
\langle A_{\ell_1,m_1}B_{\ell_2,m_2}C_{\ell_3,m_3} \rangle = \int dr_1dr_2dr_3 F_A(r_1)F_B(r_2)F_C(r_3)
\]

\[
\times \int \prod_i \frac{d k_i^3}{(2\pi)^3} 4\pi i^\ell_i j_\ell_i(k_i r_i)Y^*_{\ell_i,m_i}(\hat{k}) \langle \tilde{A}^{3D}(k_1)\tilde{B}^{3D}(k_2)\tilde{C}^{3D}(k_3) \rangle.
\]

(2.35)

We then assume translational and rotational invariance, i.e.

\[
\langle \tilde{A}^{3D}(k_1)\tilde{B}^{3D}(k_2)\tilde{C}^{3D}(k_3) \rangle = (2\pi)^3 \delta^{(3)}(k_1 + k_2 + k_3)B(k_1,k_2,k_3)
\]

(2.36)
so

\[ \langle A_{\ell_1,m_1} B_{\ell_2,m_2} C_{\ell_3,m_3} \rangle = \int \mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}r_3 F_A(r_1) F_B(r_2) F_C(r_3) \]

\[ \times \int \prod \frac{\mathrm{d}k_i^3}{(2\pi)^3} 4\pi i^\ell \int \frac{(\ell)}{i} (k_i r_i) Y_{\ell_i,m_i}^* (\hat{k}) B(k_1, k_2, k_3) \int \mathrm{d}x e^{i(\ell_1 + \ell_2 + \ell_3)x}. \] (2.37)

Again using the plane wave expansion and evaluating the angular integrals gives

\[ \langle A_{\ell_1,m_1} B_{\ell_2,m_2} C_{\ell_3,m_3} \rangle = G_{\ell_1}^{\ell_2} G_{\ell_3}^{\ell_3} \int \mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}r_3 F_A(r_1) F_B(r_2) F_C(r_3) \]

\[ \times \int \mathrm{d}x x^2 \prod \int \frac{\mathrm{d}k_i}{(2\pi)^3} (4\pi k_i)^2 j_\ell (k_i r_i) j_\ell (k_i x) B(k_1, k_2, k_3). \] (2.38)

where:

\[ G_{\ell_1}^{\ell_2} G_{\ell_3}^{\ell_3} = \int \mathrm{d}\Omega Y_{\ell_1,m_1}^*(n) Y_{\ell_2,m_2}^*(n) Y_{\ell_3,m_3}^*(n) \] (2.39)

is the Gaunt integral. Note that the bispectrum can now be expressed in terms of the reduced bispectrum of the full-sky, \( b_{\ell_1,\ell_2,\ell_3}^{\text{FULL}} \),

\[ \langle A_{\ell_1,m_1} B_{\ell_2,m_2} C_{\ell_3,m_3} \rangle = G_{\ell_1}^{\ell_2} G_{\ell_3}^{\ell_3} b_{\ell_1,\ell_2,\ell_3}^{\text{FULL}} \] (2.40)

We will focus on the reduced bispectrum. Next we replace the spherical Bessel functions, \( j_\ell (x) \), with Bessel functions of the first kind, \( J_{\ell + \frac{1}{2}} (x) = j_\ell (x) \sqrt{\frac{2x}{\pi}} \) to obtain

\[ b_{\ell_1,\ell_2,\ell_3}^{\text{FULL}} = \int \mathrm{d}r_1 \mathrm{d}r_2 \mathrm{d}r_3 F_A(r_1) F_B(r_2) F_C(r_3) \]

\[ \times \int \mathrm{d}x x^2 \prod \int \frac{\mathrm{d}k_i k_i^2}{(2\pi)^3} J_{\ell + \frac{1}{2}} (k_i r_i) J_{\ell + \frac{1}{2}} (k_i x) \frac{1}{k_i \sqrt{\ell x}} B(k_1, k_2, k_3). \] (2.41)
For notational simplicity we will define

$$f_i(r) = \frac{F_A(r)}{\sqrt{r}}.$$  \hspace{1cm} (2.42)

We can now use the relation (Loverde & Afshordi, 2008)

$$\int dx J_\nu(x)f(x) = f(\nu) - \frac{1}{2} f''(\nu) - \frac{\nu}{6} f'''(\nu) + ....$$

$$\int dx J_\nu(kx)f(x) = \frac{f(\nu k)}{k} - \frac{1}{2k^3} f''\left(\frac{\nu}{k}\right) - \frac{\nu}{6k^4} f'''\left(\frac{\nu}{k}\right) + ....$$ \hspace{1cm} (2.43)

to get

$$b_{\ell_1,\ell_2,\ell_3}^{\text{FULL}} = \int dx x^2 \prod_i \int d\nu_i J_{\ell_i+\frac{1}{2}}(k_i x) \left( f_A(r_1) - \frac{f''_A(r_1)}{2k_1^2} - \frac{\nu_1^2}{6k_1^4} f'''_A(r_1) + .... \right)$$

$$\left( f_B(r_2) - \frac{f''_B(r_2)}{2k_2^2} - \frac{\nu_2}{6k_2^4} f'''_B(r_2) + .... \right)$$

$$\left( f_C(r_3) - \frac{f''_C(r_3)}{2k_3^2} - \frac{\nu_3}{6k_3^4} f'''_C(r_3) + .... \right) B(k_1, k_2, k_3)$$ \hspace{1cm} (2.44)

with $\nu_i = \ell_i + \frac{1}{2} = k_i r_i$. Applying that relation again and retaining only the lowest terms in $\nu_i$ leads to

$$b_{\ell_1,\ell_2,\ell_3}^{\text{FULL}} = \int dx x^4 F_A(x) F_B(x) F_C(x) B \left( \frac{\nu_1}{x}, \frac{\nu_2}{x}, \frac{\nu_3}{x} \right) \left( 1 - \sum_i \frac{1}{6\nu_i^3} \left[ \left( \frac{d \ln(f_i B)}{d \ln k_i} \right)^3 + \left( \frac{d \ln f_i}{d \ln x} \right)^3 \right] + \frac{d^3 \ln(f_i B)}{d \ln^3 k_i} + \frac{d^3 \ln f_i}{d \ln^3 x} + 3 \frac{d^2 \ln(f_i B)}{d \ln^2 k_i} \frac{d \ln(f_i B)}{d \ln k_i} + 3 \frac{d^2 \ln f_i}{d \ln^2 k_i} \frac{d \ln f_i}{d \ln x} - \frac{d \ln B}{d \ln k_i} \right) + O\left( \nu^{-4} \right).$$ \hspace{1cm} (2.45)

Note that we now have $\nu_i = \ell_i + \frac{1}{2} = k_i r_i = k_i x$.  

42
2.8.2 Flat-sky

In the flat-sky regime we have

\[
\langle A(x_1^\perp)B(x_2^\perp)C(x_3^\perp) \rangle = \\
\left\langle \int dx_1^\parallel dx_2^\parallel dx_3^\parallel F_A(x_1^\parallel) A^{3D}(x_1^\parallel, x_1^\perp) F_B(x_2^\parallel) B^{3D}(x_2^\parallel, x_2^\perp) F_C(x_3^\parallel) C^{3D}(x_3^\parallel, x_3^\perp) \right\rangle
\]

(2.46)

where \(x^\parallel\) is the coordinate along the line of sight and \(x^\perp\) is perpendicular to the line of sight. Expanding in flat-sky Fourier modes gives

\[
\langle \tilde{A}(k_1^\perp) \tilde{B}(k_2^\perp) \tilde{C}(k_3^\perp) \rangle = \prod_i \int \frac{dk_i^\parallel}{2\pi} F_i(x_i^\parallel) e^{i k_i^\parallel x_i^\parallel} \left\langle \tilde{A}^{3D}(k_1^\perp, k_1^\parallel) \tilde{B}^{3D}(k_2^\perp, k_2^\parallel) \tilde{C}^{3D}(k_3^\perp, k_3^\parallel) \right\rangle.
\]

(2.47)

Assuming translational and rotational invariance we can then write

\[
\langle \tilde{A}^{3D}(k_1^\perp, k_1^\parallel) \tilde{B}^{3D}(k_2^\perp, k_2^\parallel) \tilde{C}^{3D}(k_3^\perp, k_3^\parallel) \rangle = (2\pi)^3 \delta^{(3)}(k_1 + k_2 + k_3) B(k_1, k_2, k_3)
\]

(2.48)

so

\[
\langle \tilde{A}(k_1^\perp) \tilde{B}(k_2^\perp) \tilde{C}(k_3^\perp) \rangle = \prod_i \int dx_i^\parallel \frac{dk_i^\parallel}{2\pi} F_i(x_i^\parallel) e^{i k_i^\parallel x_i^\parallel} (2\pi)^3 \delta^{(3)}(k_1 + k_2 + k_3) B(k_1, k_2, k_3).
\]

(2.49)

This can be written in terms of the flat-sky bispectrum:

\[
\langle \tilde{A}(k_1^\perp) \tilde{B}(k_2^\perp) \tilde{C}(k_3^\perp) \rangle = (2\pi)^2 \delta^{(2)}(k_1^\perp + k_2^\perp + k_3^\perp) b^{\text{FLAT}}(k_1^\perp, k_2^\perp, k_3^\perp)
\]

(2.50)
with

$$b^{\text{FLAT}}(k_1^\perp, k_2^\perp, k_3^\perp) = \prod_i \int d\xi_i \frac{d k_i^\parallel}{2\pi} F_i(\xi_i^\parallel) e^{i k_i^\parallel \xi_i^\parallel} (2\pi) \delta(k_i^\parallel + k_i^\parallel + k_3^\parallel) B \left( \sqrt{k_1^\parallel + k_1^\perp}, \sqrt{k_2^\parallel + k_2^\perp}, \sqrt{k_2^\parallel + k_3^\perp} \right).$$

(2.51)

Expanding $B \left( \sqrt{k_1^\parallel + k_1^\perp}, \sqrt{k_2^\parallel + k_2^\perp}, \sqrt{k_2^\parallel + k_3^\perp} \right)$ about $k_i^\parallel$ we obtain

$$b^{\text{FLAT}}(k_1^\perp, k_2^\perp, k_3^\perp) = \int dx F_A(x) F_B(x) F_C(x) B(k_1, k_2, k_3) \left( 1 + \frac{1}{2(k_1 x)^2} \frac{d \ln F_A}{d \ln x} \frac{d \ln F_B}{d \ln x} \frac{d \ln B}{d \ln k_1} + \frac{1}{2(k_2 x)^2} \frac{d \ln F_A}{d \ln x} \frac{d \ln F_C}{d \ln x} \frac{d \ln B}{d \ln k_2} + \frac{1}{2(k_3 x)^2} \frac{d \ln F_B}{d \ln x} \frac{d \ln F_C}{d \ln x} \frac{d \ln B}{d \ln k_3} \right).$$

(2.52)

Expanding $1/x^4 B(\nu_1, \nu_2, \nu_3)$ in equation 2.45 around the peak, $\tilde{r}$, of $F_A(x) F_B(x) F_C(x)$ we find $k^4 b^{\text{FLAT}}(k_1^\perp, k_2^\perp, k_3^\perp) \approx \ell^4 b^{\text{FULL}}_{\ell_1, \ell_2, \ell_3}$ with $l_i + 1/2 = k_i \tilde{r}$. Note that as in the power spectrum case, the convergence of these series depend both on $\ell + 1/2$ and on the projection kernels, if the projection kernels are peaked at different $r_i$ then more terms will be required in this series.

### 2.9 Appendix: Detailed Bispectrum Templates

In this appendix we briefly summarize the origin and bispectrum structure of the various significant secondary sources of non-Gaussianity. While the kSZ effect is too small to be constrained with the data used in this work we discuss the form of its bispectrum for completeness.
In this section we define a deconvolved reduced bispectrum, $\tilde{b}$, where we have factored out the beam and pixelization function, as

$$b_{\ell_1,\ell_2,\ell_3}^{X_1,X_2,X_3} \equiv w_{\ell_1} w_{\ell_2} w_{\ell_3} \tilde{b}_{\ell_1,\ell_2,\ell_3}$$

(2.53)

and for notational convenience we suppress the map indices $\{X_1, X_2, X_3\}$ on the reduced bispectrum.

### 2.9.1 Radio galaxies

Radio galaxies are point sources with strong synchrotron emission and include AGN, flat-spectrum radio quasars (FSRQs) and BL Lacertae type objects (BL Lacs). It has been shown (Toffolatti et al., 1998; González-Nuevo et al., 2005) that at CMB frequencies radio galaxies are unclustered. Thus radio galaxies follow Poisson statistics and their reduced bispectrum is given by:

$$\tilde{b}_{\text{radio-radio-radio}} = k_\nu^3 \sum_i \int_0^{S_c} dS_v^{(i)} S_v^{(i)}^3 \frac{dn^{(i)}}{dSd\Omega}$$

(2.54)

where we sum over the different populations of radio galaxies, $S_v^{(i)}$ is the flux density, $dn^{(i)}/dSd\Omega$ is the number count per steradian per flux interval, $S_c$ is the flux cut and $k_\nu$ is the conversion from flux to map temperature. We compute the theoretical level and frequency dependence using the model described in de Zotti et al. (2010). The model has three components, each described by their own number counts and with flux densities $S \propto \nu^{-\alpha}$. The components are flat-spectrum radio quasars and BL Lacs with $\alpha_{\text{FSRQ}} = \alpha_{\text{BLLac}} = 0.1$, and a population of steep spectrum sources, arising from AGNs, with $\alpha_{\text{steep}} = 0.8$. We extend de Zotti et al. (2010) by using their model for differential source counts at 100 GHz and steepening the spectral indices of the three components in agreement with the results from Tucci et al. (2011). The bispectrum
model presented is a simple generalization from the power spectrum for radio point sources which has the form

\[ C_{\ell}^{\text{radio}} = k_{\ell}^2 \sum_i \int_0^{S_c} dS_{\nu}^{(i)} S_{\nu}^{(i)} dS_{\nu}^{(i)}^2 \frac{dn^{(i)}}{dS d\Omega}. \] (2.55)

The bispectrum is weighted more towards the most luminous galaxies than the power spectrum and, given the relatively slow decrease of the source counts as a function of flux \((dn/dS d\Omega \propto S^{-1.5})\) the bispectrum of radio galaxies is dominated by the brightest sources and thus is sensitive to the level of point source masking.

### 2.9.2 Thermal Sunyaev Zel’’dovich

Light propagating from the surface of last scattering to the present day can be up-scattered by hot gas in a process known as the thermal Sunyaev Zel’’dovich effect (tSZ). The magnitude of this effect depends upon the integrated line-of-sight electron pressure, therefore this effect is dominated by galaxy groups and clusters where there is a large amount of hot intercluster gas. The observable signature depends on the frequency of observation; it can be seen as a decrement below frequencies of 220 GHz and an excess at higher frequencies. We briefly review the tSZ bispectrum as developed in Bhattacharya et al. (2012). It should be noted that we only include effects from the one-halo term as previous work (Komatsu & Kitayama, 1999) demonstrated that for the angular scales of interest in the power spectrum the two halo term is negligible. We have verified that this extends to the bispectrum. In the non-relativistic approximation the temperature fluctuation at angular position \(\theta\) from the center of a cluster of mass \(M\) at redshift \(z\) is proportional to the integral along the line-of-sight.
of the electron pressure:

\[
\frac{\Delta T(\theta, M, z)}{T} = f(x_\nu)y(\theta, M, z)
\]

\[
= f(x_\nu)\frac{\sigma_T}{m_e c^2} \int_{\text{LOS}} dr P_e(\sqrt{r^2 + D_A(z)^2|\theta|^2}, M, z)
\]

(2.56)

where \( f(x_\nu) = x_\nu(\coth(x_\nu/2) - 4) \) is the spectral function of the SZ effect, \( x_\nu = k_b T_{\text{CMB}} / h \nu \) with \( \nu \) the observing frequency, \( D_A(z) \) is the angular diameter distance, and \( P_e \) is the electron pressure. The Fourier transform of the projected SZ profile is (Komatsu & Seljak, 2002):

\[
\tilde{y}(\ell, M, z) = \frac{4\pi r_s}{\ell_s^2} \frac{\sigma_T}{m_e c^2} \int_0^\infty dx x^2 P_e(M, Z, x) \frac{\sin(x\ell/\ell_s)}{x\ell/\ell_s},
\]

(2.57)

where \( x = r/r_s \) with \( r_s \) a characteristic radius of the cluster which we take to be \( r_s = r_{\text{vir}}/c_{\text{NFW}} \) following Navarro et al. (1997), \( \ell_s = r_s/D_A(z) \). As is consistent with the literature we have included the projection (a) and Limber terms (1/\( \chi^2 \)) in our definition of \( \tilde{y}(\ell, M, z) \). The power spectrum can be calculated by summing the square of the Fourier transform of the projected SZ-profile weighted by the mass function giving:

\[
C^{\text{SZ}}_\ell = f(x_\nu_1)f(x_\nu_2) \int_0^\infty dz \frac{d^3V}{dzd\Omega} \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{y}(\ell, M, z)\tilde{y}(\ell, M, z),
\]

(2.58)

where \( \frac{dn}{d\ln M} \) is the halo mass function (for which we used Tinker et al. 2008) and \( \frac{dV}{d\xi d\Omega} \) is the comoving volume per steradian. In this work we use cluster pressure profiles obtained from simulations (Battaglia et al. 2012b) and we verified our results are not significantly changed when using observationally constrained profiles (Arnaud et al. 2010, Planck Collaboration Int. V. 2013). Similarly the one halo reduced bispectrum is obtained by summing the cube of the Fourier transform of the projected SZ-profile
weighted by the mass function giving:

\[
\tilde{b}_{\ell_1,\ell_2,\ell_3}^{\text{SZ-sSZ-sSZ}} = f(x_{\nu_1}) f(x_{\nu_2}) f(x_{\nu_3}) \int_0^\infty dz \frac{d^2V}{dz d\Omega} \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{y}(\ell_1, M, z) \tilde{y}(\ell_2, M, z) \tilde{y}(\ell_3, M, z).
\]

(2.59)

2.9.3 Dusty star-forming Galaxies

Dusty star-forming galaxies (DSFGs) are high redshift \((z \sim 2 - 3)\), dusty, star-forming galaxies; the UV radiation emitted by the newly formed massive stars is absorbed by the dust and re-emitted at longer wavelengths. In CMB maps these objects appear as point sources which are dominant at high frequencies but still important for the frequencies considered here [Hauser & Dwek, 2001]. We model these sources with two components: a Poisson component and a clustered component. We use the halo model [Cooray & Sheth, 2002] as a unified way to calculate both the theoretical DSFG template and the cross-correlations with other physical effects. In this we follow the works of Lacasa et al. (2014), Addison et al. (2012) and Xia et al. (2012) and we combine the halo model with halo occupation statistics to populate dark matter halos with DSFGs [Zheng et al., 2005]. We use the assumption that the properties of the DSFGs depend only upon redshift and use the parametric model by Béthermin et al. (2012) to model the source properties. This model incorporates two populations of galaxies, main sequence and starburst, with redshift evolution. This parametric model was fit to differential counts of mm objects from Herschel and Spitzer data [Béthermin et al., 2010; Berta et al., 2011a; Glenn et al., 2010]. The two and three halo terms, which have been measured by the Planck satellite [Planck Collaboration XXX, 2014], are only important on the largest scales [Addison et al., 2012; Lacasa et al., 2014]. As we have very limit sensitivity to scales with \(\ell < 1000\), we only consider the one-halo and Poisson terms that dominate the small scales.
For a detailed description of this bispectrum we refer the reader to Lacasa et al. (2014) and here we summarize their results. We define the $n^{th}$ order differential source intensity as:

$$\frac{dI^{(n)}(\nu_1, \nu_2, ..., \nu_n)}{dV} = \int_0^{S_{\text{cut}}} dS_{\nu_1}S_{\nu_1}dS_{\nu_2}...dS_{\nu_n} \frac{d^2n(S_{\nu_1}, z)}{dVdS}, \quad (2.60)$$

$S_{\nu_i}$ is the observer frame flux at frequency $\nu_i$, $S_{\text{cut}}$ is the flux above which all sources are masked and $d^2n(S_{\nu_1}, z)/dVdS$ is the differential count and is obtained from the Béthermin et al. (2012) model. Note that this can be related to the generalized differential source emissivity $j^{(n)}(\nu_1, \nu_2, ..., \nu_n)$ (Knox et al., 2001) used in Lacasa et al. (2014) by

$$\frac{dI^{(n)}(\nu_1, \nu_2, ..., \nu_n)}{dV} = \frac{a^n}{\chi_{2n}} j^{(n)}(\nu_1, \nu_2, ..., \nu_n); \quad (2.61)$$

in other words we absorbed the projection operator $(a)$ and the Limber terms $(1/\chi^2)$ into our differential source count. Through the halo occupation model prescription we have

$$\langle N_{\text{gal}} \rangle = \langle N_{\text{cen}} \rangle + \langle N_{\text{sat}} \rangle \quad (2.62)$$

with

$$\langle N_{\text{cen}} \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log(M) - \log(M_{\text{min}})}{\sigma_{\log(M)}} \right) \right], \quad (2.63)$$

and

$$\langle N_{\text{sat}} \rangle = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log(M) - \log(2M_{\text{min}})}{\sigma_{\log(M)}} \right) \right] \left( \frac{M}{M_{\text{sat}}} \right)^{\alpha_{\text{sat}}}, \quad (2.64)$$

where $M$ is the halo mass, and $M_{\text{min}}$, $\sigma_{\log(M)}$, $M_{\text{sat}}$ and $\alpha_{\text{sat}}$ parameterize the occupation of the halos. $M_{\text{min}}$ gives the mass at which a halo has a 50% probability of having a central galaxy and $\sigma_{\log(M)}$ describes the width of the transition from no central galaxy to one central galaxy. $M_{\text{sat}}$ and $\alpha_{\text{sat}}$ describe the number of satellite
galaxies. The parameters for the HOD model were fit to Planck power spectra in Pépin et al. (2014) and we use these values in this work. Finally we assume that the distribution of the DSFGs follows that of the dark matter profile in the halo. Thus the Fourier transform of the 3D DSFG spatial distribution, $\tilde{u}_{\text{DSFGs}}$ is

$$\tilde{u}_{\text{DSFG}}(\ell, M, z) = \tilde{u}_{\text{dm}}(\ell, M, z) = \int_0^{r_{\text{vir}}} dr \frac{\sin((\ell + 1/2)r/\chi)}{(\ell + 1/2)r/\chi} \rho(r| M, z)}{M}$$

(2.65)

where $\tilde{u}_{\text{dm}}(\ell, M, z)$ is the Fourier transformed of the normalized halo profile. Bringing all these pieces together we can write down the power spectrum of the DSFG. There are two components a one galaxy term, called the Poisson component

$$C_{\ell}^{\text{DSFG-Pois}} = \int_0^{\infty} dz \frac{d^2 V}{dz d\Omega} \frac{dI^{(2),\text{DSFG}}(\nu_1, \nu_2)}{dV}.$$

(2.66)

and two galaxy, or clustered component,

$$C^{\text{DSFG-Clustered}} = \int_0^{\infty} dz \frac{d^2 V}{dz d\Omega} \frac{dI^{(1),\text{DSFG}}(\nu_1)}{dV} \frac{dI^{(1),\text{DSFG}}(\nu_2)}{dV} \times \int_0^{\infty} d\ln M \frac{dn}{d\ln M} \tilde{u}_{\text{dm}}(\ell_1, M, z) \tilde{u}_{\text{dm}}(\ell_2, M, z) \frac{\langle N_{\text{gal}}(N_{\text{gal}} - 1) \rangle}{n_{\text{gal}}^2}.$$

(2.67)

For the bispectrum there are three DSFG components. The three galaxy term

$$\tilde{b}_{\ell_1, \ell_2, \ell_3}^{\text{DSFG-3gal}} = \int_0^{\infty} dz \frac{d^2 V}{dz d\Omega} \frac{dI^{(1),\text{DSFG}}(\nu_1)}{dV} \frac{dI^{(1),\text{DSFG}}(\nu_2)}{dV} \frac{dI^{(1),\text{DSFG}}(\nu_3)}{dV} \times \int_0^{\infty} d\ln M \frac{dn}{d\ln M} \tilde{u}_{\text{dm}}(\ell_1, M, z) \tilde{u}_{\text{dm}}(\ell_2, M, z) \tilde{u}_{\text{dm}}(\ell_3, M, z) \frac{\langle N_{\text{gal}}(N_{\text{gal}} - 1)(N_{\text{gal}} - 2) \rangle}{n_{\text{gal}}^3},$$

(2.68)
the two galaxy term,

\[
\tilde{b}_{DSFG-2gal}^{\ell_1, \ell_2, \ell_3} = \int_0^\infty \frac{d^2V}{dzd\Omega} \frac{dI^{(2),DSFG}(\nu_1, \nu_2)}{dV} \frac{dI^{(1),DSFG}(\nu_3)}{dV} \\
\times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{u}_{dm}(\ell_1, M, z) \tilde{u}_{dm}(\ell_2 + \ell_3, M, z) \langle N_{gal}(N_{gal} - 1) \rangle \frac{1}{\bar{n}_{gal}^2},
\]

(2.69)

and the one galaxy term:

\[
\tilde{b}_{DSFG-1gal}^{\ell_1, \ell_2, \ell_3} = \int_0^\infty \frac{d^2V}{dzd\Omega} \frac{dI^{(3),DSFG}(\nu_1, \nu_2, \nu_3)}{dV}.
\]

(2.70)

Typically the first two terms are called the clustered component and the last term is called the Poisson component. In this work we fix the relative amplitude of these three contributions and fit for one overall amplitude, hereafter called the DSFG-DSFG-DSFG component. The source counts for the Poisson component decrease rapidly as a function of flux and as such it is largely insensitive to the cut level used.

We also express the \(N_{gal}\) averages in terms of \(N_{cen}\) and \(N_{sat}\). Assuming that \(N_{sat}\) follows a Poisson distribution, we obtain the following (see Appendix 2.12):

\[
\langle N_{gal}(N_{gal} - 1) \rangle = 2\langle N_{sat} \rangle + \langle N_{sat} \rangle^2
\]

(2.71)

and

\[
\langle N_{gal}(N_{gal} - 1)(N_{gal} - 2) \rangle = 3\langle N_{sat} \rangle^2 + \langle N_{sat} \rangle^3.
\]

(2.72)
The halo formalism allows the cross-correlation between the DSFGs and the tSZ effect to be easily calculated. The tSZ-DSFG power spectrum is given by

\[
C_{tSZ-DSFG}^{\ell} = f(x_{\nu_1}) \int_0^\infty dz \frac{d^2V}{d\Omega d\ell} \frac{dI_{1,DSFG}(\nu_2)}{dV} \left[ \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{y}(\ell, M, z) \tilde{u}(\ell, M, z) \frac{N_{DSFG}(M)}{\bar{n}_{DSFG}} \right] + \int_0^\infty d\ln M \frac{dn}{d\ln M} b(M) \tilde{y}(\ell, M, z) \times \int_0^\infty d\ln M' \frac{dn}{d\ln M} b(M') \tilde{u}(\ell, M', z) \frac{N_{DSFG}(M')}{\bar{n}_{DSFG}} \bigg] \left\{ \ell + \frac{1}{2} \chi(z), z \right\}. \tag{2.73}
\]

There are two bispectra arising from cross correlations between the tSZ effect and DSFGs:

\[
\tilde{b}_{tSZ-DSFG-DSFG}^{\ell_1,\ell_2,\ell_3} = \int_0^\infty dz \frac{d^2V}{d\Omega d\ell} \left[ f(x_{\nu_1}) \frac{dI_{1,DSFG}(\nu_2)}{dV} \frac{dI_{1,DSFG}(\nu_3)}{dV} \right] \times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{y}(\ell_1, M, z) \tilde{u}_{dm}(\ell_2, M, z) \tilde{u}_{dm}(\ell_3, M, z) \frac{N_{gal}(N_{gal} - 1)}{\bar{n}_{gal}^2} + f(x_{\nu_1}) \frac{dI_{2,DSFG}(\nu_2, \nu_3)}{dV} \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{y}(\ell_1, M, z) \tilde{u}_{dm}(|\ell_2 + \ell_3|, M, z) \frac{N_{gal}}{\bar{n}_{gal}} \bigg] \tag{2.74}
\]

and

\[
\tilde{b}_{tSZ-tSZ-DSFG}^{\ell_1,\ell_2,\ell_3} = \int_0^\infty dz \frac{d^2V}{d\Omega d\ell} f(x_{\nu_1}) f(x_{\nu_2}) \frac{dI_{1,DSFG}(\nu_2)}{dV} \times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{y}(\ell_1, M, z) \tilde{y}(\ell_2, M, z) \tilde{u}_{dm}(\ell_3, M, z) \frac{N_{gal}}{\bar{n}_{gal}} \bigg]. \tag{2.75}
\]

### 2.9.4 Cross-bispectra among radio galaxies, the tSZ effect and DSFGs

Whilst the radio galaxies show little clustering, recent work by Gralla et al. (2014); Gupta et al. (2017) has found evidence for a correlation between radio galaxies and tSZ sources, since radio galaxies occupy the same clusters that source the tSZ effect.
Building on this we consider the cross-bispectra that arise from correlations among the tSZ effect, DSFGs and radio galaxies. There are five possible bispectra: radio-tSZ-tSZ, radio-radio-tSZ, radio-DSFG-DSFG, radio-radio-DSFG and radio-DSFG-tSZ. In this work we fix the relative amplitude of the radio-tSZ-tSZ and radio-radio-tSZ terms, hereafter this joint term is called the radio-tSZ template, and the relative amplitude of the radio-DSFG-DSFG, radio-radio-DSFG terms, hereafter the joint term is called the radio-DSFG template. These terms were combined as they cannot be distinguished with our current data sets. This is because the radio terms are expected to be small and most important at low frequencies, which can not be probed with the Planck 100 GHz data. The tSZ-DSFG templates discussed in section 2.9.3 were not combined as they are expected to be larger than the radio templates and the Planck beam at 217 GHz is smaller allowing the small scales to be more accurately measured.

To estimate the radio halo occupation, we use the model from Wake et al. (2008), which is also reexamined in Smolčić et al. (2011). In Wake et al. (2008), the authors constrain HOD parameters with data from a set of combined surveys (see Sadler et al. (2007) for more details). They are unable to constrain the parameters associated with satellite galaxies, implying most halos host only one radio galaxy, and as such we use only a simplified HOD model:

\[ N_{\text{radio}} = e^{-\frac{M^*}{M}} , \]

(2.76)

where \( M \) is the halo mass and \( M^* = 9.65 \times 10^{13} h^{-1} M_\odot \) was as measured in Wake et al. (2008). We combine this model with the halo model and force the radio galaxies to lie in the center of the cluster as was preferred in Smolčić et al. (2011) and Wake et al. (2008) so that \( \tilde{r}(\ell, M, z) = 1 \). Combining this with the results of the previous section we can write down the power spectrum for radio, tSZ and DSFG cross correlations.
\[ C_{\ell}^{\text{radio-tSZ}} = f(x_1) \int_0^\infty dz \frac{d^2V}{d\Omega} \frac{dI^{(1),\text{radio}}(\nu_2)}{dV} \left[ \int_0^\infty \frac{d\ln M}{d\ln M} \frac{dn}{d\ln M} \tilde{y}(\ell, M, z) \tilde{r}(\ell, M, z) \left( \frac{N_{\text{radio}}(M)}{n_{\text{radio}}} \right) ight. \\
+ \left. \int_0^\infty d\ln M \frac{dn}{d\ln M} b(M) \tilde{y}(\ell, M, z) \int_0^\infty d\ln M' \frac{dn}{d\ln M} b(M') \tilde{r}(\ell, M', z) \left( \frac{N_{\text{DSFG}}(M')}{n_{\text{DSFG}}} \right) P_{\text{lin}} \left( \frac{\ell + 1/2}{\chi(z)}, z \right) \right], \tag{2.77} \]

and

\[ C_{\ell}^{\text{radio-DSFG}} = \int_0^\infty dz \frac{d^2V}{d\Omega} \frac{dI^{(1),\text{DSFG}}(\nu_1)}{dV} \frac{dI^{(1),\text{radio}}(\nu_2)}{dV} \left[ \int_0^\infty \frac{d\ln M}{d\ln M} \frac{dn}{d\ln M} \tilde{u}_{\text{DM}}(\ell, M, z) \tilde{r}(\ell, M, z) \left( \frac{N_{\text{DSFG}}(M) N_{\text{radio}}(M)}{n_{\text{DSFG}} n_{\text{radio}}} \right) ight. \\
+ \left. \int_0^\infty d\ln M \frac{dn}{d\ln M} b(M) \tilde{u}_{\text{DM}}(\ell, M, z) \left( \frac{N_{\text{DSFG}}(M)}{n_{\text{DSFG}}} \right) ight. \\
\times \left. \int_0^\infty d\ln M' \frac{dn}{d\ln M} b(M') \tilde{r}(\ell, M', z) \left( \frac{N_{\text{radio}}(M')}{n_{\text{radio}}} \right) P_{\text{lin}} \left( \frac{\ell + 1/2}{\chi(z)}, z \right) \right]. \tag{2.78} \]

At the power spectrum level it is very challenging to separate the contributions from tSZ-DSFG, radio-tSZ and radio-DSFG terms. The strength of the cross-correlations between these sources is still poorly understood. Small-scale power spectrum measurements provide constraints on these terms (George et al., 2015; Dunkley et al., 2013) whilst the tSZ-DSFG cross-correlation at large scales has been explored in Planck Collaboration XXIII (2016). Our model predicts a correlation coefficient between the DSFGs and the tSZ effect of 0.28 at \( \ell = 3000 \). This is higher than was seen in George et al. (2015) who found \( 0.113^{+0.057}_{-0.051} \), but consistent with other previous work (Addison et al., 2012).
Similarly, the radio-DSFG template is:

\[
\begin{align*}
\tilde{b}^{\text{radio-DSFG}}_{\ell_1, \ell_2, \ell_3} &= f(x_{v_2}) f(x_{v_3}) \int_0^\infty d\ell \frac{d^2V}{dzd\Omega} \frac{dI^{(1),\text{radio}}(v_1)}{dV} \\
&\times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{r}(\ell, M, z) \bar{y}(\ell, M, z) \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} + f(x_{v_3}) \\
&\times \int_0^\infty d\ell \frac{d^2V}{dzd\Omega} \frac{dI^{(2),\text{radio}}(v_1, v_2)}{dV} \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{r}(|\ell_1 + \ell_2|, M, z) \tilde{y}(\ell_3, M, z) \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}}
\end{align*}
\]

(2.79)

Similarly, the radio-DSFG template is:

\[
\begin{align*}
\tilde{b}^{\text{radio-DSFG}}_{\ell_1, \ell_2, \ell_3} &= \int_0^\infty d\ell \frac{d^2V}{dzd\Omega} \frac{dI^{(1),\text{radio}}(v_1)}{dV} \frac{dI^{(1),\text{DSFG}}(v_2)}{dV} \frac{dI^{(1),\text{DSFG}}(v_3)}{dV} \\
&\times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{r}(\ell_1, M, z) \bar{u}_{\text{dm}}(\ell_2, M, z) \bar{u}_{\text{dm}}(\ell_3, M, z) \frac{N_{\text{DSFG}}(M)}{\bar{n}_{\text{DSFG}}} \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} \\
&\times \langle \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} \rangle + \int_0^\infty d\ell \frac{d^2V}{dzd\Omega} \frac{dI^{(1),\text{radio}}(v_1)}{dV} \frac{dI^{(2),\text{DSFG}}(v_2, v_3)}{dV} \\
&\times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{r}(\ell, M, z) \bar{u}_{\text{dm}}(|\ell_2 + \ell_3|, M, z) \frac{N_{\text{DSFG}}(M)}{\bar{n}_{\text{DSFG}}} \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} \\
&\times \int_0^\infty d\ell \frac{d^2V}{dzd\Omega} \frac{dI^{(2),\text{radio}}(v_1, v_2)}{dV} \frac{dI^{(1),\text{DSFG}}(v_3)}{dV} \\
&\times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{r}(|\ell_1 + \ell_2|, M, z) \bar{u}_{\text{dm}}(\ell_3, M, z) \frac{N_{\text{DSFG}}(M)}{\bar{n}_{\text{DSFG}}} \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} 
\end{align*}
\]

(2.80)

and the final template, the radio-DSFG-tSZ term, is:

\[
\begin{align*}
\tilde{b}^{\text{radio-DSFG-tSZ}}_{\ell_1, \ell_2, \ell_3} &= f(x_{v_3}) \int_0^\infty d\ell \frac{d^2V}{dzd\Omega} \frac{dI^{(1),\text{radio}}(v_1)}{dV} \frac{dI^{(1),\text{DSFG}}(v_2)}{dV} \\
&\times \int_0^\infty d\ln M \frac{dn}{d\ln M} \tilde{r}(\ell_1) \bar{u}(\ell_2) \bar{y}(\ell_3) \frac{N_{\text{DSFG}}(M)}{\bar{n}_{\text{DSFG}}} \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}}.
\end{align*}
\]

(2.81)

To calculate these templates we need to know the joint radio-DSFG halo occupation model. In this work we assume they are independent so that \(\langle N_{\text{DSFG}}(M) N_{\text{radio}}(M) \rangle = \langle N_{\text{DSFG}}(M) \rangle \langle N_{\text{radio}}(M) \rangle\). This assumption could overestimate the size of the correlated
terms as radio galaxies often correspond to strong AGN that could quench the star formation. An exploration of this is left to future work.

2.9.5 Cross-bispectra between lensing and secondary sources

As was described in section 2.3.3 we can decompose the temperature anisotropies, having expanded the effect of lensing to first order in $\phi$, as

$$\Delta T(n) = \Delta T^p(n) + \nabla \phi \cdot \nabla \Delta T^p(n) + \Delta T^s(n).$$

(2.82)

and this leads to a bispectrum with the form

$$b^\text{lensing-sec.}_{\ell_1, \ell_2, \ell_3} = -\ell_1 \cdot \ell_2 C^{TT}_{\ell_1} C^{\phi S}_{\ell_2} + 5 \text{ permutations}.$$  

(2.83)

Lewis et al. (2011) showed that the effect of higher order terms in $\phi$ lead to $\sim 10\%$ corrections to the perturbative result in equation 4.45 and they showed these higher order terms could accurately be approximated by replacing $C^{TT}$ with the lensed temperature power spectrum. Unlike the other bispectra listed here, this source of non-Gaussianity has a significant contribution from polarized maps through the lensing term. In this work we only analyze temperature maps so do not consider the polarized contribution in this work.

In the following subsections we describe our models of the cross-correlations between the lensing potential and the secondary sources. The two halo term has a larger contribution to the power spectrum than in the three-point function so we include its contribution in these terms.

**Thermal Sunyaev Zel’dovich lensing cross-correlation**

The cross-correlation between the lensing potential and thermal SZ effect is calculated in Hill & Spergel (2014) and Battaglia et al. (2015) and has two contributions, the
one and two halo terms are

\[ C_{\ell}^{\phi-SZ} = \int \frac{dz}{d\Omega} \frac{d^2V}{dzd\Omega} \int dM \frac{dn(M, z)}{d\ln M} \tilde{y}(\ell, M, z) \tilde{\phi}(\ell, M, z) + \int \frac{dz}{d\Omega} \frac{d^2V}{dzd\Omega} \times \int dM \frac{dn(M, z)}{d\ln M} b(M) \tilde{y}(\ell, M, z) \int dM' \frac{dn(M', z)}{d\ln M'} b(M') \tilde{\phi}(\ell, M', z) P_{lin} \left( \frac{\ell + 1/2}{\chi(z)}, z \right) \]

(2.84)

where \( \tilde{\phi} \) is the lensing potential profile for a cluster

\[ \tilde{\phi}(\ell, M, z) = \frac{2}{\ell(\ell + 1)} \frac{4\pi r_{s,\phi}}{l_{s,\phi}^2} \int dx x^2 \sin((\ell + 1/2)x_{s,\phi}/l_{s,\phi}) \rho(x_{s,\phi}, M, z) \frac{\Sigma_{crit}(z)}{\Sigma_{crit}(z)}, \]

(2.85)

\( r_{s,\phi} \) is a characteristic scale radius for the cluster, in our case the NFW scale radius, \( \rho \) is the cluster density, given by the NFW profile, \( \Sigma_{crit} \) is the critical surface density for lensing at redshift \( z \)

\[ \Sigma_{crit}^{-1}(z) = \frac{4\pi G \chi(z)(\chi_* - \chi(z))}{c^2 \chi_* (1 + z)}, \]

(2.86)

where \( \chi_* \) is the comoving distance to the surface of last scattering, and the other terms are as in section 2.9.2.

**DSFGs lensing cross-correlation**

The DSFG-lensing cross-correlation has been detected (e.g. Omori et al., 2017; van Engelen et al., 2015; Planck Collaboration XVIII, 2014). We use the halo model for the DSFGs as described in section 2.9.3 to model the lensing-DSFGs cross-correlation.
Through this formalism we obtain the following relations

\[ C_{\ell}^{\phi-\text{DSFG}} = \int \frac{d^2 V}{dzd\Omega} \frac{dI^{(1),\text{DSFG}}(v_1)}{dV} \int dM \frac{dn(M, z)}{d\ln M} \tilde{u}_{\text{dm}}(\ell, M, z) \tilde{\phi}(\ell, M, z) \left\langle \frac{N_{\text{gal}}(M)}{\bar{n}_{\text{gal}}} \right\rangle + \int \frac{d^2 V}{dzd\Omega} \frac{dI^{(1),\text{DSFG}}(v_1)}{dV} \int dM \frac{dn(M, z)}{d\ln M} b(M) \tilde{u}_{\text{dm}}(\ell, M, z) \left\langle \frac{N_{\text{gal}}(M)}{\bar{n}_{\text{gal}}} \right\rangle \times \int dM' \frac{dn(M', z)}{d\ln M'} b(M') \tilde{\phi}(\ell, M', z) P_{\text{lin}} \left( \ell + \frac{1}{2}, \chi(z), z \right). \]

(2.87)

Radio–Lensing cross-correlation

Given the discussion above in section 2.9.4 we should expect a cross-correlation between the \( \phi \) field and radio sources. This will have the same structure as the DSFG-lensing cross-correlation:

\[ C_{\ell}^{\phi-\text{radio}} = \int \frac{d^2 V}{dzd\Omega} \frac{dI^{(1),\text{radio}}(v_1)}{dV} \int dM \frac{dn(M, z)}{d\ln M} \tilde{u}_{\text{dm}}(\ell, M, z) \tilde{\phi}(\ell, M, z) \left\langle \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} \right\rangle + \int \frac{d^2 V}{dzd\Omega} \frac{dI^{(1),\text{radio}}(v_1)}{dV} \int dM \frac{dn(M, z)}{d\ln M} b(M) \tilde{u}_{\text{dm}}(\ell, M, z) \left\langle \frac{N_{\text{radio}}(M)}{\bar{n}_{\text{radio}}} \right\rangle \times \int dM' \frac{dn(M', z)}{d\ln M'} b(M') \tilde{\phi}(\ell, M', z) P_{\text{lin}} \left( \ell + \frac{1}{2}, \chi(z), z \right). \]

(2.88)

ISW-Lensing cross-correlation

As was noted in Verde & Spergel (2002) and seen in Planck Collaboration XIX (2014); Planck Collaboration XXI (2016), secondary CMB anisotropies caused by the late time decay of the gravitational potential, the integrated Sachs Wolfe effect (Sachs & Wolfe 1967), and their growth through non-linear structure (Rees & Sciama 1968) are correlated with CMB lensing leading to bispectrum with the above form. This bispectrum is primarily important on the largest scales and so we expect very limited sensitivity to this bispectrum with ACTPol. We use CAMB to calculate the
lensing-ISW cross-correlation, $C_{\text{ISW}-\phi}$ and combine this with equation 4.45 to form a bispectrum template.

### 2.9.6 The kinetic Sunyaev Zel’dovich effect

The kinetic Sunyaev Zel’dovich effect (kSZ) arises due to CMB photons being scattered from moving gas and this motion imparts a Doppler shift to the CMB photons. This effect can be either positive or negative (depending on the direction of the motion of gas) and so in the bispectrum analysis only kSZ squared terms are non vanishing. In this section we outline an approximate template for the kSZ bispectrum and discuss its measurability with our data set. This approach to measure the kSZ bispectrum is complimentary to other methods such as those discussed in Hill et al. (2016); Ferraro et al. (2016); Doré et al. (2004); DeDeo et al. (2005).

We use the halo model to calculate simplified one-halo templates for cross-correlations of the kSZ squared with other tracers. We stress that these are only approximate templates and that there will be significant contributions from two and three halo terms, which are not considered here. The temperature shift caused by the kSZ effect is given by Ostriker & Vishniac (1986) Sunyaev & Zeldovich (1972)

$$\frac{\Delta T(n)}{T} = -\sigma_T \int \frac{d\chi}{1 + z} e^{-\tau(z)} n_e(\chi n, \chi) \frac{v_e}{c} \cdot n$$

where $\tau(z)$ is the optical depth, $n_e$ is the electron number density and $v_e$ is the electron velocity. In our one-halo approximation the deconvolved reduced bispectrum arising from the kSZ cross secondary is

$$\tilde{b}_{\ell_1,\ell_2,\ell_3}^{\text{kSZ-kSZ}-X} =$$

$$\sigma_T^2 \int_0^\infty dz \frac{d^2V}{dzd\Omega} \int_0^\infty d\ln M \frac{dn}{d\ln M} n_e(M, z) \tilde{\tilde{n}}_e(M, z, \ell_1) \tilde{\tilde{n}}_e(M, z, \ell_2) \tilde{X}(M, z, \ell_3) \frac{\sigma_{ve}^2(M)}{3} e^{-2\tau(z)}$$

(2.90)
where \(\tilde{n}_e(M, z, \ell_1)\) is the projected line of sight integral of the electron number density, \(\sigma_{vc}(M)\) is the cluster velocity dispersion and \(\tilde{X}(M, z, \ell_3)\) is \(f(v)\tilde{y}(M, z, \ell)\) for the kSZ-kSZ-tSZ bispectrum and \(\frac{df}{dz}\tilde{u}(M, z, \ell)\) for the kSZ-kSZ-DSFGs bispectrum. We approximate the cluster velocity dispersion as the velocity of the halo given by peaks theory as (Bardeen et al., 1986)

\[
\sigma_{vc}(M) = \sigma_v(M) \sqrt{1 - \frac{\sigma_0^4}{\sigma_1^2\sigma_{-1}^2}},
\]

where:

\[
\sigma_v = (faH)^2\sigma_{-1},
\]

\[
\sigma_j = \frac{1}{2\pi} \int dk P(k)W(Rk)k^{2j+2},
\]

and \(f = \frac{d\ln(g)}{d\ln(a)}\) is the derivative of the growth function. The optical depth is given by the integral of the mean ionized electron number density to the halo

\[
\tau(z) = \sigma_T \int \frac{d\chi}{1+z}\tilde{n}_e(z).
\]

To close these equations we need the electron number density profile and, motivated by the results in Schaan et al. (2016), we consider a Gaussian distribution. The Gaussian distribution describes the profile of the electrons in the halos. The width of the Gaussian is given by halo’s the virial radius and we normalize the profile by fixing the baryon mass to be given by \(f_bM_{dm}\) and then using the mean molecular weight and the mean molecular weight per free electron to convert to the total number of electrons. The choice of the Gaussian is purely phenomenological. A physical model for the distribution should be related to the tSZ effect as the same electrons are involved in both processes.
We see that whilst the sensitivities for the individual templates is close to having constraining power, when the templates are jointly fit the similarity of the templates to our data sets leaves no constraining power.

We then performed a fisher forecast to assess their measurability. In table 2.3 we present the expected signal to noise from fitting just the three kSZ templates separately, fitting a combined kSZ template and jointly fitting these templates with the templates considered in section 4.3. When we fit the combined template we assume that the relative amplitude of the three different kSZ templates is fixed and that only a possible overall scaling can exist. This is a simplification but the results show that even with such simplifications we have no constraining power current for measuring the kSZ effect; as such we exclude this template from our analysis. It should be noted the combined template is not constrained better than the individual templates, despite being a sum of the other contributions, due to the opposite signs of the tSZ-kSZ-kSZ and DSFG-kSZ-kSZ terms. It should be noted that if we could break the degeneracies between these templates and the other secondary source templates, such that the joint fit and single fit errors are similar, then future experiments should be able to start constraining these terms. The degeneracy between these templates will be reduced for experiments with multi-frequency small-scale measurements.

<table>
<thead>
<tr>
<th>Type</th>
<th>S/N</th>
<th>S/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>kSZ-kSZ-tSZ</td>
<td>0.45</td>
<td>0.019</td>
</tr>
<tr>
<td>kSZ-kSZ-DSFG</td>
<td>0.31</td>
<td>38.1</td>
</tr>
<tr>
<td>kSZ-kSZ-radio</td>
<td>0.15</td>
<td>0.026</td>
</tr>
<tr>
<td>Combined kSZ Template</td>
<td>0.30</td>
<td>0.032</td>
</tr>
</tbody>
</table>

Table 2.3: The expected signal-to-noise the kinetic Sunyaev Zel’dovich templates.
2.10 Appendix: Pipeline Validation Tests

2.10.1 Validation Templates

In this section we describe the suite of tests that we have preformed to verify the accuracy and efficiency of our analysis. To validate our pipeline we used three types of bispectra that can be easily simulated; these are Poisson noise, pseudo-local and primordial local non-Gaussianity. We did not investigate the amplitudes of these templates in the real data and only used them for validation. These types were chosen due to their ease of simulation, which can be challenging for arbitrary non-Gaussianity.

To cross check the analysis pipeline described in section 2.5 (hereafter pipeline 1) we implemented a second pipeline based on Das et al. (2009) (hereafter pipeline 2) which is described in section 2.10.3. For most of these analyses we limit the $\ell$ range of our analysis to be $\ell < 4000$, as Pipeline 2 is restricted to these scales.

2.10.2 Non-Gaussian Simulations

In order to validate our pipelines we need to verify that we can accurately measure non-Gaussianity in the maps. In general we use the method described in Smith & Zaldarriaga (2011), hereafter called the indirect method, to generate maps with arbitrary bispectra. We briefly summarize this method here. First we generate a Gaussian map (as in 2.2.2) and then perform the following operation to its Fourier coefficients ($a^L_\ell$):

$$ a^{NL}_\ell = a^L_\ell + \frac{1}{6} \sum_{X_1,X_2} \int d^2\ell_1 d^2\ell_2 4\pi^2 \delta^{(2)}(\ell + \ell_1 + \ell_2) b^{X_1,X_2}_{\ell,\ell_1,\ell_2} C^{-1}_{\ell_1} a^L_{\ell_1} C^{-1}_{\ell_2} a^L_{\ell_2} \quad (2.95) $$

where $b_{\ell,\ell_1,\ell_2}$ is the reduced bispectrum of the non-Gaussianity that you wish to generate. This method is useful for testing our pipelines although higher order correlations

62
are not accurately reproduced. The method of generating this non-Gaussianity is similar to the method used to calculate the normalization of our estimator, and so this method would not be able to identify errors that jointly affect this method and the normalization calculation.

As a cross check of this method and for further robustness of our analysis pipeline we generate several types of non-Gaussianity via direct methods. For this, we consider the three templates Poisson, pseudo-local and local non-Gaussianity.

**Poisson Noise**

The Poisson noise bispectrum has a trivial structure:

\[ \tilde{b}_{\ell_1,\ell_2,\ell_3}^{\text{const}} = \text{Const.} \quad (2.96) \]

This template is the same as the radio point source term or the DSFG Poisson component, except we have removed any frequency dependence to simplify the template. Poisson non-Gaussianity is simulated by adding Poisson noise to the maps.

**Pseudo-Local bispectrum**

This template is inspired by the primordial local bispectrum and it is generated by adding \( A_{\text{pseudo-local}}/T_{\ell MB}^2 (\Delta T(n) - \bar{\Delta T})^2 \) to the map. This has a bispectrum of the following form:

\[ \tilde{b}_{\ell_1,\ell_2,\ell_3}^{\text{pseudo-local}} = 2C(\ell_1)C(\ell_2) + 3 \text{ permutations.} \quad (2.97) \]

This type of non-Gaussianity is not physical. We use it to validate our pipelines due to the simplicity of simulating it and as it is orthogonal to the other test templates.
Primordial Local non-Gaussianity

Primordial local non-Gaussianity is a physical primordial type of non-Gaussianity, so-called as it generated by the real space interactions of fields and is thus local. The KSW estimator (Komatsu et al., 2005) used in this work was originally used to constrain this type of non-Gaussianity. In on-going work we are using the ACTPol data to primordial non-Gaussianity, including local non-Gaussianity, and so it is important to check our pipeline on primordial templates as well. Physically it arises when two short wavelength modes interact with a longer wavelength mode. The local primordial three point structure is given by (Falk et al., 1993; Gangui et al., 1994; Verde et al., 2000; Wang & Kamionkowski, 2000; Komatsu & Spergel, 2001a)

$B^{\text{local}}_{\Phi}(k_1, k_2, k_3) = 2f_{\text{local}}^{NL}(P_{\Phi}(k_1)P_{\Phi}(k_2) + P_{\Phi}(k_1)P_{\Phi}(k_3) + P_{\Phi}(k_2)P_{\Phi}(k_3))$

$= 2A^{2}f_{\text{local}}^{NL}\left(\frac{1}{k_1^{4-n_s}} \frac{1}{k_2^{4-n_s}} \frac{1}{k_3^{4-n_s}} + \frac{1}{k_1^{4-n_s}} \frac{1}{k_3^{4-n_s}} + \frac{1}{k_2^{4-n_s}} \frac{1}{k_3^{4-n_s}}\right), \quad (2.98)$

where $n_s$ is the primordial spectral tilt and $A$ is the amplitude of fluctuations. From this we obtain the reduced bispectrum as

$B_{\ell_1, \ell_2, \ell_3}^{\text{prim}} = \frac{8}{\pi^3} \int dr \, r^2 \prod_i \left(\int dk_i k_i^2 g^{X_i}_\ell(k_i) j_\ell(k_i r)\right) B_{\Phi}(k_1, k_2, k_3), \quad (2.99)$

where $g^{X_i}_\ell(k_i)$ is the transfer function that can be computed with a Boltzmann code such as CAMB (Lewis et al., 2000).

Local non-Gaussianity can be generated by first generating a realization of primordial Gaussian fluctuations, $\Phi(x)$ and then modifying the primordial fluctuations in the following manner:

$\Phi^{NL}(x) = \Phi(x) + A_{\text{local}} \left(\Phi^2(x) - \langle \Phi^2(x) \rangle\right). \quad (2.100)$
We then apply the transfer functions to evolve these fluctuations to the present. We use algorithms described in Elsner & Wandelt (2009) and Liguori et al. (2003) to efficiently generate non-Gaussian maps at high resolution. We use their method to generate full-sky maps and then cut out a patch that is appropriate for our fields. This method exploits the fact that the \( a_{lm} \) are related to primordial fluctuations in the following manner

\[
a_{lm} = \int dr r^2 \Phi_{lm}(r) \alpha_l(r),
\]

where \( \alpha_l(r) \) are the real space transfer function and \( \Phi_{lm}(r) \) are harmonic transforms of the real space primordial potential fluctuations. We draw \( \Phi_{lm}(r) \) from a Gaussian distribution, with the appropriate radial correlations, perform a spherical harmonic transform of these, square them, subtract the variance and transform back. Then we use equation (2.101) to evolve these perturbations and attain the late time \( a_{lm} \). This quantity is then added back to the original map scaled to the appropriate \( f_{NL} \) to add local non Gaussianity. We then transform these to a HEALpix map (Górski et al., 2005) and cut out a patch to use in our flat-sky analysis. We refer the reader to Liguori et al. (2003) for further details on this method.

### 2.10.3 Pre-whitening Analysis Pipeline

This pipeline is based on Das et al. (2009) and uses the same real space and Fourier masks as described in section 2.5. We pre-whiten the maps by using a disc differencing method; this involves convolving two copies of the maps with discs of radius \( R = 1.5' \) and \( 3R \) and then differencing them. We then add back a fraction of the original maps to these maps and then convolve the maps with a Gaussian. This pre-whitening reduces the mode coupling when we apply the point source and the mask (as above). When we move to Fourier space we deconvolve these effects to attain the desired maps. This method is limited to probing a reduced range of \( \ell \) space as deconvolving.
the disc-differencing involves a sinc function, and the signal near the zero of the sinc function cannot be accurately recovered. For this reason we restrict the use of this method to testing our pipelines at $\ell < 4000$.

### 2.10.4 Fisher Information Comparisons

It is important to assess how the variance of our estimators compares to the theoretical limits. A Fisher analysis shows that the expected variance of our estimator should be (in the diagonal covariance case)

$$
\sigma^2(A_{sec}^\alpha) = F_{\alpha,\alpha}^{-1}
$$

$$
F_{\alpha,\beta} = \frac{f_{\text{sky}}^3 f_{\text{sky}}^3}{6} \int \prod_i d^2t_i (2\pi)^2 \delta(t_1 + t_2 + t_3) b_{\ell_1 \ell_2 \ell_3} b_{\ell_1 \ell_2 \ell_3} C_{\ell_1} C_{\ell_2} C_{\ell_3}.
$$

We generate sets of 100 Gaussian simulations of maps as would be measured by the ACTPol PA2, including beam effects and anisotropic noise as described section 2.2.2. In table 2.4 we show the expected standard deviation of our estimators against the measured standard deviation. As mentioned above pipeline 2 only works for $\ell < 4000$ and so when we consider $\ell > 4000$ we limit our analysis to pipeline 1. We see generally good agreement between the different methods, with pipeline 2 having slightly higher variance. There is close agreement between our estimator variances with the Fisher estimations, which implies that our methods are close to satisfying the Cramér-Rao bound.

### 2.10.5 Simulated non-Gaussian Maps

It is important to verify that our estimators are able to accurately recover known levels of non-Gaussianity. To do this we first consider the three types that we can directly simulate: the constant, pseudo-local and local form. We compared direct
Table 2.4: A comparison between the expected and measured standard deviations of our estimator when applied to simulated Gaussian maps. The consistency of our estimator is a check of the optimality of our implementation, especially our approximate inverse covariance filtering.

<table>
<thead>
<tr>
<th>Type</th>
<th>Max $\ell$</th>
<th>$\sigma_{\text{Fisher}}$</th>
<th>$\sigma_{\text{measured, Pipeline 1}}$</th>
<th>$\sigma_{\text{measured, Pipeline 2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poisson Noise</td>
<td>3000</td>
<td>8.2</td>
<td>8.0</td>
<td>8.8</td>
</tr>
<tr>
<td>Poisson Noise</td>
<td>7500</td>
<td>0.19</td>
<td>0.21</td>
<td>–</td>
</tr>
<tr>
<td>Pseudo Local</td>
<td>3000</td>
<td>65</td>
<td>68</td>
<td>87</td>
</tr>
<tr>
<td>Local</td>
<td>3000</td>
<td>430</td>
<td>420</td>
<td>450</td>
</tr>
<tr>
<td>tSZ-tSZ-tSZ</td>
<td>3000</td>
<td>0.23</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>tSZ-tSZ-tSZ</td>
<td>7500</td>
<td>0.085</td>
<td>0.086</td>
<td>–</td>
</tr>
</tbody>
</table>

non-Gaussian simulations to non-Gaussianity generated by the method described in Smith & Zaldarriaga (2011) and section 2.10.2. By confirming the accuracy of the Smith & Zaldarriaga (2011) method, we can then use it to verify the accuracy of estimators on the remaining templates, which are much more challenging to simulate directly.

In figure 2.7 we compare the accuracy of our simulation methods for several different types of non-Gaussianity with various different cut offs obtained from sets of 100 simulations of PA2 observations. We see that there is generally good agreement between the different methods and pipelines. However we see that the residuals are predominantly slightly negative, this indicates a slight bias in our analysis at the level of $\sim 1\%$. We believe this arises from ignoring the linear term in our normalization. For the measurements reported here this effect is not significant but for future precision measurements this should be accounted for. It should be noted that the variances of the two non-Gaussian simulation methods for large non-Gaussianity is different. This is because the method in Smith & Zaldarriaga (2011) accurately reproduces the power spectrum and bispectrum but does not reproduce the higher order statistics, which results in an inaccurate variance for the bispectrum (as the variance of the bispectrum is a six point function).
Figure 2.7: A set of graphs demonstrating how our pipelines can recover an input level of non-Gaussianity, input $A_i$, for a set of different templates with different max $\ell$. The amplitudes, $A_i$, are overall scalings of our templates and so are dimensionless. We compare our two methods of simulating non-Gaussian templates, the “direct” method and the “indirect” method, as described in section 2.10.2. Having verified that our estimators are unbiased on this set of templates we then use the “indirect” method to validate our full range of templates, which cannot easily be verified by direct simulation methods.

Another useful diagnostic is to ensure that our pipelines measure similar levels of non-Gaussianity on individual maps, as well as for the statistical ensemble. In figure 2.8 we view the recovered amplitudes, $A_i$, for our local and constant estimator on a map-by-map basis for our two analysis pipelines. We see that in general there is reasonable agreement between the two methods. Together, these tests give us confidence that our estimators are unbiased and efficient.
Figure 2.8: A map-by-map comparison of our estimators applied to maps with local type and Poisson noise non-Gaussianity simulated with the direct method. The input level of non-Gaussianity for local type this is $A_{\text{local}} = 2000$ and for Poisson noise this is $A_{\text{Poisson}} = 10.6$. It can be seen the two pipelines agree well on a map by map basis, increasing the confidence in the accuracy of our pipelines.

### 2.11 Appendix: Template Errors

When the non-Gaussianity is weak the variance on the template is given by the inverse of the normalization term. However as we have detections of these templates, and further as we wish to estimate the template amplitude rather than just the significance of the deviation from null, it is important to either verify that these errors are still dominant or, in the case that they are subdominant, calculate the correct error bars. For simplicity we will consider the case when we only have a single map, but the generalization is simple. From our estimator the full covariance is given as:

$$
\langle (\hat{A}_i - \hat{A}_i^\text{mean})(\hat{A}_j - \hat{A}_j^\text{mean}) \rangle = \\
\sum_{i',j'} N_{i,j}^{-1} N_{j,i}^{-1} \sum_{\ell_1,\ell_2} \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) \delta^{(2)}(\ell_4 + \ell_5 + \ell_6) b_{i_1,i_2,i_3} b_{j_1,j_2,j_3} C_{\ell_1}^{-1} C_{\ell_2}^{-1} C_{\ell_3}^{-1} C_{\ell_4}^{-1} C_{\ell_5}^{-1} C_{\ell_6}^{-1} \\
\left( C_{\ell_1} C_{\ell_2} C_{\ell_3} \delta(\ell_1 + \ell_4) \delta(\ell_2 + \ell_5) \delta(\ell_3 + \ell_6) + \delta(\ell_1 + \ell_4) b_{\ell_1,\ell_2,\ell_3} b_{\ell_4,\ell_5,\ell_6} + \delta(\ell_1 + \ell_4) C_{\ell_1} T_{\ell_2,\ell_3,\ell_5,\ell_6} \\
+ S_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6} + V_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6} + \text{permutations} \right) \\
$$

(2.103)
where $S$ is the fully connected six point function, $T$ is the trispectrum and $V^{SSV}$ is the halo sample variance or the super sample variance \cite{Hamilton2006}. The first term is the ‘Gaussian’ term described in equation 2.17. To calculate these contributions we use the halo model and keep only the one-halo term. This results in the following form for the trispectrum:

$$T_{\ell_2, \ell_3, \ell_5, \ell_6} = \sum_{i, j, k, l} \int_0^\infty \frac{d^2V}{dz d\Omega} \int_0^\infty d \ln M \frac{dn}{d \ln M} X_{\ell_2}^i X_{\ell_3}^j X_{\ell_5}^k X_{\ell_6}^l$$

(2.104)

where $X_{\ell}^i$ is the Fourier transform of the integrated line of sight halo property, i.e. $X^1$ could be the Compton Y parameter and $X^2$ could be the DSFG profile etc. We sum over all the contributions to the trispectrum from the tSZ effect, DSFGs and radio galaxies (including both the Poisson and clustered terms). Similarly for the six point function:

$$S_{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6} = \sum_{i, j, k, l, m, n} \int_0^\infty \frac{d^2V}{dz d\Omega} \int_0^\infty d \ln M \frac{dn}{d \ln M} X_{\ell_1}^i X_{\ell_2}^j X_{\ell_3}^k X_{\ell_4}^l X_{\ell_5}^m X_{\ell_6}^n$$

(2.105)

where we again sum over all of the contributions listed above. The super sample variance term can be though of additional variance induced by the presence of modes longer than the survey volume. For example, if the region is located near a peak of such a long mode then more regions will have density fluctuations greater than the level required for collapse, resulting in more halos. We refer the reader to \cite{Takada2013} and \cite{Schaan2014} for a more detailed description of the super sample variance term. As is shown in \cite{Kayo2013} and \cite{Schaan2014} the one-halo term is most important for the bispectrum and has the following form:

$$V^{SSV}_{\ell_1, \ell_2, \ell_3, \ell_4, \ell_5, \ell_6} = \sum_{i, j, k, l, m, n} \int_0^\infty \frac{d^2V}{dz d\Omega} \sigma_s^2(V_s) \frac{dn}{d \ln M} b(m) \int_0^\infty d \ln M \frac{dn}{d \ln M} b(m) X_{\ell_4}^l X_{\ell_5}^m X_{\ell_6}^n$$

(2.106)
where $b(m)$ is the linear bias and $\sigma_s(V_s)$ encapsulates the finite survey volume by:

$$\sigma^2_s(V_s) = \frac{1}{4\pi^2} \int d^2k |W(k)|^2 P_{\text{lin}}(k, z)$$

(2.107)

where $W(k)$ is the 3D Fourier transform of the 3D survey window function (Krause & Eifler, 2017).

With these pieces, equation 2.103 could be directly calculated by using the halo model to calculate $b_{\ell_1,\ell_2,\ell_3}, b_{\ell_4,\ell_5,\ell_6}, S_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6}$ etc. However such a direct calculation is very challenging and so in this work we use an ensemble average method to calculate these terms. In particular we need to exploit some form of factorization as the six point matrix $S_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6}$ for modern surveys is extremely large ($O(10^{40})$ elements for our data) and so cannot be calculated directly. Instead we use methods inspired by the non-Gaussian simulation methods described in Smith & Zaldarriaga (2011) and section 2.10.2. As is done in Smith & Zaldarriaga (2011) and above in equation 2.95 we first calculate:

$$a_{\ell}^{NL} = \frac{1}{2} \sum_{X_1, X_2} \int d^2\ell_1 d^2\ell_2 4\pi^2 \delta^{(2)}(\ell + \ell_1 + \ell_2) b_{\ell,\ell_1,\ell_2} C_{\ell_1}^{-1} a_{\ell_1} C_{\ell_2}^{-1} a$$

(2.108)

This has the property that $\langle a_{\ell_1} a_{\ell_2} a_{\ell_3}^{NL} \rangle = 4\pi^2 \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) b_{\ell_1,\ell_2,\ell_3}$. We calculate this for two different Gaussian realizations, $a^{(1)}, a^{(2)}$ and then calculate the following quantity:

$$\int \prod_{i=1}^6 d^2\ell S_{\ell_1,\ell_2,\ell_3,\ell_4,\ell_5,\ell_6} C_{\ell_1}^{-1} a^{(1)} C_{\ell_2}^{-1} a^{(1)} C_{\ell_3}^{-1} a^{NL(1)} C_{\ell_4}^{-1} a^{(2)} C_{\ell_5}^{-1} a^{(2)} C_{\ell_6}^{-1} a^{NL(2)}$$

(2.109)

and similar terms for the other components of the six point function. As with the calculation of the estimator, we compute this by utilizing the separability of the integrand of the double integral. This means for each redshift and mass in our numerical calculation of $S$ (or the other six point terms) we factorize the above quantity and
efficiently calculate the sums by FFTs. Explicitly we would evaluate the term in equation 2.105 in the follow manner. First, for notational clarify, we define

\[ f^{(1)}(1), L(M, z) = \sum_i \sum_{\ell_1} X_i^{\ell_1} C_{\ell_1}^{-1} a^{L(1)}, \]  
\[ f^{(1),NL}(M, z) = \sum_i \sum_{\ell_1} X_i^{\ell_1} C_{\ell_1}^{-1} a^{NL(1)}. \]  

\( f^{(2)}(1), L(M, z) \) and \( f^{(2),L}(M, z) \) are defined in an analogous manner. Then equation 2.105 can be written as

\[ \sigma_{6, pnt}^2 \propto \int_0^\infty dz \int_0^\infty \frac{d^2V}{d^2\Omega} \int_0^\infty dM \frac{dn}{d\ln M} \int_0^\infty dz d\Omega \int_0^\infty d\ln M \frac{dn}{d\ln M} f^{(1)}(1), L(M, z) f^{(1),L}(M, z) f^{(1),NL}(M, z) \times \]
\[ f^{(2),L}(M, z) f^{(2),L}(M, z) f^{(2),NL}(M, z) \]  

(2.111)

It is easily seen that the ensemble average of equation 2.109 generates the corresponding term in equation 2.103. The calculation of the \( n \)th order HOD moments required for the DSFG terms is described in Appendix 2.12. By using two different Gaussian simulations we find that the majority of the contributions converge very rapidly, to within 10 percent for a single pair of simulations. We found that a handful of contributions require large numbers of simulations to converge; these terms arise from Poisson contributions to the trispectrum term such as

\[ C_{\ell_1, \ell_4} T_{\ell_2, \ell_3, \ell_5, \ell_6} = \sum_i \int_0^\infty dz \frac{d^2V}{d^2\Omega} \frac{df^{(3)}(v_1, v_2, v_5)}{dz} \int_0^\infty d\ln M \frac{dn}{d\ln M} u(\ell_2 + \ell_3 + \ell_5) X_i^{\ell_6}. \]  

(2.112)

As these terms are sub dominant, we neglect their contribution so that we can use orders of magnitude fewer simulations to obtained converged errors.
2.12 Appendix: $N^{th}$ order HOD moments

When calculating $n$ point functions of DSFG fluctuations we need expectations of drawing $n$ different galaxies from a cluster, i.e.:

$$\langle N(N-1)(N-2)\ldots(N-(n-1)) \rangle$$

(2.113)

Simulations show that the number of galaxies in a cluster has first a step, then a plateau and then a power law (Kravtsov et al., 2004; Berlind et al., 2003). This is understood as the galaxies being split into two: a central and satellite galaxies (with the central galaxy not necessarily being in the center). As the mass of the cluster increases the probability of having a central galaxy increases, this is the step and plateau, above a certain mass the cluster can also host satellites, whose number follows a power law. The HOD model parameterizes the probability of central and satellite galaxies. Only clusters with central galaxies can host satellite galaxies, so the probability distribution of satellite is conditional on the central galaxies, i.e.:

$$P(N_s) = \sum_{N_c} P(N_s|N_c)P(N_c) = P(N_s|N_c = 1)P(N_c = 1),$$

(2.114)

where $N_c$ is the number of central galaxies and $N_s$ is the number of satellites. We model the distribution of central galaxies as a Bernoulli distribution, as given in eq. 2.63 and the satellite galaxies with a Poisson distribution, as described in eq. 2.64.

With these pieces we can compute the desired expectations values. It can be shown
that for \( n \geq 2 \) the expectation values can be written in the following form:

\[
\langle N(N-1)(N-2)...(N-(n-1)) \rangle \\
= \langle (N_c + N_s)(N_c + N_s - 1)(N_c + N_s - 2)...(N_c + N_s - (n-1)) \rangle \\
= \left( f(N_c, N_s)N_c(N_c - 1) + nN_c \prod_{i=0}^{i=n-2} (N_s - i) + \prod_{i=0}^{i=n-1} (N_s - i) \right).
\]

(2.115)

The first term in the above always vanishes as \( N_c \in 0, 1 \) and using the Poisson statistics of the satellites we find:

\[
\langle N(N-1)(N-2)...(N-(n-1)) \rangle = n\langle N_c \rangle \langle N_s \rangle^{n-1} + \langle N_s \rangle^n
\]

(2.116)
Chapter 3

Neutrinos and the bispectrum

This work was conducted in collaboration with Dr Jia Liu and Dr Mathew S. Madhavacheril under the supervision of Professor David Spergel. I wrote the code to compute the bispectrum. I then used this code to obtain our constraints. It is currently being prepared for submission to Physical Review D.

3.1 Introduction

In the past few years there has been increased interest in statistics beyond the two point function. The convergence field is highly non-Gaussian and thus there is significant amounts of information beyond two point statistics. Combining two point statistics with higher points statistics, such as Minkowski functionals (Petri et al., 2013, 2015), clipping transforms (Giblin et al., 2018; Simpson et al., 2016), peak counts (Liu et al., 2015; Peel et al., 2017; Kacprzak et al., 2016) and bispectrum measurements (Kayo et al., 2013; Takada & Jain, 2004), has been shown to tighten cosmological parameters by up to 50%. In this work we will focus on the bispectrum, which is harmonic equivalent of the three point function and, if the convergence field were purely Gaussian then the bispectrum should vanish.
Since the first three-point shear measurements were made (Bernardeau, F. et al., 2002; Jarvis et al., 2004), there has been significant work to understand how systematics impact bispectrum measurements such that now unbiased measurements of cosmological parameters can be performed with the bispectrum (Semboloni et al., 2011b; Fu et al., 2014). Complementary to the shear bispectrum, there has been work on the CMB lensing bispectrum (Pratten & Lewis, 2016; Böhm et al., 2016), including work by Namikawa (2016) that found combining CMB lensing bispectrum measurements with power-spectrum measurements would provide a $\sim 30\%$ improvement on constraints on dark energy parameters $w$ and the sum of the masses of the neutrinos. Motivated by this we explore how shear bispectrum measurements are affected by massive neutrinos and whether constraints on the sum of the masses of the neutrinos can be improved with bispectrum measurements.

### 3.2 Methodology

#### 3.2.1 Simulations

The Cosmological Massive Neutrino Simulations (MassiveNuS) Liu et al. (2018) consist of a large suite of 101 N-body simulations with three varying parameters $\Sigma m_\nu$, $A_s$, and $\Omega_m$. The simulations use the public code Gadget-2 (Springel 2005), with a box size of $512 \, h^{-1}\text{Mpc}$ and $1024^3$ CDM particles, accurately capturing structure growth at $k < 10 \, h \, \text{Mpc}^{-1}$. MassiveNuS adopts a fast linear response algorithm (Ali-Ha¨ımoud & Bird, 2013; Bird et al., 2018), where neutrinos are treated using linear perturbation theory and their clustering is sourced by the full non-linear matter density. This method avoids the shot noise and high computational costs that are

---

1. The MassiveNuS data products, including galaxy and CMB lensing convergence maps, N-body snapshots, halo catalogues, and merger trees, are publicly available at [http://ColumbiaLensing.org](http://ColumbiaLensing.org).

2. The neutrino patch kspace-neutrinos is publicly available at [https://github.com/sbird/kspace-neutrinos](https://github.com/sbird/kspace-neutrinos).
usually associated with particle neutrino simulations. The code has been tested robustly and agreements with particle neutrino simulations are found to be within 0.2% for $\Sigma m_\nu < 0.6$ eV.

Galaxy and CMB lensing convergence maps are generated with the ray-tracing code LensTools\(^3\) [Petri, 2016]. The N-body snapshots are first cut into 4 planes, each with comoving thickness $126 \, h^{-1}\text{Mpc}$, regularly spaced light rays from the center of the $z = 0$ plane are then shot backwards in redshift, spreading over a $3.5^2$ deg\(^2\) solid angle, and their trajectories are tracked until the source planes at $z=0.5$, 1.0, 1.5, 2.0, 2.5 (five galaxy source planes) and $z=1100$ (CMB source plane). This ray-tracing calculation does not assume the Born approximation and thus automatically includes the post-Born terms Dodelson & Zhang (2005). In total, through randomly rotating and shifting lens planes, 10,000 convergence map realisations are generated per cosmological model per redshift, each with map size $3.5^2$ deg\(^2\) and 512\(^2\) pixels. For each realisation, the maps at different source redshifts are ray-traced through the same large scale structure and hence are properly correlated.

To model the covariance matrices, we also generate an additional set of simulations at the fiducial model ($\Sigma m_\nu = 0$, $A_s = 2.1 \times 10^{-9}$, and $\Omega_m = 0.3$), with different initial conditions. This is necessary to avoid the correlation between the model and the covariance noises during likelihood estimation, which can artificially underestimate the error size.

### 3.2.2 Convergence maps

As described in section 3.2.1, we use ray-tracing through our simulations to generate convergence maps. The convergence map is a weighted projected of the matter density

\(^3\)https://pypi.python.org/pypi/lensertools/
<table>
<thead>
<tr>
<th>Redshift</th>
<th>0.5</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{n}_{\text{gal}} )</td>
<td>8.83</td>
<td>13.25</td>
<td>11.15</td>
<td>7.36</td>
<td>4.26</td>
</tr>
</tbody>
</table>

Table 3.1: The projected source counts per arcmin\(^2\) used in this work.

\[
\kappa^i(\theta) = \int dz W(z) \delta(\chi(z)\theta, z) \tag{3.1}
\]

where \(\delta(\chi(z)\theta, z)\) is the total matter over density at redshift \(z\), \(\chi(z)\) is the comoving distance, \(W^i(z)\) is the lensing kernel for the sources in the \(i^{th}\) \(\in\{0.5, 1.0, 1.5, 2.0, 2.5, 1100\}\) redshift bin. In this work we assumed that all the sources are distributed delta function planes in redshift, thus the lensing kernels are

\[
W^i(z) = \frac{3}{2} \Omega_m H_0^2 \frac{(1+z)\chi(z)}{H(z)c} \frac{\chi z_i - \chi z}{\chi z_i} \tag{3.2}
\]

where \(\Omega_m\) is the present day fractional matter density, \(c\) is the speed of light, \(H(z)\) is the Hubble’s constant and \(z_i\) is the redshift of the source plane. Note that eq. 3.1 assumes the Born-approximation but this is not assumed in our simulations, as is described in section 3.2.1.

As our simulated maps are small, 3.5\(^2\) deg\(^2\), we will work in the flat-sky approximation. Thus we can decompose the convergence maps into Fourier coefficients,

\[
\kappa^i(\theta) = \int d^2 \ell \kappa^i_\ell e^{i \theta \cdot \ell}. \tag{3.3}
\]

Our simulations give us the true \(\kappa\) field, however when perform observations we only have access to noisy reconstructions. For the galaxy-shear convergence maps, the noise arises due to the discrete sampling of galaxies and the intrinsic shape noise of the observed galaxies. In this work we simulate noise levels appropriate for the
LSST experiment by adding Gaussian noise to our maps with variance $\sigma_s^2/\bar{n}_{\text{gal}}^i$, where $\sigma_s = 0.3$ is the intrinsic shape noise and $\bar{n}_{\text{gal}}^i$ is the projected number of sources per arcmin$^2$. The source counts used in this work are shown in table 3.1 and were chosen to be consistent with the projected levels for LSST (LSST Science Collaboration et al., 2009b).

3.2.3 Binned Power-spectrum Estimator

In the flat-sky regime, the power-spectrum $C_{\ell}^{X_1,X_2}$ is defined as

$$\langle \kappa_{\ell_1}^{X_1} \kappa_{\ell_2}^{X_2} \rangle = (2\pi)^2 \delta^{(2)}(\ell_1 + \ell_2) C_{\ell_1}^{X_1,X_2,X_3}. \quad (3.4)$$

We estimate the binned power-spectrum as

$$\hat{C}_i^{X_1,X_2} = \frac{1}{N_i} \int_{\ell_i}^{\ell_{i+1}} d^2 \ell \kappa_{\ell_1}^{X_1} \kappa_{\ell_2}^{X_2} \hat{\kappa}_{\ell_1} \hat{\kappa}_{\ell_2} \quad (3.5)$$

where $\hat{C}_i^{X_1,X_2}$ is the estimated binned power-spectrum, $\ell_i$ ($\ell_{i+1}$) is the lower (upper) boundary of the $i^{th}$ bin, and $N_i$ is the number of modes in each bin. We have pixelated convergence maps and so we discretize the above integrals and replace them with sums. For the power-spectrum analysis we used bins of width, $\Delta \ell$, of 100, a minimum $\ell$, $\ell_{\text{min}}$, of 150 and a maximum $\ell$, $\ell_{\text{max}}$, of 3150.

3.2.4 Binned Bispectrum Estimator

The calculation of the full bispectrum for modern survey is computational prohibitive and instead compression methods are used. In this work we use a binned bispectrum estimator based on the work of Bucher et al. (2016) but adapted to the flat-sky regime. Recall that, under the conditions of homogeneity and isotropy, the bispectrum can
be written as (Spergel & Goldberg 1999b; Hu 2000b)

\[
\langle \kappa_{\ell_1} \kappa_{\ell_2} \kappa_{\ell_3} \rangle = (2\pi)^2 \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) b_{\ell_1,\ell_2,\ell_3}^{X_1,X_2,X_3} \tag{3.6}
\]

where \( b_{\ell_1,\ell_2,\ell_3}^{X_1,X_2,X_3} \) is the reduced bispectrum, which only depends on the magnitude of the \( \ell \) modes. The binned bispectrum is then defined as

\[
b_{i,j,k}^{X_1,X_2,X_3} = \frac{1}{N_{i,j,k}} \int_{\ell_i}^{\ell_{i+1}} \frac{d^2 \ell_1}{(2\pi)^2} \int_{\ell_j}^{\ell_{j+1}} \frac{d^2 \ell_2}{(2\pi)^2} \int_{\ell_k}^{\ell_{k+1}} \frac{d^2 \ell_3}{(2\pi)^2} (2\pi)^2 \delta^{(2)}(\ell_1 + \ell_2 + \ell_3) \kappa_{\ell_1} \kappa_{\ell_2} \kappa_{\ell_3} \tag{3.7}
\]

where \( N_{i,j,k} \) is number of triplets in each bin, \( \ell_i \) (\( \ell_{i+1} \)) is the lower (upper) boundary of the \( i \)th bin and \( b_{i,j,k}^{X_1,X_2,X_3} \) is the estimated binned bispectrum. The binned bispectrum can be efficiently implemented by first creating a set of filtered maps

\[
W_i^X(n) = \int_{\ell_i}^{\ell_{i+1}} \frac{d^2 \ell}{(2\pi)^2} \kappa_{\ell}^X e^{i \ell \cdot n} f / \tag{3.8}
\]

Then the binned bispectrum estimator is reduced to a simple sum over products of filter maps

\[
b_{i,j,k}^{X_1,X_2,X_3} = \frac{1}{N_{i,j,k}} \int d^2 n W_i^{X_1}(n) W_j^{X_2}(n) W_k^{X_3}(n). \tag{3.9}
\]

As in power-spectrum case for our pixelated maps we replace the integrals with sum and we implement the Fourier transforms with the FFTW3 library (Frigo & Johnson 2005). For the bispectrum analysis we used bins of width, \( \Delta \ell \), of 300, a minimum \( \ell \), \( \ell_{\text{min}} \), of 150 and a maximum \( \ell \), \( \ell_{\text{max}} \), of 3150.
Figure 3.1: The difference in power of the $z = 1.0$ shear power spectrum between two cosmologies and the fiducial cosmology ($\sum m_\nu = 0.1, \Omega_m = 0.3, A_s = 2.1 \times 10^{-9}$) for 10 splits of the 10,000 simulations (dotted lines). The solid line is the difference in power from the RBF interpolation that uses the subset of cosmologies described in section 3.2.5.

### 3.2.5 Interpolation

To generate parameter constraints we will need to evaluate the likelihoods at points in parameter space beyond the 101 cosmologies we have simulated. In order to do this we will interpolate the power-spectrum and bispectrum measurements from the simulated cosmologies. To do this we used radial basis function (RBF) interpolation with a multiquadric kernel (see Buhmann & Buhmann 2003 for more details on RBF interpolation). We used the scipy implementation of this method (Jones et al. 2001).

RBF interpolation is very useful for interpolating from multidimensional scattered data and generally gives more accurate interpolation. However, as RBF interpolation uses all the input points to interpolate at a new point, if a small number of points are very noisy then this noise can propagate to the interpolated point. We found that the mean power spectrum for 19 of the simulated cosmologies were not converged. An example of one such cosmology is shown in figure 3.1b, where we plot the difference between the power-spectrum of the $z = 1.0$ shear maps at two cosmologies and the
fiducial cosmology ($\Sigma m_\nu = .1, \Omega_m = 0.3, A_s = 2.1 \times 10^{-9}$) for 10 different splits of the 10,000 simulations. There is $O(1)$ variation between the splits at the large scales, for comparison in figure 3.1a we plot the same quantity for a converged cosmology. The cosmologies for which this is an issue is those whose power-spectrum (and bispectrum) are very similar to fiducial cosmology. At first it is surprising that we have un-converged quantities after 10,000 simulations, however this occurs as our patch is very small such that some of the low $\ell$ bins have a very small number of modes.

In order to avoid the un-converged nature of some of the cosmologies propagating to interpolated we exclude those cosmologies from our interpolation. The removal of cosmologies that are preferentially similar to the fiducial cosmology risks biasing our results. In order to verify this was not an issue we verified that our interpolation accurately reproduced the values of the power-spectrum and bispectrum at excluded cosmologies. In general we found very good agreement between the interpolated values and the values of the excluded cosmologies. An example of this is shown in 3.1b where we plot the interpolated the power-spectrum at an excluded cosmology. We see that across all scales the interpolated power-spectrum matches that of the simulations. This gives us confidence that excluding these simulations will not bias our results.
We used Gaussian likelihoods to describe the distribution of the binned power-spectrum and bispectrum. Thus

\[
\ln L \propto \frac{1}{2} \sum_{i,j,X,Y} \left( \bar{C}_i^X - \hat{C}_i^X \right) \Sigma_{X,Y, i,j}^{-1} \left( \hat{C}_i^Y - \bar{C}_i^Y \right) \quad (3.10)
\]

\[
\ln L \propto \frac{1}{2} \sum_{i,j,X,Y} \left( \bar{b}_{i,j,k}^X - \hat{b}_{i,j,k}^X \right) \Sigma_{B, X,Y, (i,j,k), (i',j',k')}^{-1} \left( \hat{b}_{i',j',k'}^X - \bar{b}_{i',j',k'}^X \right) \quad (3.11)
\]

\[
\ln L \propto \frac{1}{2} \sum_{i,j,X,Y} \left( \bar{J}_i^X - \hat{J}_i^X \right) \Sigma_{J, X,Y, i,j}^{-1} \left( \hat{J}_i^Y - \bar{J}_i^Y \right), \quad (3.12)
\]

where \( \hat{J} \) is vector of joint power-spectrum and bispectrum measurements, barred quantities denote the mean of the observable, and \( \Sigma_P, \Sigma_B \) and \( \Sigma_J \) are the covariances of the power-spectrum, bispectrum and joint measurements. The assumption of a Gaussian likelihood for the bispectrum is a strong assumption and we examined the probability distribution function (PDF) of single bins of the binned-bispectrum in order to verify that this was a reasonable approximation.

We assume the covariance matrices have no cosmology dependence and we calculate them from a separate set of simulations, as described in section 3.2.1. Assuming
a cosmology independent covariance matrix is correct if the convergence maps were Gaussian random fields \cite{Carron2013}. As the convergence maps are not Gaussian random fields, we should include cosmology dependence in the covariance matrix. It is neglected here due to the computational complexity of including it and because, given the size of our constraints, will be a small effect. We remove the bias in the inverse covariance matrix with the correction factor \cite{Hartlap2007}

\[
\Sigma^{-1}_{\text{unbiased}} = \frac{n - 1}{n - p - 2} \Sigma^{-1}_{\text{sample}}
\]  

(3.13)

where \(\Sigma^{-1}_{\text{sample}}\) is the inverse of the covariance matrix from the simulations, \(n\) is the number of simulations and \(p\) is the number of parameters.

Finally we found that when we computed the mean observables from a subset of the simulations our constraint contours we consistently smaller than when all the simulations we used. This can be seen in figure 3.2a, where we calculate power-spectrum constraints using all the simulations and using four splits of 2500 simulations. It is thought that this effect arises as there is noise in the measurements of the means. To reduce the impact of this problem, which is more significant for tightly correlated variables, we instead implement the likelihoods as

\[
\ln L \propto \frac{1}{2} \left| \sum_{a!b} \sum_{i,j,X,Y} (\hat{O}^X_i - \bar{O}^X_i) \sum_{O}^{-1} (\hat{O}^Y_i - \bar{O}^Y_i) \right|
\]

(3.14)

where \(\bar{O}^X_i\) is mean of observable, \(\hat{O}^X_i\) constructed from a split, \(a\), of the simulations. In our implementation we split the simulations into ten splits, calculate the means for each of the splits and calculate the likelihood only using cross terms. As the 'noise' in the means should be independent between the splits this method should remove the bias. This approach has two drawbacks: the first is that it requires more evaluations of the likelihood (a factor of \(n_{\text{split}}^2\) more evaluations, where \(n_{\text{split}}\) is the number of splits) and by ignoring the auto-terms you are throwing away information.
The second point means that the uncertainty on the contours is increased (roughly by a factor of $(n_{\text{split}}^2/(n_{\text{split}}^2 - n_{\text{split}}))^{1/2}$). The results of using this approach are shown in figure 3.2b. We see that there is no change in the size of the constraint contours as the number of simulations is increased.

We use flat priors that uniformly weight values within the region covered by our simulations and give zero probability to cosmologies outside. Thus

\[
P \left( \sum m_\nu \right) = \begin{cases} 
\text{const}, & \text{if } 0 \text{ eV} \leq \sum m_\nu \leq 0.62036 \text{ eV} \\
0, & \text{otherwise}
\end{cases}
\]  
\tag{3.15}

\[
P (\Omega_m) = \begin{cases} 
\text{const}, & \text{if } 0.18409 \leq \Omega_m \leq 0.41591 \\
0, & \text{otherwise}
\end{cases}
\]  
\tag{3.16}

\[
P (A_s) = \begin{cases} 
\text{const}, & \text{if } 1.2886 \leq A_s < 2.9114 \\
0, & \text{otherwise}
\end{cases}
\]  
\tag{3.17}

Finally we use emcee (Foreman-Mackey et al., 2013) to sample from the parameter posteriors and the corner package (Foreman-Mackey, 2016) to display the results.

### 3.3 The effect of neutrino mass on the power-spectrum

Before presenting the bispectrum results we briefly review the effect of neutrinos on the power-spectrum, for a detailed review of their effects see (Lesgourgues & Pastor, 2006). On the largest scales massive neutrinos behave like cold dark matter whereas on scales smaller than $k_{FS} = 0.0072(\sum m_\nu/0.1\text{eV})^{1/2}(\Omega_m/0.315)^{1/2} \text{ h Mpc}^{-1}$ they free stream. As can be seen in figure 3.3a the main effect of neutrino mass on the power-spectrum is to suppress the amplitude of the power-spectrum, by ~
Figure 3.3: The shear power spectrum for three redshift configurations is plotted for the case of massless neutrinos (solid lines) and the case with $\sum m_\nu = 0.1$ eV (dotted lines). We compare the power-spectrum from simulations to power-spectra calculated using CAMB. The bottom plots show the fractional difference between the massive and massless neutrino power-spectra. The error bars show how well these configurations can be measured with an LSST-like experiment that observes half of the sky.

6% for 0.1 eV neutrinos. The suppression shows a scale dependence, which arises as massive neutrinos contribute to the energy density but do not cluster on small scales. Massive neutrinos also effect the growth of structure and thus power-spectra at different redshifts are affected differently. This effect can be seen in figure 3.3a and is explored further in section 3.5. Finally, in figure 3.3b we plot the predictions of CAMB with HALOFIT (Takahashi et al., 2012) for the shear power-spectrum. We can see these power-spectra agree well with the results from the simulations. The two most significant differences are two: the effect of neutrinos mass on the largest scales is greater in the simulations, and secondly the simulation power-spectra show
Figure 3.4: A comparison between constraints from a Fisher forecast with CAMB and constraints from the interpolated-likelihood analysis that is built using CAMB power-spectra. The contours are the 68% and 95% confidence levels.

a slight lack of power at small scales, which arises due to the resolution limit of the simulations and is described in more detail in Liu et al. (2018).

We performed a series of cross checks in order to validate both the performance of the interpolation and to compare our results to the commonly used HALOFIT predictions (Takahashi et al., 2012).

The first test compares the consistency of Fisher constraints with constraints from our interpolated-likelihood. We calculate the Fisher contours using CAMB (Lewis et al., 2000). For a fair a comparison we then use CAMB to generate power-spectra at the cosmologies of our simulations. We then interpolate these CAMB power-spectra in an identical manner to the simulations. In figure 3.4 we plot the constraint contours from the CAMB Fisher forecast and from the CAMB interpolated-likelihood. In general we see good agreement between the two contours. The distribution are not identical; whilst the upper limits seem consistent the lower limits are quite dif-
Figure 3.5: A comparison between constraints from CAMB and our simulations. The contours are the 68% and 95% confidence levels.

The second test compares the constraints from CAMB with the constraints from our simulations, and the results is show in figure 3.5. It can be seen that the CAMB constraints are significantly tighter than the simulation constraints. This is thought to be mostly driven by the drop off in power seen in the simulations at small scales. This drop off can be seen in the $z = 0.5$ power-spectra in figures 3.3a and 3.3b. This effect is explored further in (Liu et al., 2018), arises due to to the resolution of the simulations. In figures 3.3a and 3.3b it can be seen that the fractional change in power-spectra with neutrino mass is roughly the same between the simulations and CAMB. This means that, due to the small-scale drop-off in power seen in the simulations, the absolute change in CAMB power-spectra at small scales is significantly larger and thus, for the same covariance matrix, constraints from CAMB will be tighter. As
Figure 3.6: In figures 3.1b and 3.6a the squeezed and equilateral slices for two cosmologies, \( \sum m_\nu = 0.1, \Omega_m = 0.3, A_s = 2.1 \times 10^{-9} \) and \( \sum m_\nu = 0.0, \Omega_m = 0.3, A_s = 2.1 \times 10^{-9} \) is plotted. We consider the bispectrum at four different redshift configurations, the \( z = 0.5 \) auto-correlation bispectrum, \( z = 1.5 \) auto-correlation bispectrum, the \( z = 2.5 \) auto-correlation bispectrum and bispectrum from the cross correlation between \( z = 0.5 \), \( z = 1.5 \) and \( z = 2.5 \). The bottom panels show the fractional difference between the two cases The error bars show how well these configurations can be measured with an LSST-like experiment that observes half of the sky.

the covariance matrix is generated from the simulations, and thus has the drop off in power, the CAMB constraints will be tighter than they should be.

3.4 The effect of neutrino mass on the bispectrum

In figures 3.6a and 3.6b we show the effect of neutrino mass on the equilateral and squeezed slices of the shear bispectrum. The most significant effect of neutrino mass on the shear-bispectrum is a suppression of the amplitude. We find that, compared to the zero neutrino mass case, the amplitude of most configurations is reduced by
10% for $\sum m_\nu = 0.1$eV. [Levi & Vlah (2016)] computed the effect of neutrino mass on the matter bispectrum to second order in perturbation theory. They found on small scales the equilateral bispectrum was suppressed by $-13.5\Omega_\nu/\Omega_m$, which for $\sum m_\nu = 0.1$eV is a suppression of 9.9% and is consistent with the suppression seen in figure 3.6b. [Ruggeri et al. (2018)] examined the effect of neutrino mass on the matter bispectrum with simulations and found, for $\sum m_\nu = 0.17$eV, the amplitude of the matter bispectrum was reduced by 17% on the smallest scales, which is consistent with our results. Thus the amplitude of the bispectrum is almost twice as sensitive to neutrino mass than the power-spectrum, which is reduced by $\sim 6\%$ as $\sum m_\nu$ is increased from 0eV to 0.1 eV. However, whilst the bispectrum is more sensitive to neutrino mass the bispectrum is harder to measure (for LSST the signal to noise of the bispectrum is a factor of $\sim 5$ less than the power-spectrum). This issue can be partly seen in figure 3.6b where the error bars are the expected error bars if 50% of the sky is observed with the galaxy number counts given in table 3.1.

We find that, for both the squeezed and equilateral configuration, the effect of neutrinos is greatest on the largest scales, but we find only a very weak dependence on configuration. However, it should be noted that our ability to study the configuration dependence of the bispectrum is limited due to the size of simulated maps. The small patch size means that we do not have access to large scale modes $\ell < 150$ and necessitates the use of large bispectrum bins, $\Delta \ell = 300$. The lack of large scale modes means that we are unable to probe very squeezed configurations whilst the large bins means that we are potentially washing out structure in the bispectrum. In future work we would like to investigate the importance of these issues.

We find that the neutrino mass effects the bispectrum differently at different redshifts. This is seen in figures 3.6a and 3.6b where the squeezed and equilateral slices of the bispectrum are plotted for four different redshifts. We find that for high redshift sources the main effect of neutrino mass is to just reduce the bispectrum amplitude.
Figure 3.7: 68% parameter constraints from the power-spectrum, bispectrum and joint analysis.

Whereas for lower redshift sources we find a weak scale dependence. The importance of this redshift information is discussed in 3.5.

### 3.5 Constraints

In figure [3.7] we present the constraints obtained from the bispectrum. We find a strong degeneracies between $\sum m_\nu$ and $A_s$, and between $\sum m_\nu$ and $\Omega_m$. This degeneracy arises as the shape of the bispectrum depends only weakly on the cosmological parameters. Instead the main effect of neutrino mass on the shear bispectrum is to reduce the amplitude and the main effect of increasing either the matter density or
the amplitude of fluctuations is to increase the amplitude, thus leading to a degeneracy between these parameters. If $\sum m_\nu = 0.1$ eV, then LSST-shear bispectrum would provide a weak constraint on the sum of the neutrino mass $\sum m_\nu = 0.1^{+0.30}_{-0.01}$. The constraints from the bispectrum are weaker than those from the power-spectrum, for which the corresponding constraint is $\sum m_\nu = 0.1^{+0.15}_{-0.05}$. We find that power-spectrum provides tighter constraints across all three of the cosmological parameters, and this can be seen in figure 3.7.

Despite the similar degeneracies, we find that combining the power-spectrum and bispectrum measurements significantly improves the ability to constrain the parameters, as can be seen in figure 3.7. We find that the joint measurements improves the constraint on $\sum m_\nu$ by 30%, to $\sum m_\nu = 0.1^{+0.12}_{-0.01}$, on $\Omega_m$ by 30%, to $\Omega_m = 0.3^{+0.0037}_{-0.00035}$ and the constraint on $A_s$ by 40%, to $A_s \times 10^9 = 2.1^{+0.10}_{-0.02}$.

In figures 3.8a and 3.8b we review the effect of tomography on power-spectrum constraints and discuss the importance of tomography in bispectrum measurements. In the case without tomography we use the approximation that all of the galaxies lie with the redshift 1.0 bin, thus it has $\bar{n}_{\text{gal}} = 44.8$ gal per arcmin$^2$. Tomography greatly enhances the ability to constrain neutrino mass with the power-spectrum. We see a similar result in figure 3.8a. The growth information from tomography helps break the $A_s$ and $\sum m_\nu$ degeneracy and so significantly improves the power-spectrum constraint. For the bispectrum, we find that using tomography provides a significant improvement in the ability to constrain $A_s$ and a slight improvement in the ability to constrain $\sum m_\nu$. The bispectrum is sensitive to the growth history so one would expect greater improvement with tomography than we see. However, as is seen in figures 3.6a and 3.6b the very weak shape dependence induced by neutrino mass, combined with noise levels that are comparable to the signal, means that the degeneracy with $A_s$ cannot be broken.
Figure 3.8: A comparison between power-spectrum and bispectrum constraints on $\Omega_M$ and $\sum m_\nu$ from a single redshift measurement at $z = 1.0$ against tomographic measurements from five redshift bins, $z \in \{0, 0.5, 1, 1.5, 2, 2.5\}$. The total number of galaxies per arcmin$^2$ is the same for the two cases.

3.6 Conclusions

We explored the impact of neutrino mass on the weak lensing bispectrum. We find that the shape of the bispectrum is largely insensitive to the sum of the neutrino masses and the dominant effect of massive neutrinos is to suppress the amplitude of the bispectrum. As this effect is highly degenerate with $A_s$, and as the signal to noise of the bispectrum is significantly less than the power-spectrum, the bispectrum alone does not provide better constraints on neutrino mass than the power-spectrum. However, bispectrum and power-spectrum measurements are highly complementary and analyzing them joint produces 30% tighter parameter constraints. Whilst the results here are focused on a LSST like experiment bispectrum measurements could, and have been, be used to improve current parameter constraints [Fu et al., 2014]. It will be the subject of future work to investigate how much information the bispectrum adds for different noise levels and thus we cannot currently say what precisely what level of improvement would be obtained from adding bispectrum measurements to constraints from current surveys like the Dark Energy Survey. With that said, the naive scaling of the bispectrum S/N (when compared to the power spectrum) suggests that bispectrum measurements are increasingly beneficial as the noise is decreased.
As only an amplitude reduction is the main effect of neutrino mass on the bispectrum, it is possible that the information of the bispectrum could be encoded simply through the skewness, which is easier to compute and analyse. However, previous work (Levi & Vlah, 2016) found that in perturbation theory the squeezed matter bispectrum was very sensitive to neutrino mass. In this work, we could not probe the largest scales and so were not sensitive to very squeezed configurations. If highly squeezed convergence bispectra are very sensitive to neutrino mass then the bispectrum results could not be completely compressed to the skew.

The analysis presented here is a overly optimistic for several reasons. Firstly our use of delta functions for the galaxies is highly idealized as the sources are expected to be spread over broad redshift kernels (LSST Science Collaboration et al., 2009b; Chang et al., 2013). Secondly we do not consider the impact of systematics such as source-lens clustering (Bernardeau, 1997; Hamana et al., 2001), intrinsic alignments (Hirata & Seljak, 2003; Crittenden et al., 2001), baryonic effects (Semboloni et al., 2011a) and multiplicative biases (Massey et al., 2013; Huterer et al., 2006; Schaan et al., 2017). Finally we have only considered the effect of three cosmological parameters. It is known that, for the power spectrum, curvature, the Hubble parameter and altered dark energy histories are degenerate with the sum of the masses of the neutrinos (Font-Ribera et al., 2014; Benoît-Lévy et al., 2012; Hamann et al., 2012; Mishra-Sharma et al., 2018). As a rigorous assessment of these effects is beyond the scope of this work, we stress that our constraints are not realistic projections for LSST. Rather, as these effects are neglected for both the power-spectrum and bispectrum, it is a demonstration of the potential information gain from including bispectrum measurements.
Chapter 4

Constraining Primordial Tensor Non-Gaussianity with ACTPol and Planck

This chapter was done under the supervision of Professor David Spergel and has benefited from discussions with Dr Alex van Engelen, Dr Joel Meyers and Dr Daan Meerburg. The results of this chapter have been presented at the ACT (2018) collaboration meeting.

4.1 Introduction

The cosmic microwave background (CMB) has been a crucial tool for studying the primordial universe. Precise measurements of the CMB temperature and polarization power spectra from space (e.g. PLANCK, WMAP), the ground (e.g. SPT, ACT, BICEP, POLARBEAR) and balloons (e.g. BOOMERanG) have helped study the primordial power spectrum and constrain the ratio of primordial tensor to scalar power.
Simultaneously we have developed the analytic and analysis tools to probe the higher order statistics of the primordial fluctuations (Maldacena, 2003); through which we have developed new ways to study the early universe physics (Komatsu et al., 2005). Many primordial physical processes can create non-Gaussianity and signatures of these effects would be informative for characterizing the physics of these epochs (see Liguori et al. 2010, Yadav & Wandelt 2010, Chen 2010 for a review of these effects).

There have been already been many investigations into primordial non-Gaussianity (e.g. Senatore et al. 2010b, Creminelli et al. 2006, Komatsu et al. 2003). The Planck satellite has provided the strongest constraints on primordial non-Gaussianity (from both the bispectrum and trispectrum) (Planck Collaboration XVII, 2016). Primordial bispectra sourced by purely scalars are the most common theoretical models and has been the focus of recent analyses, with the Planck satellite constraining a vast space of models. Several of the Planck constraints that are close to the cosmic variance limit (CVL) for CMB only measurements (Planck Collaboration XVII, 2016).

Recent work has highlighted the potential interest in bispectrum from interactions involving scalar and tensors, such as the tensor-scalar-scalar bispectrum discussed in Meerburg et al. (2016). In addition tensor-tensor-tensor non-Gaussianity has been proposed as a way probing high energy gravity theories (Maldacena & Pimentel, 2011, Soda et al., 2011, McFadden & Skenderis, 2011), primordial magnetic fields (Brown & Crittenden, 2005, Shiraishi et al., 2012) and as an observable of a class of inflationary models (Agrawal et al., 2017, Barnaby et al., 2012, Cook & Sorbo, 2013, Ferreira & Sloth, 2014). Scalar non-Gaussianity can only produce parity even bispectra, those with $\ell_1 + \ell_2 + \ell_3 = \text{even}$, however non-Gaussianity involving tensor fields naturally produce parity even and parity odd bispectra, those with $\ell_1 + \ell_2 + \ell_3 = \text{odd}$ (Shiraishi et al., 2011, Kamionkowski & Souradeep, 2011). Finally the different transfer functions of tensor non-Gaussianity means that these templates
have significantly different shapes from the scalar non-Gaussianity and so can be
constrained almost independently (Shiraishi et al., 2013).

So far there has been only a few studies to constrain these types of non-Gaussianity. The amplitude of parity violating equilateral non-Gaussianity has been constrained with WMAP and Planck data (Shiraishi et al., 2015; Planck Collaboration XVII, 2016) and recently (Shiraishi et al., 2018) provided the first constraint on the tensor-scalar-scalar bispectrum. If discovered such types of non-Gaussianity would provide strong constraints on theoretical models and provide an exciting new handle to constrain primordial physics.

In this work we combine $\sim 550$ deg$^2$ from the ACTPol experiment (Louis et al., 2017) with the overlapping data from the Planck satellite to derive constraints on bispectrum from tensor-scalar interactions. We will use the KSW (Komatsu et al., 2005) estimator to a set of theoretical templates for tensor-scalar non-Gaussianitiy. This has been used in many previous analyses (e.g. Spergel et al., 2007; Planck Collaboration XXIV, 2014). We constrain templates motivated by the familiar local, equilateral and orthogonal shapes (see e.g Komatsu, 2010 for an overview of these templates).

In section 4.2 we outline the KSW estimator that we use and then in section 4.3 we describe the theoretical shapes of our templates. In section 4.4 we describe the data used in this work and briefly outline the analysis pipeline before presenting our results and conclusions in sections 4.5 and 4.6. In Appendix 4.7 we describe methods to simulated separable primordial non-Gaussian templates and use these methods to test our analysis pipeline.
4.2 Bispectrum estimator and Analysis Pipeline

4.2.1 Flat-sky bispectrum estimator

Current generation CMB experiments measure both the intensity and linear polarization of CMB. Typically we parameterize measurements of the electric field perpendicular to the observation direction \( \vec{n} \) in the following manner

\[
\langle \vec{E}_i \vec{E}_j \rangle \propto \frac{1}{2} (\sigma_1 U + \sigma_3 Q + II)_{i,j},
\]

where \( E_i \) is the component of the electric field in direction \( \vec{e}_i \) (where \( \vec{e}_1 \) and \( \vec{e}_2 \) form an orthonormal basis); \( Q, U \) and \( I \) and the stokes parameters, \( \sigma_1, \sigma_3 \) are the first and third Pauli matrices and \( I \) is the 2D identity matrix. Stokes I measures the total intensity and is related to the temperature fluctuations in the CMB. Stokes Q measures the difference in power in direction \( \vec{e}_1 \) compared to \( \vec{e}_2 \) and Stokes U measure the difference in power between the \( \vec{e}_1 \pm \vec{e}_2 \) directions. Here we have ignored the possibility of Stokes V, which essentially measures the contribution from circularly polarized light, as Compton scattering does not generate circularly polarised light. In this work we use will be focusing on small regions of sky and at high \( \ell \) and so we are able to make the flat sky approximation. In this regime the temperature anisotropies (Stokes I) are decomposed as follows

\[
\frac{\Delta T(\vec{n})}{T} = \int \frac{d\ell^2}{4\pi^2} a^T(\vec{\ell}) e^{i\vec{n} \cdot \vec{\ell}}. \quad (4.2)
\]

and the polarization can be decomposed as

\[
a^E(\vec{\ell}) + ia^B(\vec{\ell}) = \int \frac{d\ell^2}{(2\pi)^2} (Q + i U)e^{-2i\phi_\ell}e^{-\vec{\ell} \vec{n}} \quad (4.3)
\]
\[
a^E(\vec{\ell}) - ia^B(\vec{\ell}) = \int \frac{d\ell^2}{(2\pi)^2} (Q - i U)e^{2i\phi_\ell}e^{-\vec{\ell} \vec{n}}. \quad (4.4)
\]
For notational clarity from here we will write $a(\vec{\ell})$ as $a_{\vec{\ell}}$. The flat sky approximation is accurate to $< 1\%$ for scales greater than $\ell \sim 200$. The bispectrum is equal to the ensemble average of the three $a_{\vec{\ell}}^X$ so

$$B^{(X_1,X_2,X_3)}(\vec{\ell}_1, \vec{\ell}_2, \vec{\ell}_3) = \langle a_{\vec{\ell}_1}^{X_1} a_{\vec{\ell}_2}^{X_2} a_{\vec{\ell}_3}^{X_3} \rangle$$  \hspace{1cm} (4.5)$$

As shown by Hu (2000a); Babich et al. (2004) (and we briefly review this in section 4.3.2) the primordial bispectrum, in the flat sky, can be written as

$$B^{X_1,X_2,X_3}(\vec{\ell}_1, \vec{\ell}_2, \vec{\ell}_3) = 4\pi^2 \delta^2(\vec{\ell}_1 + \vec{\ell}_2 + \vec{\ell}_3) b^{X_1,X_2,X_3}(\vec{\ell}_1, \vec{\ell}_2, \vec{\ell}_3)$$  \hspace{1cm} (4.6)$$

where $b^{X_1,X_2,X_3}(\vec{\ell}_1, \vec{\ell}_2, \vec{\ell}_3)$ is the reduced bispectrum.

The minimum variance estimator, one that saturates the Cramér-Rao bound, for the amplitude $\hat{f}_{NL}$ of non-Gaussianity with reduced bispectrum $b^{X_1,X_2,X_3}(\vec{\ell}_1, \vec{\ell}_2, \vec{\ell}_3)$ can be written as (Babich et al., 2004; Creminelli et al. 2006; Senatore et al., 2010a; Verde et al., 2013)

$$\hat{f}_{NL} = \frac{1}{N} \sum_{X_i} \sum_{X_i'} \left[ \prod_i \int \int d^2 \vec{l}_i \int \int d^2 \vec{l}'_i \delta^2(\vec{l}_1' + \vec{l}_2' + \vec{l}_3') \times \right.$$

$$\left. \left( \left( C_{\vec{l}_1}^{-1} \right)^{X_i} a_{\vec{\ell}_1}^{X_i} \left( C_{\vec{l}_2}^{-1} \right) a_{\vec{\ell}_2}^{X_i} \left( C_{\vec{l}_3}^{-1} \right)^{X_i} a_{\vec{\ell}_3}^{X_i} - 3 \left( C_{\vec{l}_1}^{-1} \right)^{X_i} a_{\vec{\ell}_2}^{X_i} \left( C_{\vec{l}_3}^{-1} \right)^{X_i} a_{\vec{\ell}_3}^{X_i} \right) + \ldots \mathrm{cyclic} \right) \right],$$

$$\hspace{1cm} (4.7)$$

where is a suitable normalization chosen to produce unit response to the theoretical bispectrum. The last term of eq. 4.7 is known as the linear term. As was discussed in Creminelli et al. (2006) and Bucher et al. (2016), this term is important for squeezed configurations. In Coulton et al. (2017) we found that if the lowest $\ell$ modes are masked then linear terms will be reduced to a negligible level. In this work we will also mask the lowest $\ell$ modes as is discussed in section 4.4 and so we can ignore the contributions of the linear term. We will exploit the approximate separability of the
primordial bispectra to evaluate the estimator and its normalization as described in \cite{Komatsu2005, Smith2011}. The variance of the estimator is inversely proportional to the normalization.

\[ \text{Var}(\hat{A}) \propto \frac{1}{N}, \quad (4.8) \]

where the proportionality constant depends on the sky coverage. More details of our implementation of this estimator are given in \cite{Coulton2017}.

### 4.3 Bispectrum Templates

In this work we consider bispectrum templates arising from primordial scalar only and primordial tensor-scalar-scalar non-Gaussianity. We also consider the possible contamination from Lensing x secondary sources. Dust and synchrotron radiation can potentially also contaminate these measurements but we leave an investigation of their impact to future work.

#### 4.3.1 Primordial Templates in the flat sky regime

We briefly review the approximations derived in the appendix of \cite{Babich2004} and used in \cite{Meerburg2016} to calculate bispectrum in the flat sky approximation.

The temperature and E modes fluctuations \(a^X(\ell) \quad (X \in \{T, E\})\) are related to the primordial scalar fluctuations, \(\psi(k)\), by \cite{Shiraishi2013}

\[ a^T_\psi(\ell) = w^T_\ell \int_0^{\tau_0} d\tau \int \frac{d^3k}{(2\pi)^3} \psi(k)e^{-ikz}D S^T_\psi(k, \tau)(2\pi)^2 \delta^2(k\parallel D - \ell) \quad (4.9) \]

\[ a^E_\psi(\ell) = w^E_\ell \int_0^{\tau_0} d\tau \int \frac{d^3k}{(2\pi)^3} \psi(k)e^{-ikz}D \frac{\ell^2}{(Dk^z)^2 + \ell^2} S^E_\psi(k, \tau)(2\pi)^2 \delta^2(k\parallel D - \ell) \quad (4.10) \]
and temperature, E and B mode fluctuations sourced by primordial tensor fluctuations, \( h^\pm(k) \), are given by

\[
a^T_h(\ell) = w^T_T \int_0^{\tau_0} d\tau \int \frac{d^3k}{(2\pi)^3} \sum_{\pm} h^\pm(k) e^{-ikzD} \left\{ \frac{\ell^2}{(Dk^z)^2 + \ell^2} S^T_h(k, \tau)(2\pi)^2 \delta^2(k^\parallel D - \ell) \right. \\
\left. + \frac{\ell^2}{(Dk^z)^2 + \ell^2} S^P_h(k, \tau)(2\pi)^2 \delta^2(k^\parallel D - \ell) \right\} 
\]  
(4.11)

\[
a^E_h(\ell) = w^E_E \int_0^{\tau_0} d\tau \int \frac{d^3k}{(2\pi)^3} \sum_{\pm} h^\pm(k) e^{-ikzD} \left\{ 2 - \frac{\ell^2}{(Dk^z)^2 + \ell^2} \right\} \frac{\ell^2}{(Dk^z)^2 + \ell^2} S^P_h(k, \tau)(2\pi)^2 \delta^2(k^\parallel D - \ell) 
\]  
(4.12)

\[
a^B_h(\ell) = w^B_B \int_0^{\tau_0} d\tau \int \frac{d^3k}{(2\pi)^3} \sum_{\pm} h^\pm(k) e^{-ikzD} \left\{ 2 - \frac{k^z D}{\sqrt{(Dk^z)^2 + \ell^2}} \right\} \frac{k^z D}{\sqrt{(Dk^z)^2 + \ell^2}} S^P_h(k, \tau)(2\pi)^2 \delta^2(k^\parallel D - \ell) 
\]  
(4.13)

where \( S^a_b(k, \tau) \) (\( a \in \{T, P\} \) and \( b \in \{\psi, h\} \)) are the temperature and polarization source functions for the scalars and tensors, \( \tau \) is the conformal time, \( \tau_0 \) is the conformal time today, \( D = \tau_0 - \tau \), \( k^z \) is the line of sight component and \( k^\parallel \) is the component in the plane of the sky. With these components we can compute the bispectrum, demonstrated here for the temperature only and purely scalar bispectrum

\[
\langle a(\ell_1) a(\ell_2) a(\ell_3) \rangle = w^T_{\ell_1} w^T_{\ell_2} w^T_{\ell_3} \times \frac{3}{3} \left[ \int d\tau_n \frac{d^3k_n}{(2\pi)^3} e^{i k_n D} (2\pi)^2 \delta(D_n k^\parallel - \ell) \right] S^T_{\psi}(k_1, \tau) S^T_{\psi}(k_2, \tau) S^T_{\psi}(k_3, \tau) \langle \psi(k_1) \psi(k_2) \psi(k_3) \rangle. 
\]  
(4.14)

Under the assumptions of statistical isotropy and translational invariance the primordial three point function can be expressed as

\[
\langle \psi(k_1) \psi(k_2) \psi(k_3) \rangle = B_{\psi\psi\psi}(k_1, k_2, k_3) = (2\pi)^3 \delta^3(k_1 + k_2 + k_3) B_{\psi\psi\psi}(k_1, k_2, k_3). 
\]  
(4.15)
Now we assume that the bispectrum does not vary significantly over the surface of last scattering; this is true for the smooth bispectrum considered here, but may not be true for bispectra with features \cite{Babich:2004}. With this assumption we can approximate $k_{\parallel}$ with $\ell/(\tau_0 - \tau_R)$, where $\tau_R$ is the conformal time at recombination, and take the primordial bispectrum out of the time integrals giving

$$\langle a(\ell_1)a(\ell_2)a(\ell_3) \rangle = w_{\ell_1}^T w_{\ell_2}^T w_{\ell_3}^T (2\pi)^2 (D_R)^2 \delta^2 (\ell_1 + \ell_2 + \ell_3)$$

$$\prod_{n=1}^3 \left[ \int \frac{dk^2_n}{(2\pi)} e^{ik^2_n D} \delta^2 (k_1^2 + k_2^2 + k_3^2) \right]$$

where we defined $k_i^R = \sqrt{(k_i^2)^2 + (\ell/D_R)^2}$ with $D_R = \tau_0 - \tau_R$ and the flat sky transfer function $\Delta^T_\phi$ is defined as

$$\Delta^T_\phi = \int_0^{\tau_0} \frac{d\tau}{D^2} \delta^2 (k_i^R, \tau) e^{-ik_i^R D} \cdot$$

(4.17)

This result is easily generalize to give the results for general bispectra giving

$$\langle a^{X_1}(\ell_1)a^{X_2}(\ell_2)a^{X_3}(\ell_3) \rangle = w_{\ell_1}^{X_1} w_{\ell_2}^{X_2} w_{\ell_3}^{X_3} (2\pi)^2 \delta^2 (\ell_1 + \ell_2 + \ell_3)$$

$$\times \int \frac{dk_1^2}{(2\pi)} \frac{dk_2^2}{(2\pi)} \frac{dk_3^2}{(2\pi)} \Delta^X_{Y_1}(k_1^2, \ell_1) \Delta^X_{Y_2}(k_2^2, \ell_2) \Delta^X_{Y_3}(k_3^2, \ell_3) B_{Y_1,Y_2,Y_3}(k_1,k_2,k_3) 2\pi \delta (k_1^2 + k_2^2 + k_3^2)$$

(4.18)
with $X_i \in \{T, E, B\}$ and $Y_i \in \{\psi, h\}$, $B_{Y_i,Y_2,Y_3}(k_1,k_2,k_3)$ is the general primordial bispectrum from tensors and scalars and we have defined the transfer functions $\Delta^X_Y$ as

$$
\Delta^E_\psi = \int_0^{\tau_0} d\tau \frac{\ell^2}{D^2} \left( S^\psi_T(k^R_i, \tau) e^{-i k_i^\bot D} \right)
$$

(4.19)

$$
\Delta^T_h = \int_0^{\tau_0} d\tau \frac{\ell^2}{D^2} \left( S^h_T(k^R_i, \tau) e^{-i k_i^\bot D} \right)
$$

(4.20)

$$
\Delta^E_h = \int_0^{\tau_0} d\tau \frac{\ell^2}{D^2} \left( 2 - \frac{\ell^2}{(k^\bot D)^2 + \ell^2} \right) S^h_T(k^R_i, \tau) e^{-i k_i^\bot D}
$$

(4.21)

$$
\Delta^B_h = \int_0^{\tau_0} d\tau \frac{k^\bot D}{D^2} \frac{2i}{\sqrt{(k^\bot D)^2 + \ell^2}} S^h_T(k^R_i, \tau) e^{-i k_i^\bot D}.
$$

(4.22)

We will use these pieces in the following sections to compute bispectra templates for primordial scalar-tensor bispectra. For the remainder of this section we define the reduced bispectrum without the beam and pixelization function as

$$
b_{E1,E2,E3}^{X1,X2,X3} = w_{E1}^{X1} w_{E2}^{X2} w_{E3}^{X3} \tilde{b}_{E1,E2,E3}^{X1,X2,X3},
$$

(4.23)

where $w_E^X$ is beam and pixel window function for map $X$.

### 4.3.2 Primordial Templates Scalar-Scalar-Scalar

Primordial scalar non-Gaussianity has been well studied and constrained in many previous works and the current best constraints are given in Planck Collaboration XVII (2016). We note that traditional analyses have been performed in terms of primordial potential fluctuations $\Phi(k)$ which are related to the field fluctuations in the matter dominated era as $\Phi = \frac{3}{5} \psi$. For consistency with previous work we will use $\Phi$ when discussing these bispectra. In this work we consider three primordial templates: local, orthogonal and equilateral.
From this previous work we have the full sky result for scalar-scalar-scalar bispectra as (Komatsu & Spergel, 2001a; Babich & Zaldarriaga, 2004; Yadav et al., 2007)

\[
\tilde{b}_{\ell_1,\ell_2,\ell_3}^{X_1,X_2,X_3} = \frac{8}{\pi^3} \int dr r^2 \prod_i \left( \int dk_i k_i^2 g_{\ell}^{X_i}(k_i) j_{\ell_i}(k_ir) \right) B_{\Phi\Phi\Phi}(k_1, k_2, k_3),
\]

where \(g_{\ell}^{X_i}(k_i)\), for \(X_i \in \{T, E\}\), is the full sky transfer function, calculated by CAMB (Lewis et al., 2000) and defined as

\[
g_{\ell}^{T,\psi}(k) = \int_0^{\tau_0} d\tau_0 \langle \psi_T(k, \tau) j_\ell(k(\tau_0 - \tau)) \rangle
\]

\[
g_{\ell}^{E,\psi}(k) = \int_0^{\tau_0} d\tau_0 \langle \psi_P(k, \tau)(1 + \partial_x^2)x^2 j_\ell(x) \rangle,
\]

where \(x = k(\tau_0 - \tau)\). As a validation of our calculation, we cross check our flat sky results against the full sky results and find good agreement over all the \(\ell\) range considered here \(\ell > 200\).

**Local Bispectrum**

Primordial local non-Gaussianity is so-called as it generated by the real space interactions of fields, and is thus local. Physically is arises when two short wavelength modes interact with a longer wavelength mode. It has been shown that this type of non-Gaussianity is small in single field inflationary theories (Maldacena, 2003; Creminelli & Zaldarriaga, 2004) however it is generic in inflationary theories with multiple fields. Thus a detection of local primordial non-Gaussianity would rule out the possibility of single field inflation. The local three point structure is given by (Falk et al., 1993; Gangui et al., 1994; Verde et al., 2000; Wang & Kamionkowski, 2000; Komatsu &
\[ B_{\Phi}^{local}(k_1, k_2, k_3) = 2 f_{NL}^{local}(P_{\Phi}(k_1)P_{\Phi}(k_2) + P_{\Phi}(k_1)P_{\Phi}(k_3) + P_{\Phi}(k_2)P_{\Phi}(k_3)) \]

\[ = 2 A^2 (2\pi)^2 f_{NL}^{local} \left( \frac{1}{k_1^{4-n_s} k_2^{4-n_s} k_3^{4-n_s}} + \frac{1}{k_1^{4-n_s} k_2^{4-n_s} k_3^{4-n_s}} + \frac{1}{k_2^{4-n_s} k_3^{4-n_s} k_3^{4-n_s}} \right), \]

(4.27)

where \( n_s \) is the primordial spectral tilt and \( A \) is the amplitude of fluctuations. This shape receives large contributions when one mode is much longer than the others, in what is known as the squeezed limit, \( k_1 << k_2 \approx k_3 \). The local bispectrum has been well studied and the current best limits come from the Planck experiment (Planck Collaboration XVII, 2016) and are \( 0.8 \pm 5.0 \).

**Equilateral Bispectrum**

The equilateral primordial shape receives large contributions from interacting modes within the horizon. Contributions to this term are maximal just before the modes leave the horizon and so this shape is peaked when all three modes are approximately equal in magnitude, \( k_1 \approx k_2 \approx k_3 \). This template is given by the following three point structure Creminelli et al. (2006)

\[ B_{\Phi}^{equil}(k_1, k_2, k_3) = 6 A^2 (2\pi)^2 f_{NL}^{local} \left( \frac{1}{k_1^{4-n_s} k_2^{4-n_s} k_3^{4-n_s}} - \frac{1}{k_1^{4-n_s} k_2^{4-n_s} k_3^{4-n_s}} + \frac{2}{(k_1 k_2 k_3)^{2(4-n_s)/3}} \right) + \text{5 permutations}, \]

(4.28)

This bispectrum shape approximates many shapes that arise from theories of inflation where there is derivative coupling, such as in \( P(X) \) theories e.g. DBI inflation Silverstein & Tong (2004). The current best limits for this are \( -4 \pm 43 \) Planck Collaboration XVII (2016).

105
Orthogonal Bispectrum

This bispectrum is orthogonal, in the sense of the bispectrum cosine Babich et al. (2004), to the equilateral shape. This shape is approximated by the point structure given by Senatore et al. (2010a); Cheung et al. (2008) as

\[
B_{\Phi}^{\text{orth}}(k_1, k_2, k_3) = 6A^2(2\pi)^2 f_{\text{local}}^{NL} \left( -\frac{3}{k_1^{(4-n_s)}} k_2^{(4-n_s)} - \frac{3}{k_1^{(4-n_s)}} k_3^{(4-n_s)} \right. \\
- \frac{3}{k_2^{(4-n_s)} k_3^{(4-n_s)}} - \left. \frac{8}{(k_1 k_2 k_3)^{2(4-n_s)/3}} + 3 \left( \frac{1}{k_1^{(4-n_s)/3}} \frac{1}{k_2^{(4-n_s)/3}} \frac{1}{k_3^{(4-n_s)/3}} + 5 \text{ permutations} \right) \right). \\
\]

(4.29)

Such non-Gaussianity arises in effective field theories of inflation and in Galilean inflation Burrage et al. (2011). The current best limits for this are $-26 \pm 21$ Planck Collaboration XVII (2016).

4.3.3 Primordial tensor-scalar-scalar non-Gaussianity

Whilst Maldacena (2003) showed that primordial tensor-scalar-scalar non-Gaussianity is generated in single field inflation, they also showed that it would be small (proportional to the slow roll parameters). More recent work has shown that in solid inflationary models Akhshik (2015) and in Axion-Gauge field inflationary scenarios Dimastrogiovanni et al. (2018), this type of non-Gaussianity can be generated at measurable levels. Further in single field inflationary models with non-Einstein gravity, tensor-scalar-scalar non-Gaussianity can be generated Gao et al. (2013).

This section follows the derivation of Meerburg et al. (2016). In Meerburg et al. (2016) they express primordial general bispectrum as

\[
\langle \psi(k_1)\psi(k_2)h^{s_3}(k_3) \rangle = 4\pi^4 A_s^2 \sqrt{f_{\text{NL}}} \left( \phi \right) \delta(k_1 + k_2 + k_3) I(k_1, k_2, k_3) e_{s_3}^{a,b}(k_3) k_1^a k_2^b \\
\]

(4.30)
where $I(k_1, k_2, k_3)$ is the primordial shape, $r$ is the tensor to scalar ratio, $\epsilon_{ij}^x(k)$ is the transverse traceless polarization tensor, and $f_{NL}^{h\phi\psi}$ is the amplitude of the non-Gaussianity. To calculate this bispectrum we need to evaluate the contraction of the momenta with the polarization tensor. When $\hat{k} = \hat{z}$ the polarization tensor has the following form

$$
\epsilon_{ij}^z(\hat{z}) = \frac{1}{\sqrt{2}} \begin{pmatrix}
1 & \mp i & 0 \\
\mp i & 1 & 0 \\
0 & 0 & 0
\end{pmatrix}
$$

(4.31)

which can be rotated using the standard rotation matrix, $S(\hat{n})$ to any direction. For $k_i = (k_1^x, k_1^y, k_1^z) = k_i(\cos(\phi_i) \sin(\theta_i), \sin(\phi_i) \sin(\theta_i), \cos(\theta_i))$ we have [Weinberg (2008)]

$$
S_{a,b}(\hat{n}) = \begin{pmatrix}
\cos(\phi_i) \cos(\theta_i) & -\sin \phi_i & \cos(\phi_i) \sin(\theta_i) \\
\sin(\phi_i) \cos(\theta_i) & \cos \phi_i & \sin(\phi_i) \sin(\theta_i) \\
-\sin(\theta_i) & 0 & \cos(\theta_i)
\end{pmatrix} = \begin{pmatrix}
\frac{k_i^x k_1^x}{k_i} & -\frac{k_i^y}{\sqrt{k_i^x + k_i^y}} & \frac{k_1^x}{k_i} \\
\frac{k_i^y}{\sqrt{k_i^x + k_i^y}} & \frac{k_1^y}{k_i} & \frac{k_1^y}{k_i} \\
\frac{k_i^z}{\sqrt{k_i^x + k_i^z}} & -\frac{k_i^z}{\sqrt{k_i^x + k_i^z}} & \frac{k_i^z}{k_i}
\end{pmatrix}
$$

(4.32)

With these pieces we can calculate the reduced bispectrum. The different sign between the even and odd parity tensor terms leads to two distinct contributions to the reduced bispectrum. The first from the even parity only terms, so a bispectrum without any B mode legs, is

$$
\tilde{b}^{X_1, X_2, X_3}(\ell_1, \ell_2, \ell_3) = 4\pi^4 A_s^2 \sqrt{r} f_{NL}^{h\phi\psi} \times 
$$

$$
\int \int \int \frac{dk_1^2 dk_2^2 dk_3^2}{2\pi^2} I(k_1^R, k_2^R, k_3^R) \Delta_{X_1}^{X_1}(k_1^z, \ell_1) \Delta_{X_2}^{X_2}(k_2^z, \ell_2) \Delta_{X_3}^{X_3}(k_3^z, \ell_3) 
$$

$$
\times (2\pi) \delta(k_1^z + k_2^z + k_3^z) \left[ k_2^2 \ell_3^2 + k_3^2 (\ell_2)^2 - 2k_2k_3 \ell_2 \cdot \ell_3 + \frac{1}{D_R^2} (\ell_2 \times \ell_3)^2 \right] \frac{-\sqrt{2}}{k_1^R k_2^R (k_3^R)^2} + \text{cyclic}
$$

(4.33)
The second contribution arises when there is one B mode leg

\[ b^{X_1X_2B}(\ell_1, \ell_2, \ell_3) = 4\pi^4 A_N^2 \sqrt{f_{NL}} \times \]
\[ \int \int \int \frac{dk_1^z dk_2^z dk_3^z}{2\pi} I(k_1^R, k_2^R, k_3^R) \Delta_{\psi}^{X_1}(k_1^z, \ell_1) \Delta_{\psi}^{X_2}(k_2^z, \ell_2) \Delta_{h}^{X_3}(k_3^z, \ell_3) \]
\[ \times (2\pi) \delta(k_1^z + k_2^z + k_3^z)(\ell_2 \times \ell_3) \left[ k_2^z \ell_1 \cdot \ell_3 - k_1^z \ell_2 \cdot \ell_3 \right] \frac{-2\sqrt{2}}{\ell_3^2 k_3^R k_1^R k_2^R}. \]

(4.34)

This result is a slight extension of the work in Meerburg et al. (2016) to include all the contributions from temperature and polarization. In the Meerburg et al. (2016) they considered a $\psi\psi h$ template motivated by Maldacena (2003). In this work we consider shape functions described in section 4.3.2 so with the local, orthogonal and equilateral form. Choosing these functional forms is somewhat ad-hoc as there are very few theoretical models for these types of non-Gaussianity. However as was seen in Agrawal et al. (2017) and Shiraishi et al. (2013) some tensor non-Gaussianities are well approximated equilateral templates, thus we expect that constraints on these templates should cover at least part of the theoretical space of templates.

4.3.4 Primordial tensor-tensor-scalar non-Gaussianity

Primordial tensor-tensor-scalar non-Gaussianity has quite limited theoretical motivation. In single field inflation, Maldacena (2003) showed that it would be generated at unobservable levels. Recent work by Bartolo & Orlando (2017) showed the tensor-tensor-scalar non-Gaussianity arises in inflationary models where the inflation is coupled to Chern-Simons term and Gao et al. (2013) showed tensor-tensor-scalar that can arise in single field inflationary models with modified gravity. Further work has shown that tensor-tensor-scalar non-Gaussianity could be a useful probe of the speed on primordial gravitational waves (Noumi & Yamaguchi, 2014)
In this work, we consider templates motivated by those in Maldacena (2003).

Explicitly we assume a primordial shape of the following form

\[
\langle \psi(k_1)h^{s_2}(k_2)h^{s_3}(k_3) \rangle = r f^{\psi hh} 4\pi^4 \delta(k_1 + k_2 + k_3) I(k_1, k_2, k_3) \epsilon_{i,j}^{s_2}(k_2) \epsilon_{i,j}^{s_3}(k_3), \tag{4.35}
\]

where \( I(k_1, k_2, k_3) \) is a function describing the template shapes (in our case the local, equilateral and orthogonal shapes) that depends only on the magnitude of the wave-vectors and \( \epsilon_{i,j}^{s}(\hat{k}) \) is, as above, the transverse traceless polarization tensor.

Using the same method as above we can evaluate the trace of the product of the polarization tensor. There are three distinct cases: terms with two B mode legs, terms with one B mode leg and terms with no B mode legs. These lead to different contributions to the bispectrum as the polarization sums are evaluated differently for the different cases. For the case with two B modes the polarization sum is

\[
\sum \sum (\text{Sign}(s_2) \text{Sign}(s_3) \epsilon_{i,j}^{s_2}(k_2)) \epsilon_{i,j}^{s_3}(k_3) \tag{4.36}
\]

and this leads to a bispectrum contribution of

\[
\bar{b}^{X_{1,B,B}}(\ell_1, \ell_2, \ell_3) = 4\pi^4 A_s^2 r f^{\psi hh} \int \int \int \frac{dk_1^z}{2\pi} \frac{dk_2^z}{2\pi} \frac{dk_3^z}{2\pi} I(k_1^R, k_2^R, k_3^R) \Delta_X^{s_1}(k_1^z, \ell_1) \Delta_h(k_2^z, \ell_2) \Delta_h(k_3^z, \ell_3) 2\pi \delta(k_1^z + k_2^z + k_3^z) \left[ \ell_2 \cdot \ell_3 + k_2^z k_3^z D^2 \left( 1 - 2\left(\frac{\ell_2 \times \ell_3}{(\ell_2 \ell_3)^2}\right)^2 \right) \right]. \tag{4.37}
\]

For the case with one B modes the polarization sum is

\[
\sum \sum (\text{Sign}(s_3) \epsilon_{i,j}^{s_3}(k_3))(\epsilon_{i,j}^{s_3}(k_3)) \tag{4.38}
\]
and this leads to a bispectrum contribution of
\[
\tilde{b}_{X_1,X_2,B}(\ell_1, \ell_2, \ell_3) = 4\pi^4 A_{r}^2 f_{NL}^{hh\psi} \int \int \int \frac{dk_1^z \, dk_2^z \, dk_3^z}{2\pi^2} \frac{dk_1^R \, dk_2^R \, dk_3^R}{2\pi^2} \frac{dk_1^R \, dk_2^R \, dk_3^R}{2\pi^2} I(k_1^R, k_2^R, k_3^R) \Delta_{X_1}^{X_1}(k_1^z, \ell_1) \Delta_{X_2}^{X_2}(k_2^z, \ell_2) \times \Delta_{h}^{B}(k_3^z, \ell_3) 2\pi \delta(k_1^z + k_2^z + k_3^z) \left( k_2^z \ell_1 \cdot \ell_3 - k_1^z \ell_2 \cdot \ell_3 \right) .
\]

Finally for the case with no B modes the polarization sum is
\[
\sum_{\ell_1} \sum_{\ell_2} \sum_{\ell_3} (\epsilon_{i,j}^{s_2}(k_2))(\epsilon_{i,j}^{s_3}(k_3))
\]

and this leads to a bispectrum contribution of
\[
\tilde{b}_{X_1,X_2,X_3}(\ell_1, \ell_2, \ell_3) = 4\pi^4 A_{r}^2 f_{NL}^{hh\psi} \int \int \int \frac{dk_1^z \, dk_2^z \, dk_3^z}{2\pi^2} \frac{dk_1^R \, dk_2^R \, dk_3^R}{2\pi^2} \frac{dk_1^R \, dk_2^R \, dk_3^R}{2\pi^2} I(k_1^R, k_2^R, k_3^R) \Delta_{X_1}^{X_1}(k_1^z, \ell_1) \Delta_{X_2}^{X_2}(k_2^z, \ell_2) \times \Delta_{h}^{X_3}(k_3^z, \ell_3) 2\pi \delta(k_1^z + k_2^z + k_3^z) \left[ 2k_2^zh_3^z \ell_1 \cdot \ell_3 + (D^R k_3 k_2)^2 + (D^R k_3 k_2)^2 + \left( \frac{\ell_2 \cdot \ell_3}{D^R} \right)^2 \right] .
\]

### 4.3.5 Primordial tensor-tensor-tensor non-Gaussianity

Primordial tensor non-Gaussianity could be a key method of studying the early universe and is present, in observable levels, in a variety of inflationary models (Agrawal et al., 2017; Barnaby et al., 2012; Cook & Sorbo, 2013; Ferreira & Sloth, 2014). A model-independent template of the equilateral-type, polarized tensor-tensor-tensor bispectrum has been constrained using Planck data to \(400 \pm 1500\) (Planck Collaboration XVII, 2016). Given that the limits we can obtain using the ACTPol data are significantly worse than the existing limits, we do not perform any searches for tensor-tensor-tensor templates.
4.3.6 Foreground bispectra

There are several potential bispectra foregrounds that could contaminate our measurements. Galactic foregrounds are highly non-Gaussian (Planck Collaboration Int. XIX, 2015) and could bias our results. However, as we use the *Planck* cleaned maps and do not use the largest scales, $\ell > 100$, we expect this to have a negligible effect on this work. The bispectrum of these galactic foregrounds will become increasingly important for future measurements that use larger fractions of sky and with lower noise levels, as such these bispectra are being explored in ongoing work. Extra-galactic sources such as the thermal Sunyaev Zel’dovich effect, dusty star forming galaxies and lensing also are have measurable bispectra (Crawford et al., 2014; Coulton et al., 2017). Whilst the bispectrum from extra-galactic sources can have significant contributions in the shapes considered in this work (Lacasa et al., 2014) only the lensing and reionization terms will have significant contributions in the polarized maps. Thus we only consider the effect of these sources.

**Lensing-secondary sources bispectra**

Lensing-secondary source bispectra have been studied extensively in the literature (Spergel & Goldberg, 1999a; Goldberg & Spergel, 1999; Hu, 2000a; Lewis et al., 2011) and so here we just briefly review their origin and structure. We can decompose the anisotropies into contributions from the early universe and late time sources

$$\Delta T(n) = \Delta T^P(n + \partial \phi) + \Delta T^S(n),$$  \hspace{1cm} (4.42)

where $\Delta T^P$ is the fluctuations at the surface of last scattering, $\phi$ is the lensing potential, and $\Delta T^S$ are the contributions from late time sources such as radio galaxies or clusters via the Sunyaev Zel’dovich effect. If we expand in the lensing potential to
first order

\[ \Delta T(n) = \Delta T^P(n) + \partial \phi \cdot \partial \Delta T^P(n) + \Delta T^s(n) \quad (4.43) \]

\[ P^{a,b}(n) = P^{a,b,P}(n) + \partial \phi \cdot \partial P^{a,b,P}(n) + P^{a,b,S}(n). \quad (4.44) \]

From this we see that we can get non vanishing bispectra arising from terms of the form \( \langle \Delta T^P(n) \partial \phi \cdot \partial \Delta T^P(n) \Delta T^s(n) \rangle \) as the secondary sources are correlated with the structures which cause the lensing. More precisely these generate the following reduced bispectra

\[ \tilde{b}^{TTT}_{\ell_1,\ell_2,\ell_3} = - \ell_1 \cdot \ell_2 C^{TT}_{\ell_1} C^{\phi S}_{\ell_2} + 5\text{permutations} \quad (4.45) \]

\[ \tilde{b}^{TTE} = - \ell_1 \cdot \ell_3 C^{TT}_{\ell_1} C^{\phi P}_{\ell_3} - \ell_3 \cdot \ell_2 C^{TE}_{\ell_3} C^{\phi S}_{\ell_2} \]

\[ \quad - \ell_1 \cdot \ell_2 C^{TE}_{\ell_1} C^{\phi S}_{\ell_2} [\cos(2\phi_3) \cos(2\phi_1) + \sin(2\phi_3) \sin(2\phi_1)] + (\ell_1 \leftrightarrow \ell_2) \quad (4.46) \]

\[ \tilde{b}^{TTB} = - \ell_1 \cdot \ell_2 C^{TE}_{\ell_1} C^{\phi S}_{\ell_2} [\sin(2\phi_3) \cos(2\phi_1) + \sin(2\phi_3) \cos(2\phi_1)] + (\ell_1 \leftrightarrow \ell_2) \quad (4.47) \]

\[ \tilde{b}^{TEE} = - \ell_2 \cdot \ell_3 C^{TE}_{\ell_2} C^{\phi P}_{\ell_3} - \ell_1 \cdot \ell_3 C^{TE}_{\ell_3} C^{\phi S}_{\ell_2} [\cos(2\phi_2) \cos(2\phi_1) + \sin(2\phi_2) \sin(2\phi_1)] \]

\[ \quad - \ell_1 \cdot \ell_2 C^{EE}_{\ell_1} C^{\phi S}_{\ell_2} [\cos(2\phi_3) \cos(2\phi_2) + \sin(2\phi_3) \sin(2\phi_2)] + (\ell_2 \leftrightarrow \ell_3) \quad (4.48) \]

\[ \tilde{b}^{EEE} = - \ell_1 \cdot \ell_2 C^{EE}_{\ell_1} C^{\phi P}_{\ell_2} [\cos(2\phi_3) \cos(2\phi_1) + \sin(2\phi_3) \sin(2\phi_1)] + 5\text{permutations}, \quad (4.49) \]

where \( C^{\phi S} \) is the cross correlation between the lensing potential and the secondary source. The reduced bispectra \( b^{TEE} \) and \( b^{TTE} \) contain both of the above terms. Lewis et al. (2011) showed that the effect of higher order terms in \( \phi \) lead to \( \sim 10\% \) corrections to the perturbative results in equation (4.45) and they showed these higher order terms could accurately be approximated by replacing \( C^{XX} \) with the lensed power spectrum.
4.4 Data Sets and Pipeline

In this work we use data from the Atacama Cosmology Telescope and combine this with data from the *Planck* Satellite.

4.4.1 ACTPol Maps

We use data in the ‘D56’ field, a patch of sky on the equator with coordinates $-7.2^\circ < \text{dec} < 4^\circ$ and $352^\circ < \text{RA} < 41^\circ$. This is part of the data described in [Louis et al. (2017)](Louis et al. (2017)); in particular we use only the wide field and not the deep fields. In this work we use the data from the two arrays (called PA1 and PA2 hereafter) which observed the sky at 148 GHz. We use both the temperature and polarization maps from the experiment. For the temperature maps we mask all the point sources whose fluxes were measured to be above 15 mJy with discs of radius $5'$ and perform no masking of clusters or of galactic dust. For polarized maps no masking of point sources or clusters is performed.

We apply a $k$-space mask to these maps which removes modes with $\ell < 500$ and $\ell < 350$ in temperature and polarization respectively. The temperature mask is used as the largest scales are dominated by non-white atmospheric noise, see [Louis et al. (2017)](Louis et al. (2017)) for a more detailed discussion. In polarization we mask the largest scales as the transfer function deviates from one and these deviations are not accounted for in this analysis.

4.4.2 Planck Maps

In this work, we use the part of Spectral Matching Independent Component Analysis (SMICA) maps that overlap with ‘D56’ field, which is described in section . The SMICA maps are produced by combining the individual *Planck* frequency maps in a manner to remove galactic and extra-galactic foregrounds. The SMICA foreground
cleaning method is a non-parametric method that models the foregrounds with as a small number of components with arbitrary frequency dependence and arbitrary power spectra. The parameters for this model are fit from the measured auto and cross power-spectra. This model can then be used to weight the spherical harmonics of the frequency maps to obtain a cleaned CMB map. For more details we refer the reader to Cardoso et al. (2008); Planck Collaboration XII (2014); Planck Collaboration IX (2016). Before using the maps, we reprojected Planck healpix maps onto the ACTPol pixels by first expanding them in spherical harmonic coefficients, rotating these coefficients from galactic to equatorial coordinates using healpy, and then evaluating these coefficients on the ACTPol pixels using libsharp (Reinecke & Seljebotn, 2013). Due to the computation complexity associated with analyzing multiple maps, in this work we only use the Planck SMICA B mode map.

We use the SMICA confidence mask to mask point sources, clusters and galactic emission; the details of the masks can be found in Planck Collaboration IX (2016). The Planck SMICA polarization maps have beams of 5 arcmin. Finally we apply a k-space mask to these maps which removes modes with \( \ell < 150 \) in polarization respectively. The lowest \( \ell \) modes are masked as they are most effected by the real space mask. Further we found that after masking these modes the linear correction term became negligible. For our templates the calculation of the linear term is very expensive and so in this work we sacrifice the low \( \ell \) signal to avoid calculating the linear term.

### 4.4.3 Analysis Pipeline

In this work we use the pipeline that was described and verified in Coulton et al. (2017), which we briefly outline here. First we apply a point source mask, with levels described in sections 4.4.2 and 4.4.2, no masking is applied to the ACTPol polarization maps. We then use an in-painting routine to iteratively apodize the
<table>
<thead>
<tr>
<th>Type</th>
<th>Measured Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tensor-Scalar-Scalar $\sqrt{f_{NL}}$</td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td>$-237 \pm 459$</td>
</tr>
<tr>
<td>Equilateral</td>
<td>$-1040 \pm 1360$</td>
</tr>
<tr>
<td>Orthogonal</td>
<td>$2000 \pm 1630$</td>
</tr>
<tr>
<td>Tensor-Tensor-Scalar $f_{NL}$</td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td>$24000 \pm 23900$</td>
</tr>
<tr>
<td>Equilateral</td>
<td>$51000 \pm 103000$</td>
</tr>
<tr>
<td>Orthogonal</td>
<td>$-162000 \pm 108000$</td>
</tr>
</tbody>
</table>

Table 4.1: Preliminary limits obtained from combining ACTPol 148GHz T+E with Planck SMICA B mode with $\ell_{max} = 5000$ and $\ell_{min} = 350$ and $\ell_{min} = 150$ for ACTPol and Planck respectively.

point source mask (Gruetjen et al., 2015). This routine fills in each masked pixel with the average of the neighbors and the process is iterated until the process has converged. After this we apply a mask based on the square root of the hit counts to apodize the map and suppress the noisiest pixels. We then apply a cosine squared mask to remove areas of the map with few than 5500 hits in the ACTPol maps. For the polarized maps we use the same hit count and cosine mask and then use the pure E and B estimator described in Louis et al. (2013); Smith & Zaldarriaga (2007); Smith (2006) to calculate the E and B components. In the final step we apply the $k$-space masks described in sections 4.4.2 and 4.4.2.

4.5 Results

In table 4.1 we present our preliminary results obtained by using the ACTPol T and E maps in combination with the Planck SMICA B mode map. We find that there is no evidence for any tensor-scalar-scalar or tensor-tensor-scalar non-Gaussianity. It should be noted that these preliminary results are not jointly fit with the lensing cross secondary source templates discussed in 4.3.6. To prevent biases this needs to be done and is the subject of on-going work.
4.6 Conclusion

We have described a set of theoretical templates to constrain part of the space of possible tensor-scalar non-Gaussianity. In Appendix 4.7 we described a method for generating general separable non-Gaussianity in the flat-sky regime and used it to validated our pipeline. The preliminary results in section 4.5 describe the current state of this work. In on-going work we are testing the robustness of these results, fitting these templates jointly with the lensing-ISW templates and will use the ACTPol data alone to constraint the different forms of non-Gaussianity. This will allow a cross check of systematics. With robust constraints physical theories such as solid inflation, which predicts non-zero tensor-scalar-scalar non-Gaussianity, can be constrained, though to provide the tightest constraints model specific templates should be used. However as our constraints are very large (such that power-spectrum constraints will provide stricter constraints on many models) and still preliminary, we leave an exploration of this to future work.

Beyond the short term, primordial tensor non-Gaussianity is an interesting probe of the early universe. Unlike scalar-scalar-scalar non-Gaussianity there is a large amount more we can learn about tensor non-Gaussianity from the CMB. Upcoming experiments will vastly improve the constraints on these types of non-Gaussianity and potentially provide a new look at the primordial universe.

4.7 Appendix: Simulating non-Gaussianity in the Flat-Sky

In order to validate our estimator it is useful to simulate non-Gaussian maps. In Coulton et al. (2017) we tested our analysis pipeline using three types of non-Gaussianity. In this work we constrain odd parity templates and it is beneficial to verify that
our estimator is unbiased for these templates, as well as the parity even templates tested before. In this section we describe how tensor-scalar non-Gaussianity can be simulated in the flat-sky regime and then test our estimator on simulated odd-parity maps. In the first subsection we describe a general method for simulating primordial bispectra that are separable. In the second subsection we use this method to generate maps with tensor-scalar-scalar non-Gaussianity and use them to test our pipeline.

4.7.1 Flat-sky simulation methods

We wish to simulate bispectra with the structure given in eq. 4.18 and our approach to simulating primordial non-Gaussianity is a flat sky adaptation of Liguori et al. (2003). The desired non-Gaussianity has to be sufficiently smooth (Babich et al., 2004) so that we can treat the bispectrum as constant over the width of the last scattering surface. Finally we only consider generating bispectra whose primordial bispectrum is separable or equivalently

\[
B_{Y_1,Y_2,Y_3}(k_1, k_2, k_3) = f_{NL}F_{1,x}(k_1^x)F_{1,y}(k_1^y)F_{1,z}(k_1^z)F_{2,x}(k_2^x)F_{2,y}(k_2^y)F_{2,z}(k_2^z)\times
\]

\[
F_{3,x}(k_3^x)F_{3,y}(k_3^y)F_{3,z}(k_3^z)P_{Y_1}(k_1)P_{Y_2}(k_2) + \text{permutations.} \quad (4.50)
\]

In order to avoid generating tensor fields we instead utilise the fact that we only need to simulate the projections \( \sum_\pm h^x \) and \( \sum_\pm \pm h^\pm \). These fields behave as a scalar and pseudo-scalar field and so are simpler to simulate. In the first step we generate Gaussian realisations of \( \zeta \), \( \sum_\pm h^x \) and \( \sum_\pm \pm h^\pm \) on a grid.

\[
\zeta^{(G)}(k) \sim N(0, P(k)) \quad (4.51)
\]

\[
h^{x(G)} = \sum_\pm h^x(k) \sim N(0, P_t(k)) \quad (4.52)
\]

\[
h^{d,(G)} = \sum_\pm \pm h^\pm(k) \sim N(0, P_t(k)) \quad (4.53)
\]
where $P(k)$ is the scalar primordial fluctuation power spectrum, $P_t(k)$ is the tensor primordial fluctuation power spectrum and we defined $h^p$ and $h^d$ for notational simplicity. In this work we fix $P_t(k) = rP(k)$, where $r$ is the tensor to scalar ratio.

The dimensions of the grid in $(x, y)$ plane are fixed by the desired output $\ell$ modes as $k_x = \ell_x/ (\tau_0 - \tau_r)$ and $k_y = \ell_y/ (\tau_0 - \tau_r)$. The grid dimension in the $z$ direction is set by the $k_z$ integration. We discretize the $k_z$ integrals in eq. (4.18) and evaluate the integrals with Simpson’s integration. The number of points used in the $k_z$ integration sets the dimension of the grid.

We then transform copies of these Gaussian realizations to generate a non-linear component. For scalar-scalar-scalar non-Gaussianity we perform the following transformation

$$
\xi^{(NL)}(k) = f_{NL} F_{1,x}(k_x^1) F_{1,y}(k_y^1) F_{1,z}(k_z^1) \\
\times \int d^3 \tilde{k} F_{2,x}(\tilde{k}_x^x) F_{2,y}(\tilde{k}_y^y) F_{2,z}(\tilde{k}_z^z) \xi^{(G)}(k) F_{3,x}(k_x^x + \tilde{k}_x^x) F_{3,y}(k_y^y + \tilde{k}_y^y) F_{3,z}(k_z^z + \tilde{k}_z^z) \xi^{(G)}(k + \tilde{k}). 
$$

(4.54)

The convolution can be performed via two Fourier transforms. We then generate maps by evolving the perturbations as follows

$$
da^X(\ell) = \int \frac{dk_z}{2\pi} \Delta^X(k_z, \ell) \left( \xi^{(G)}(k) + \xi^{(NL)}(k) \right).
$$

(4.55)

To generate tensor-scalar non-Gaussianity we can simply replace one or more of the $\xi^{(G)}$ fields with $h^{s(G)}$ or $h^{d(G)}$ in eq. (4.54).

Whilst the method described here is inefficient, requiring several 3D Fourier transforms per non-Gaussian map, it means that we can generate maps with any smooth and separable primordial bispectrum.
4.7.2 Odd Parity Estimator Validation

We use the method described above to validate our estimator. We generate maps with local tensor-scalar-scalar bispectra described in section 4.3.3. For the BTT bispectrum the primordial template is

\[
B^\ell_{\zeta-h}^{\zeta-h}(k_1, k_2, k_3) = 16\pi^4 A_s^2 \sqrt{r f_{NL}} i \sqrt{2} \frac{(k_x^2 k_3^2 - k_2^2 k_3^2)}{(k_3^2 + k_3^2) k_3^2} \left( -k_2^z \left( k_3^z + k_3^2 \right) + k_3^z \left( k_3^x k_3^y + k_3^x k_3^y \right) \right) \times (P_\Phi(k_1)P_\Phi(k_2) + P_\Phi(k_1)P_\Phi(k_3) + P_\Phi(k_2)P_\Phi(k_3))
\]

(4.56)

This has the required form given in eq. 4.54 and after expanding the brackets it is straightforward to read off the \( F_i \) kernels.

Using these pieces we generate non-Gaussian maps and analyze them with our pipeline. We use maps of size \( 8^2 \) deg\(^2\) to test our pipeline. In figure 4.1 we show the results of applying our estimator to 100 non-Gaussian maps. We also simulated non-Gaussian maps using the method described in Smith & Zaldarriaga (2011) and the results of using our estimator on these maps is also shown in figure 4.1. We find that our estimator recovers the input non-Gaussianity for Smith & Zaldarriaga (2011) method but is a factor of 2 off for the method described above. We are working on understanding the source of this discrepancy. This discrepancy is not seen for the Smith & Zaldarriaga (2011) method, however as our estimator normalization is related to the Smith & Zaldarriaga (2011) simulation method it is not surprising that these agree. As a byproduct of this, we can also verify that the variance of the estimator, evaluated when \( f_{NL} = 0 \), is consistent with the Fisher estimate. We find the standard deviation of the estimator is \( \sigma_{\text{measured}}(\sqrt{r f_{NL}}) = 409 \) and the Fisher prediction is \( \sigma_{\text{Fisher}}(\sqrt{r f_{NL}}) = 416 \). This suggests that the discrepancy seen in figure 4.1 arises in our simulation method and is the subject of ongoing work.
Figure 4.1: A preliminary test of the new non-Gaussianity simulation method, as described in §4.7.1, to simulated tensor-scalar-scalar BTT non-Gaussianity. We see that there is an unknown issue that causes a discrepancy between the simulated and recovered value. We also simulate non-Gaussianity via the method described in Smith & Zaldarriaga (2011).
Chapter 5

Exploring Galactic foregrounds
with the bispectrum

This chapter was done under the supervision of Professor David Spergel and has benefited from discussions with Dr Sigurd Naess and Dr Steve Choi. The results of this work have been presented at a Gravity Group meeting in Princeton.

5.1 Introduction

One of the main goals of current and up-coming cosmic microwave background (CMB) surveys is to detect the imprint of primordial tensor modes. Primordial tensor modes are theorised to have been generate in the early universe during inflation (Grishchuk 1975; Starobinskij 1979; Rubakov et al. 1982) and these tensor modes then leave an imprint on the CMB at the surface of last scattering (Fabbri & Pollock 1983; Abbott & Wise 1984). Their contribution to the temperature anisotropies (T mode) and curl-free component of the polarised CMB (E mode) has found to be masked by the dominant scalar modes (Spergel et al. 2007). Instead cosmologist seek to measure the tensor modes by measuring the curl component of the polarised CMB (B mode) (Seljak 1997; Seljak & Zaldarriaga 1997). Scalar modes do not produce B mode
polarisation and so the detection of the B modes from the surface of last scattering would be strong evidence for primordial tensor modes (Polnarev, 1985; Zaldarriaga & Seljak, 1997; Kamionkowski et al., 1997) .

Whilst the level of the primordial B mode signal is unknown, the bounds from current experiments mean that it is unlikely that primordial signals will be the dominant sky signal at any frequency (BICEP2 Collaboration et al., 2016b). Currently there are three known sources of signal which could mask the primordial signal. These foregrounds are: polarised dust emission, polarised synchrotron emission and gravitational lensing. Polarized dust emission arises from asymmetric dust grains that are aligned with the galactic magnetic field. The dust emission can be described my a modified black body

\[ I_{\text{dust}} \propto \nu^{\beta_{\text{dust}}} B_\nu(T_{\text{dust}}), \]  

where \( B_\nu(T) \) is the Planck function, \( T_{\text{dust}} \) is the dust temperature (and the Planck best fit value is \( T_{\text{dust}} = 15.9 K \)), and \( \beta_{\text{dust}} \) is the dust spectral index which is \( \beta_{\text{dust}} = 1.48 \pm 0.01 \) for temperature and \( \beta_{\text{dust}} = 1.53 \pm 0.02 \) for polarization (Planck Collaboration et al., 2018). Dust emission is the dominant foreground at frequencies above \( \sim 150 \) GHz (Draine, 2004; Draine & Fraisse, 2009). Polarised synchrotron emission arises from electrons spiralling the galactic magnetic field and is dominant at lower frequencies (Kogut et al., 2007). At first order the Galactic synchrotron radiation’s spectral behaviour can be described by a power law

\[ I_{\text{sync}} \propto \nu^{\beta_{\text{sync}}}, \]  

where \( \beta_{\text{sync}} \) is the sky mean spectral index and the current best fit values are \( -3.09 \pm 0.05 \) at 23 GHz (Kogut, 2012) for temperature and \( \beta_s = -3.13 \pm 0.13 \) for polarisation (Planck Collaboration et al., 2018). It is uncertain as to why the temperature and polarised components have different spectral indices. However, there is strong
evidence that the synchrotron spectral index, for both polarised and temperature emission, varies across the sky (Miville-Deschênes et al., 2008; Fuskeland et al., 2014) and shows spectral steepening (Kogut, 2012). Gravitational lensing B modes are generated as light propagates from the surface of last scattering to the observer. The light is lensed by intervening matter and the lensing shears the E mode signal to B mode signal (Seljak, 1996; Blanchard & Schneider, 1987).

The primordial B mode signal distinguishes itself from these other sources as it has a blackbody spectrum with temperature $2.726 \pm 0.010$ K (Mather et al., 1994) and is thought to have highly Gaussian fluctuations. Dust and synchrotron emission have distinctly different spectra and whilst gravitational lensing signals have an identical spectra, they are highly non-Gaussian. Many different approaches to foreground cleaning have been proposed, see Ichiki (2014) for a review of some of these approaches; several of which have successfully been applied to Planck data (Planck Collaboration IX, 2016; Planck Collaboration XII, 2014). Methods to remove lensing, known as delensing, are described in Seljak & Hirata (2004); Sehgal et al. (2017). The properties of galactic foregrounds are less well constrained than the lensing foregrounds and are the focus of this work.

To remove foregrounds with high precision the foregrounds need to be accurately characterised. Recent measurements have started to constrain the spatial and spectral properties of the dust and synchrotron. Planck measurements have characterised the distribution of polarisation fractions, their correlation with intensity measurements and studied how these properties related to the galactic magnetic field (Planck Collaboration Int. XIX, 2015). Further they found that the galactic dust E mode power has roughly a factor of two more power than the B mode and that the dust T-B power-spectrum is non-zero (Planck Collaboration et al., 2018). Work by Planck Collaboration Int. XIX (2015); Rotti & Huffenberger (2016) have shown that the dust polarised emission varies significantly across the sky. Outside the galactic plane
polarised synchrotron emission is mainly from large filaments with high polarisation fractions and these have been extensively studied (e.g. Kogut et al. 2007; Planck Collaboration XXV 2016) including in the context of CMB B mode foregrounds (Vidal et al. 2015). Building on the work of Kogut et al. (2007); Page et al. (2007), Choi & Page (2015) explored the correlation between the synchrotron emission and the dust emission. Most recently Rana et al. (2018) explored the properties of the temperature bispectrum synchrotron emission at 408 MHz.

Complimenting these measurements there has been extensive theoretical work on the properties of the foregrounds. Theoretical models (Caldwell et al., 2017) and numerical magnetohydrodynamics (Kritsuk et al., 2017; Vansyngel et al., 2017) have explored the physics behind the observed ratio of dust EE to BB power. Work by Kamionkowski & Kovetz (2014) suggested that galactic dust is likely to be highly anisotropic, in particular it should have a local hexadecapolar type deviations. This was further explored in Philcox et al. (2018), where this anisotropy was measured in simulations and the detectability of the local hexadecapole with future surveys was discussed.

Once foreground cleaned CMB maps are obtained tools are needed to validate that any remaining signal is not residual foregrounds. The importance of validating cleaned maps was highlighted in recent work (e.g. Madhavacheril & Hill 2018) that found the Planck foreground cleaned temperature maps contain significant residual thermal Sunyaev Zel’dovich signal. Traditional approaches rely on cross correlations with foreground dominated maps to constrain residual foregrounds. Whilst cross correlations are a very useful tool, alone they may not be sufficient to validate a potential primordial signal, particularly if the power spectra have been used to remove the foregrounds or if foregrounds de-correlate with frequency, though there is no evidence to show that they do (Sheehy & Slosar 2018). Kamionkowski & Kovetz (2014) and Rotti & Huffenberger (2016) suggested using measurements of the anisotropy to char-
acterise remaining foregrounds, as the primordial signal is expected to be isotropic but residual foregrounds are not.

Our work explores how the bispectrum can be used to characterise the foregrounds and test for residual signals, exploiting the fact that the primordial signal is expected to be highly Gaussian. The bispectrum is the harmonic equivalent of the three point function and should vanish for Gaussian signals. The bispectrum is the lowest order non-Gaussian statistic. We first seek to characterise the bispectrum of the galactic foregrounds using data from the Planck satellite. Then we demonstrate how measurements of the bispectrum can be used to test for residual foregrounds in foreground-cleaned maps. This work builds on previous bispectrum measurements of galactic foregrounds Komatsu & Seljak (2002), which used bispectrum measurements of foregrounds to avoid biases in primordial non-Gaussianity estimators. Understanding the bispectrum from galactic foregrounds, and especially any residual foregrounds in cleaned maps, is very important for constraints on primordial non-Gaussianity.

We use a binned bispectrum approach (Bucher et al., 2010, 2016). The binned bispectrum is blind approach that requires no theoretical model of the signal and can constrain non-smooth signals. The cost of this broad approach is that it can be less optimal constraints that more targeted approaches. The details of our implementation of the binned bispectrum are described in section 5.2. In section 5.3 we describe the data sets used in this work and briefly outline our analysis pipeline. Our results are presented in 5.4 and then are discussed, along with our conclusions, in section 5.5.

5.2 Bispectrum Estimator

In this section we will briefly overview the binned bispectrum estimator for the parity even and odd cases. For both our estimators we must decompose the maps into spherical harmonic components. The measurements of the stokes I component, $\Delta T(\mathbf{n})$,
are decomposed using spin-zero harmonics as

\[ a_{\ell,m} = \sum_n Y_{\ell,m}(n) \Delta T. \]  

(5.3)

When no mask is applied the the stokes Q and U polarisation components can be decomposed as

\[ (Q \pm iU)(n) = \sum_{\ell,m} a_{\pm 2,\ell,m \pm 2} Y_{\ell,m}(n) \]  

(5.4)

\[ a_{E,\ell,m} = -(a_{2,\ell,m} + a_{-2,\ell,m})/2 \]  

(5.5)

\[ a_{B,\ell,m} = i(a_{2,\ell,m} - a_{-2,\ell,m})/2 \]  

(5.6)

When a sky mask is applied the decomposition mixes E and B modes together. In this work we use pure E and B estimators, as described in Smith & Zaldarriaga (2007); Grain et al. (2012), to obtain maps which are free of this leakage.

### 5.2.1 Parity Even Estimator

The bispectrum is given by the ensemble average of three spherical harmonic coefficients

\[ \langle a_{\ell_1,m_1}^X a_{\ell_2,m_2}^Y a_{\ell_3,m_3}^Z \rangle = B_{X,Y,Z}^{X,Y,Z}(\ell_1,\ell_2,\ell_3,m_1,m_2,m_3). \]  

(5.7)

where \( B_{X,Y,Z}^{X,Y,Z}(\ell_1,\ell_2,\ell_3,m_1,m_2,m_3) \) is the bispectrum between maps \( X,Y \) and \( Z \). Under the conditions of homogeneity and isotropy the bispectrum can be decomposed as (Komatsu & Spergel, 2001b)

\[ \langle a_{\ell_1,m_1}^X a_{\ell_2,m_2}^Y a_{\ell_3,m_3}^Z \rangle = \sqrt{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)} \left( \begin{array}{ccc} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{array} \right) \left( \begin{array}{ccc} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{array} \right) b_{X,Y,Z}^{X,Y,Z}(\ell_1,\ell_2,\ell_3) \]  

(5.8)
where
\[
\begin{pmatrix}
\ell_1 & \ell_2 & \ell_3 \\
m_1 & m_2 & m_3
\end{pmatrix}
\]
(5.9)

is the Wigner 3j symbol and \(b_{\ell_1,\ell_2,\ell_3}\) is the reduced bispectrum. The Wigner 3j symbol is related to the Clebsh-Gordan coefficients that describe the coupling of two angular momenta. The Wigner 3j symbol vanishes unless the triangle conditions are satisfied. They are require that
\[
|\ell_i - \ell_j| \leq \ell_k \leq \ell_i + \ell_j.
\]
(5.10)

This estimator is zero unless \(\ell_1 + \ell_2 + \ell_3 = \text{even}\) and so is only sensitive to even parity configurations. The foregrounds are neither isotropic nor homogeneous and by focusing on the reduced bispectrum we lose any anisotropic information (this is also an issue for most power spectra approaches). It is left to future work to explore the anisotropic contributions to the bispectra.

We use a binned bispectrum method similar to that described in [Bucher et al. (2016)]. Using the fact that the Wigner 3j symbols can be evaluated using the Gaunt integral
\[
\mathcal{G}^{m_1,m_2,m_3}_{\ell_1,\ell_2,\ell_3} = \int d^2n Y_{\ell_1,m_1}(n) Y_{\ell_2,m_2}(n) Y_{\ell_3,m_3}(n)
\]
\[
= \sqrt{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1) \over 4\pi} \left| \begin{pmatrix}
\ell_1 & \ell_2 & \ell_3 \\
m_1 & m_2 & m_3
\end{pmatrix}
\right| \begin{pmatrix}
\ell_1 & \ell_2 & \ell_3 \\
0 & 0 & 0
\end{pmatrix},
\]
(5.11)

we can estimate the reduced bispectrum as
\[
b_{\ell_1,\ell_2,\ell_3}^{X,Y,Z} = \frac{1}{N_{\ell_1,\ell_2,\ell_3}} \sum_m \int d^2n Y_{\ell_1,m_1}(n) Y_{\ell_2,m_2}(n) Y_{\ell_3,m_3}(n) a_{\ell_1,m_1}^X a_{\ell_2,m_2}^Y a_{\ell_3,m_3}^Z,
\]
(5.12)
where $N_{\ell_1,\ell_2,\ell_3}$ is the appropriate weight. The binned estimator is a simple modification of the above formula. The maps are filtered in harmonic space to contain only modes with $\ell$ within the bin. The filtered map with $\ell$ satisfying $\ell_i < \ell \leq \ell_{i+1}$ is denoted as $W_i^X(n)$ and is given explicitly by

$$W_i^X(n) = \sum_{\ell_i < \ell \leq \ell_{i+1}} \sum_{m} \int d^2n Y_{\ell,m}(n) a_{\ell,m}^X.$$  

The binned bispectrum estimator is then given by

$$\hat{b}_{i,j,k}^{X,Y,Z} = \frac{1}{N'_{i,j,k}} \int d^2n W_i^X(n) W_j^Y(n) W_k^Z(n).$$  

The estimator normalisation is given by

$$N'_{i,j,k} = \sum_{\ell_i < \ell_1 \leq \ell_{i+1}} \sum_{\ell_j < \ell_2 \leq \ell_{j+1}} \sum_{\ell_k < \ell_3 \leq \ell_{k+1}} \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}.$$  

For Gaussian signals the variance of the estimator is given by

$$V_{i,j,k}^{X,Y,Z,X',Y',Z'} = \frac{1}{N'_{i,j,k} N'_{i,j,k}} \sum_{\ell_1 < \ell_{i+1}} \sum_{\ell_2 < \ell_{j+1}} \sum_{\ell_3 < \ell_{k+1}} \frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi} \times \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 C_{\ell_1}^{X,X'} C_{\ell_2}^{Y,Y'} C_{\ell_3}^{Z,Z'} g_{i,j,k},$$  

where $C_{X,Y}$ is the power spectrum between map $X$ and map $Y$ and $g_{i,j,k}$ is 6 all its arguments are equal, 2 if any two of the arguments are equal and 1 if none of them are equal. When masks are applied to the data the variance of the estimator is altered. In this work the variance is calculated using 100 simulations.
5.2.2 Parity Odd Estimator

A non-parity violating process should only produce parity even T and E bispectra (Shiraishi et al., 2011; Kamionkowski & Souradeep, 2011), however as Planck Collaboration et al. (2018) found BT correlations in the dust, it is possible that the dust also exhibits parity-odd T and E bispectra. This motivated us to consider parity odd bispectra, those with $\ell_1 + \ell_2 + \ell_3 = \text{odd}$. These estimators were originally investigated in the context of parity violating inflationary models and searches for primordial magnetic fields (Shiraishi et al., 2012, 2013), however in the absence of parity violating signals bispectra involving odd number of B modes naturally have odd parity (Meerburg et al., 2016). The reduced bispectrum estimator described above is only sensitive to bispectra with even parity, those that satisfy $\ell_1 + \ell_2 + \ell_3 = \text{even}$. To measure these bispectra we introduce a second estimator based on the work of (Shiraishi et al., 2014, 2015). Our odd parity binned estimator vanishes for bispectra that do not satisfy the triangle conditions or have $\ell_1 + \ell_2 + \ell_3 = \text{even}$. For notional convenience we define the following function

$$h_{\ell_1,\ell_2,\ell_3}^{m_1,m_2,m_3} = \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)(2\ell_3 + 1)}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} 2 & -1 & -1 \\ -2 & 1 & 1 \end{pmatrix} = \int d^2n (-2Y_{\ell_1,m_1}(n)Y_{\ell_2,m_2}(n)Y_{\ell_3,m_3}(n) - 2Y_{\ell_1,m_1}(n)Y_{\ell_3,m_2}(n)-1Y_{\ell_3,m_3}(n)) .$$

(5.17)

We define the odd parity reduced bispectrum as

$$\langle a_{\ell_1}^X a_{\ell_2}^X a_{\ell_3}^X \rangle = \frac{1}{6} \left( h_{\ell_1,\ell_2,\ell_3}^{m_1,m_2,m_3} + h_{\ell_3,\ell_1,\ell_2}^{m_1,m_2,m_3} + h_{\ell_2,\ell_3,\ell_1}^{m_1,m_2,m_3} \right) h_{\ell_1,\ell_2,\ell_3}^{\text{odd},X,Y,Z} .$$

(5.18)
Now we can write down an estimator for the full odd parity bispectrum as

\[
\hat{b}_{\ell_1,\ell_2,\ell_3}^{odd,X,Y,Z} = \frac{1}{6N_{\ell_1,\ell_2,\ell_3}^{o}} \left( h_{\ell_1,\ell_2,\ell_3}^{m_1,m_2,m_3} + h_{\ell_5,\ell_1,\ell_2}^{m_3,m_1,m_2} + h_{\ell_3,\ell_5,\ell_1}^{m_2,m_3,m_1} \right) a_{\ell_1}^{X} a_{\ell_2}^{Y} a_{\ell_3}^{Z},
\]

(5.19)

In a similar manner to the parity even bispectrum we define a set of filtered maps. The filtered maps are obtained via a spin weighted spherical harmonic transform and are defined by

\[
s_{W_{i}^{X}} = \sum_{\ell_1 < \ell < \ell_{i+1}} \sum_{m} \int d^{2}x_{i} Y_{\ell_{i},m}(\mathbf{n}) a_{\ell_{i}}^{X} \]

(5.20)

Then the binned odd parity bispectrum estimator is given by

\[
\hat{b}_{i,j,k}^{odd,X,Y,Z} = \frac{1}{6N_{i,j,k}^{o}} \int d^{2}n \left( -2W_{i}^{X} W_{j}^{Y} W_{k}^{Z} - 2W_{i}^{X} W_{i-1}^{Y} W_{j}^{Z} \\
+ W_{k}^{X} W_{i}^{Y} W_{j}^{Z} - 2W_{k}^{X} W_{i}^{Y} W_{j}^{Z} + 2W_{j}^{X} W_{k}^{Y} W_{j}^{Z} - 2W_{j}^{X} W_{i}^{Y} W_{k}^{Z} \right). \]

(5.21)

In analogy to the parity even case, the normalization is

\[
N_{i,j,k}^{o} = \frac{1}{6} \sum_{\ell_{1} < \ell_{1} \leq \ell_{i+1}} \sum_{\ell_{2} < \ell_{2} \leq \ell_{j+1}} \sum_{\ell_{3} < \ell_{3} \leq \ell_{k+1}} \left( h_{\ell_{1},\ell_{2},\ell_{3}}^{m_1,m_2,m_3} + h_{\ell_{5},\ell_{1},\ell_{2}}^{m_3,m_1,m_2} + h_{\ell_{3},\ell_{5},\ell_{1}}^{m_2,m_3,m_1} \right) \times \left( h_{\ell_{1},\ell_{2},\ell_{3}}^{m_1,m_2,m_3} + h_{\ell_{5},\ell_{1},\ell_{2}}^{m_3,m_1,m_2} + h_{\ell_{3},\ell_{5},\ell_{1}}^{m_2,m_3,m_1} \right)^{2}
\]

(5.22)

and the variance is

\[
V_{\ell_1,\ell_2,\ell_3}^{oX,Y,Z,X',Y',Z'} = \frac{1}{N_{i,j,k}^{o} N_{i,j,k}^{o}} \sum_{\ell_{1} < \ell_{1} \leq \ell_{i+1}} \sum_{\ell_{2} < \ell_{2} \leq \ell_{j+1}} \sum_{\ell_{3} < \ell_{3} \leq \ell_{k+1}} \left( h_{\ell_{1},\ell_{2},\ell_{3}}^{m_1,m_2,m_3} + h_{\ell_{5},\ell_{1},\ell_{2}}^{m_3,m_1,m_2} + h_{\ell_{3},\ell_{5},\ell_{1}}^{m_2,m_3,m_1} \right)^{2} \times C_{\ell_{1}}^{X,X'} C_{\ell_{2}}^{Y,Y'} C_{\ell_{3}}^{Z,Z'} g_{i,j,k}.
\]

(5.23)
5.2.3 Linear Term

The estimators defined above are strictly only correct in the case that the data is homogenous and isotropic. In the case of a real analysis this condition is broken by the detector noise, masking and the galactic foreground signals. To account for these effects the estimator must be altered to include a linear correction term. As this term has been discussed thoroughly in the literature Creminelli et al. (2006); Yadav et al. (2008); Bucher et al. (2016) here we just summarise the results. The parity even estimator given in eq 5.14 is altered to

$$\hat{b}_{i,j,k}^{X,Y,Z} = \frac{1}{N_{i,j,k}} \int d^2n W_i^X(n)W_j^Y(n)W_k^Z(n) - W_i^X(n)\langle W^G_j^Y(n)W^G_k^Z(n) \rangle$$

$$- W_j^Y(n)\langle W^G_i^X(n)W^G_k^Z(n) \rangle - W_k^Z(n)\langle W^G_i^X(n)W^G_j^Y(n) \rangle$$

(5.24)

and the parity odd estimator given in eq. 5.21 is altered in an identical manner.

In our current calculation of the linear term we include the effects of masking and anisotropic noise but not the anisotropy of the sky signal. This is potentially very important and will be discussed further in sections 5.3.3 and 5.5.

5.3 Data Sets and Pipeline

5.3.1 Data sets

For this work we used data from the Planck 2015 data release (Planck Collaboration X, 2016; Planck Collaboration IX, 2016). For our fiducial studies of the foregrounds we use the Commander synchrotron map and the 353 GHz map as a tracer of the dust. For a detailed description of the commander component separation method we refer the reader to Eriksen et al. (2006, 2008) and here we very briefly overview the method. Commander is a Bayesian, parametric map-based method which sepa-
rates a set of different frequency maps into their components. The map components are modeled by an amplitude for each pixel and a spectral parameterization. The amplitudes and spectral indexes in each pixel are then obtained by an MCMC Gibbs sampling algorithm. The Commander method uses temperature data from the Planck Satellite, WMAP and a reprocessed version of the 408 MHz radio continuum all-sky map HASLAM map and only Planck data for the polarization maps [Haslam et al., 1982; Bennett et al., 2013; Remazeilles et al., 2015].

For our synchrotron analysis we used the Commander maximum likelihood synchrotron temperature map, which has a reference frequency of 408 MHz and is smoothed such that the effective beam full width half-maximum (FWHM) is 60 arcmin, and the Commander maximum likelihood synchrotron polarization maps, which have reference frequencies of 30GHz and are smoothed to have an effective FWHM of 40 arcmin. The effective beam FWHM of the 353 GHz maps is 4.86 arcmin.

In section 5.4.4, where we investigate cleaned CMB maps, we used maps produced by Planck’s four different cleaning methods: Commander, SMICA, SEVEM and NILC. These four methods approach the task of foreground cleaning from four different approaches that range from data orientated, minimal assumption methods, such as NILC and SEVEM, to parametric methods, such as Commander. For the purpose of this work it is sufficient to note that the different assumptions, approaches and spaces (i.e. wavelet, pixel and harmonic space) in which these methods work mean that these methods are a good representation of all the different cleaning methods available in the literature. For more detailed information on the foreground separation methods we refer the reader to Planck Collaboration X (2016) for a general overview, Delabrouille et al. (2009) for NILC, Fernández-Cobos et al. (2012) for SEVEM and Cardoso et al. (2008) for SMICA.
5.3.2 Pipeline

Before estimating the bispectrum from the maps we several pre-processing steps. Firstly we apply a galactic sky mask. Our fiducial mask is the 70% sky mask provided by the *Planck* experiment but in [5.4] we investigate the effect of the sky cut by using the 20%, 40%, 60% and 80% *Planck* masks ([Planck Collaboration I](#) 2016). The mask is apodized with a 1 degree Gaussian to reduce the effect of mode coupling. We apply the mask as without it our results would be entirely dominated by the bright signal from the galactic plane and because we are interested in characterising the foregrounds in sky regions of interest to cosmology. Next we apply the *Planck* point source mask and a second galactic mask that masks the brightest 2% of the sky within in the 70% sky mask. We discuss the motivation behind this mask in section [5.3.3](#) We then iteratively infill it to apodize the mask. [Gruetjen et al.](#) (2015) showed that infilling point sources allowed optimal bispectrum measurements, see [Bucher et al.](#) (2016) for a more detailed discussion of the effect of infilling on the binned bispectrum. We then use the [Reinecke & Seljebotn](#) (2013) and [Górski et al.](#) (2005) packages combined with the pure E/B decomposition method from [Grain et al.](#) (2012) to obtain the spherical harmonic coefficients. Finally as the maps released as part of the *Planck* 2015 data release have had the modes with $\ell < 30$ in polarisation masked due to residual systematics ([Planck Collaboration VII](#) 2016), we filter our maps for both polarisation and temperature to remove all modes with $\ell < 40$.

5.3.3 Signal Inhomogeneity

We have made three suboptimal analysis choices: our low $\ell$ cut in temperature, the second dust mask described in section [5.3.2](#) and choosing to use the 353 GHz maps rather than the Commander dust map, which should be a cleaner map of the dust. These choices were necessary as without them we found we were unable to accurately model the estimator’s variance. The dust is highly anisotropic and thus contributes
Figure 5.1: The average 1-D bispectrum of 100 realizations of a Poisson random field compared with the theoretical value.

Figure 5.2: The average 1-D chi-squared of 100 realizations of a Poisson noise and 100 realizations of a Gaussian CMB.

significantly to the linear term, see 5.2.3. Currently we are not able to calculate the dust’s contributions to the linear term as this requires simulating the anisotropic distribution of the dust. Accurately calculating this term is the subject of on-going work and is necessary to obtain optimal results. The 2% dust mask was chosen as it was found that with this mask, and the low ℓ cut, we can model the variance of our estimator to with ∼ 10%.

5.4 Foreground Bispectra

In this work we use two projections of the data for ease of visualisation. They are the one dimensional bispectrum

\[
\hat{b}^{1D,X,Y,Z}_l = \sum_{3\ell_i^2 < \ell_j^2 + \ell_k^2 \leq 3\ell_{l+1}} \hat{b}^{X,Y,Z}_{i,j,k}
\]  

(5.25)
Figure 5.3: The 353 GHz TT, EE and BB pseudo-Cl power-spectrum obtained from 70% of the sky.

Figure 5.4: The parity even 1-D bispectrum and 1-D chi-squared for the 353 GHz map. We used 70% of the sky for these measurements. The dotted lines represent negative values of the bispectrum.

and the one dimensional chi-squared

\[
\chi^2_{1D} = \sum_{3\ell_i^2 + \ell_j^2 + \ell_k^2 \leq 3\ell_{t+1}} \hat{b}^X_{i,j,k} (V^{X,Y,Z,X,Y,Z})^{-1}_{i,j,k} \hat{b}^{X,Y,Z}_{i,j,k}.
\]  

(5.26)
Figure 5.5: 1-D chi-squared from half-ring difference maps for 70% of the sky.

Figure 5.6: EET (solid lines) and BBT (dotted lines) 1-D 353 GHz bispectra for different fractions of the sky.

Figure 5.7: Equilateral slice of 353 GHz bispectrum 70% of the sky.

Figure 5.8: The parity odd 1-D chi-squared for the 353 GHz maps. 70% of the sky was used for these measurements.
While the 1-D bispectrum is useful for visualisation purposes, it can obscure the signal as the variance for each bin is set by the noisiest bin included in the sum. The one dimensional $\chi^2$ is based on the Gaussian covariance and assumes that the bispectra bins are independent. The application of a mask introduces mode coupling between the bispectra bins violating the independence assumptions, however we choose the width of our bins such that the bin to bin coupling constrained to be less than 5% such that this assumption is approximately true. For the largest mask, $f_{\text{sky}} = 80\%$, we use bin widths of $\Delta \ell = 25$ and for the smallest mask, $f_{\text{sky}} = 20\%$, we use $\Delta \ell = 45$. Note that we deconvolve the beams for all the bispectra presented in this work.

Before presenting our results, we first present a simple application of the 1-D statistics presented above in order to aid the interpretation of our results and as a simple verification of our pipeline. We apply these 1-D statistics to a set of 100s maps with Poisson point sources and a set of 100 Gaussian realizations of the CMB. In figure 5.1 we plot the average 1D bispectrum from the 100 Poisson maps and we plot the theoretical expectation (which is just a constant). In figure 5.2 we plot the average 1D chi-squared for the 100 Poisson maps and for the 100 Gaussian maps. As expected the Gaussian maps $\chi^2$ is consistent with 1 across all the bins, which says there is no evidence of non-Gaussianity. This is not the case for the Poisson maps which deviate significantly from 1.

### 5.4.1 353 GHz Bispectra

It is useful to first review the 353 GHz power spectrum. In figure 5.3 we plot TT, EE and BB pseudo-Cl power-spectra for the 353 GHz maps. We see that the BB mode power is a factor of $\sim 2$ less than the EE power and than the power spectra are well described by power laws $C_\ell \propto \ell^\alpha$ with $\alpha \sim -2.5$ for the all three spectra. Now we move to the bispectrum results. In figure 5.4a we present the 1-D parity even binned bispectrum of the Planck 353 GHz maps. In this projection, we see a strong signal
for several bispectra combinations, notably TTT, BTT, ETT, BBT, EET and EEE. The remaining combinations seem to show no clear signal in figure 5.4a, however by examining the 1-D parity even chi-squared, shown in figure 5.4b, we see that this is a projection effect. From the 1-D chi-squared we see that all the map combinations show a non-zero signal.

Having observed non-zero bispectrum it is necessary to verify that is a sky signal. To do this we use the difference of the two-half ring data splits Planck Collaboration VIII (2016). This combination should have no signal and only noise. In figure 5.5 we plot the 1-D chi-squared for the noise splits. We see that no strong evidence of non-Gaussianity in any of the splits. The TTT noise split shows weak evidence of non-Gaussianity. It is the subject of on-going work to validate the origin of this excess. The significance is not high compared to the level of the 353 GHz and so we assume this deviation is not important for our results. Secondly we wish to verify this is galactic dust and not another source of non-Gaussianity, such as extra-galactic point sources Argüeso et al., 2003, Lacasa et al., 2014, Crawford et al., 2014, Coulton et al., 2017. To do this we plot the signal as a function of masked sky. The masks are constructed from the 353 GHz maps such that the cleanest parts of the sky are selected, thus as we reduce the used fraction of sky we should see a reduction in the size of the bispectrum signal. In figure 5.6 we plot the BBT and EET bispectrum signals as a function of the sky cut and find that as the fraction of the sky is reduced the signals are reduced in amplitude.

In Planck Collaboration Int. XXX (2016) it was found that the dust EE power was twice that of the BB power. We explored whether this was the case for bispectrum where we can examine the EEE and BBB auto-bispectra and their cross correlations with temperature. In figure 5.4a we see that the EEE parity even bispectrum is significantly larger (at the largest scales by a factor of 10) than the parity even BBB bispectrum. In 5.6 we plot the BBT and EET parity even bispectra cross-correlations.
for different sky cuts. We find that the amplitude of the EET bispectrum is greater than the BBT bispectrum for all sky cuts by a factor of 2. It is of interest to examine the ETT and BTT bispectra. We find a non zero parity even BTT bispectrum. This is unusual, but not expected given the existence of the BT power-spectrum signal. As expected this BTT signal is significantly weaker than the ETT bispectrum at all scales. The dust power spectra are well characterized by a power laws (Planck Collaboration et al., 2018) \( C_\ell \propto \ell^\alpha \) with \( \alpha \sim -2.5 \) and we explore the scale dependence of the bispectrum in figure 5.7 where we plot the equilateral slice of the dust bispectrum. The 1-D bispectrum plotted elsewhere merges many different configurations and can hide the scale dependence of the bispectrum and thus we used the equilateral slice to probe the scale dependence. We see that the TTT and EEE bispectra are roughly described by \( b(\ell, \ell, \ell) \propto \ell^\alpha \) where \( \alpha \sim -4 \). The other configurations do not have sufficient signal to noise in the equilateral configuration to be characterized.

Finally, we then measured the parity odd 353 GHz bispectrum and the resulting 1-D chi-squared are shown in figure 5.8. We find measurable levels of parity odd non-Gaussianity for many of the bispectrum configurations, most significantly the BET, BEE and BTT. It is unexpected to see non-zero levels of parity odd BBE bispectrum as at the power-spectrum level Planck Collaboration et al. (2018) found no evidence of EB correlations.

### 5.4.2 Synchrotron Bispectra

In figures 5.9a and 5.9b we present the 1-D parity even binned bispectrum and 1-D chi-squared of the Planck commander synchrotron maps. We find that there is strong evidence for the TTT bispectrum and weaker evidence for ETT and slight evidence for BTT bispectra. These bispectra are curious as they show an increase on smaller scales. It is possible that this arises from the filaments seen in the synchrotron maps. However there are half mission splits provided for the Commander synchroton
Figure 5.9: The parity even 1-D bispectrum and 1-D chi-squared for the commander synchrotron map. We used 70% of the sky for these measurements. The dotted lines represent negative values of the bispectrum.

Figure 5.10: The 1-D chi-squared for 353 GHz -synchrotron B mode cross correlations. The solid lines are for the parity even estimator and the dotted lines are for the parity odd case.
353GHz T and B map cross lower frequency B map parity-even bispectrum

353GHz T and E cross lower frequency B map parity-odd bispectrum

Figure 5.11: The 1-D chi-squared for the cross bispectra between the 353 GHz maps and the Planck 100, 143 and 217 GHz B mode map for 70% of the sky. The solid lines are the signal and the dashed lines are noise maps constructed from half-ring maps.

Parity Even

Parity Odd

Figure 5.12: The 1-D chi-squared for the cross bispectra between the 353 GHz T, E and B map and the SMICA B mode map for 70% of the sky
Figure 5.13: The 1-D chi-squared for the cross bispectra between the 353 GHz T, E and B map and the B mode map from the different cleaning maps for 70% of the sky.

(a) Parity even cross-bispectrum between the 353 GHz B and E maps and the component separated B map

(b) Parity odd cross-bispectrum between two 353 GHz B maps and the component separated B map

Figure 5.14: The 1-D chi-squared for parity even cross bispectra between 353 GHz T, E and B maps and SMICA T and E maps on 70% of the sky.

(a) SMICA T map

(b) SMICA E map
maps so it is hard to validate that this is not a result of non-Gaussian noise in the low frequency maps. The polarised synchrotron maps are significantly noisier than the dust polarization maps and this limits our ability to probe the synchrotron bispectrum. We find no evidence for any odd-parity non-Gaussianity.

5.4.3 Dust - Synchrotron Cross Bispectra

In Choi & Page (2015) they studied the correlation between the dust and synchrotron B mode signals. Physically this correlation arises as both effects are influenced by the galactic magnetic field. Analogously to that we investigated the dust-synchrotron B mode bispectrum. We find that the dust-synchrotron B mode bispectrum is consistent with zero for all of the masks except the 80% sky mask. The 1-D chi-squared for the parity even and parity odd dust-synchrotron B mode bispectra are shown in figure 5.10. We find weak evidence for a 353-synchrotron bispectrum. With lower noise maps, particularly the synchrotron B map, we hope to improve the measurement of this bispectrum.

5.4.4 Cleaned Map Bispectra

Motivated by the signals seen above, we investigated whether bispectra can be used to search for residual foregrounds. Before considering foreground cleaned maps it is necessary to see if we see a bispectrum in the Planck single frequency maps. The bispectrum in these maps will be significantly reduced when compared to the results in section 5.4.1. In figures 5.11a and 5.11b we examined the 1-D chi-squared for cross bispectra between the 353 GHz T, E and B and the B mode maps from the 100, 143 and 217 GHz channels. Using the strong foregrounds in the 353 GHz maps we look for correlations in the B map. We find strong evidence for the parity even BTB bispectra in these maps and evidence for odd parity TEB bispectrum at 143 and 217 GHz, with a hint of a signal in the 100GHz map.
The strong signal seen in 5.11a and 5.11b means that the bispectrum could be used to test for residual foregrounds. In figures 5.12a and 5.12b we plot the cross bispectra between the 353 GHz T,E and B maps with the SMICA B mode map. As can be seen there is no strong evidence of a residual bispectrum. There is a slight, 3.1σ, low ℓ excess above the Gaussian level in the parity even BBE combination and low ℓ excess in the parity odd BBB bispectrum. We explore this more in figures 5.13a and 5.13b where we plot the BBE parity even and BBB parity odd bispectrum for the four different component separation methods. We find this excess, with varying strengths, in all the foreground cleaning methods. This signal is not seen in the half ring splits.

Similarly we can test for residuals foregrounds in the temperature and polarisation maps and the results are shown in figures 5.14a and 5.14b. In figure 5.14a we see a large residual foreground in TEE combination, however when we examine the same configuration replacing the T map with the T noise split map we still see this signal. This implies that the single is arising from the noise. There are two possibilities: that the noise is non-Gaussian or the noise is very anisotropic and not accounted for with our linear term. Our linear term calculation modelled the anisotropy of the noise by modulating the noise maps with a square root of the hits maps. The component separated maps weight the different input maps very differently across the maps and it could be this introduces anisotropy which we do not capture. To identify this we need to process noise simulations through the component separation pipelines.

5.5 Discussion and Conclusions

The results presented in section 5.4 suggest that there are large levels of non-Gaussianity in the galactic foregrounds, particularly the galactic dust. It also suggests that these bispectra could be useful for cross checking residual foregrounds
in cleaned temperature maps, particularly as lower r constraints increase the difficulty of foreground cleaning (see e.g. Remazeilles, 2018). Whilst we found no strong evidence for residual foreground non-Gaussianity in the component separated maps, we found a hint of a residual. To further investigate this we need be able to properly treat the linear term. This includes modeling the anisotropy of the galactic dust and processing noise simulations through the component separation pipelines.

In this paper we have focused on characterizing the bispectra signals and discussing the utility of the bispectrum to constrain residual foregrounds, however there is far more information available. Planck Collaboration Int. XIX (2015) examined the physical origins of the polarized dust signal. Bispectrum measurements have the can also be used to study the physical properties of the dust and complement the results of Planck Collaboration Int. XIX (2015). In particular, in Burkhart et al. (2009) they demonstrate first using simulations, and then in Burkhart et al. (2010) using measurements of the Small Magellanic Cloud, that column density bispectra can be used to constrain Alfvén and sonic Mach number. Thus in the future bispectrum measurements can be used to constrain the magnetic and kinetic properties of the interstellar medium.
Chapter 6

Conclusion

What can we conclude from the diverse range of results presented in this thesis? Firstly that the bispectrum is a very flexible tool; we have used it to study the physics of the very early universe, the large scale structure of the universe and the properties of the galaxy. Further these bispectrum measurements are very complementary to existing power-spectrum measurements, providing new information with which we can further enhance our understanding of the universe.

Secondly, as two of the projects presented here demonstrate the proof of principle of the techniques, there is clearly lots of interesting directions for future work. These directions include further developing the bispectrum methods to efficiently extract cosmological information, developing theoretical models of primordial tensor signals, applying these techniques to larger and more precise data sets and many more avenues.

Finally and more broadly, current and upcoming data sets will allow use to make high precision tests of our cosmological picture. To test Λ CDM as thoroughly as possible, we will increasingly need to go beyond the power spectrum and the bispectrum is one, of many possible, approaches to this.
Bibliography

Alpher, R. A. & Herman, R. C. 1948, Phys. Rev., 74, 1737
Bernardeau, F. 1997, Astron. Astrophys, 382, 10


Buhmann, M. D. & Buhmann, M. D. 2003, Radial Basis Functions (New York, NY, USA: Cambridge University Press)


Einstein, A. 1915, Ständiger Beobachter der Preussischen Akademie der Wissenschaften, Part 2, 844

Einstein, A. 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152.


Friedman, A. 1922, Zeitschrift für Physik, 10, 377

Friedmann, A. 1924, Zeitschrift für Physik, 21, 326


Grishchuk, L. P. 1975, Soviet Journal of Experimental and Theoretical Physics, 40, 409


Ichiki, K. 2014, Progress of Theoretical and Experimental Physics, 2014, 06B109


Kacprzak, T., et al. 2016, Not. R. Astron. Soc, 000, 0


155


Mishra-Sharma, S., Alonso, D., & Dunkley, J. 2018


Namikawa, T. 2016, Physical Review D, 93, 121301


Polnarev, A. G. 1985, Soviet Ast., 29, 607


Shiraishi, M. 2013, Probing the Early Universe with the CMB Scalar, Vector and Tensor Bispectrum

160


Shiraishi, M., Nitta, D., & Yokoyama, S. 2011, arXiv:1108.0175, Progress of Theoretical Physics, 126, 937


Weinberg, S. 2008, Cosmology, Cosmology (OUP Oxford)


