University of Cardiff

PhD Thesis

The investigation, development and testing of novel methods for the statistical characterisation of:

- cosmic microwave background data, aimed at isolating and quantifying departures from the standard cosmological model; and,
- large scale galaxy clustering data, aimed at refining estimates of key parameters required for the advancement of galaxy formation theory.

Jo Short MSci

2011

Thesis Summary and Declaration/Statements

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.
Signed: .... Jo Short .... Date: .... 30 September 2011 ....

Statement 1

This thesis is being submitted in the partial fulfillment of the requirements for the degree of PhD.
Signed: .... Jo Short .... Date: .... 30 September 2011 ....

Statement 2

This thesis is the result of my own independant work/investigation, except where otherwise stated. Other sources are acknowledged giving explicit references.
Signed: .... Jo Short .... Date: .... 30 September 2011 ....

Statement 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be available to outside organisations.
Signed: .... Jo Short .... Date: .... 30 September 2011 ....

Statement 4: Previously Appoved Bar on Access

I hearby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee**
Signed: .... Jo Short .... Date: .... 30 September 2011 ....

Contents

Acknowledgements **xiii**

Abstract **xiv**

1 Introduction to Cosmology **1**

1.1 Cosmological Principles 2

1.2 Hubble's Law 2

1.2.1 Redshift 3

1.2.2 Comoving coordinates 4

1.3 Equation of state 5

1.3.1 Perfect Fluid 5

1.3.2 Relativistic fluid 5

1.4 Einstein's General Relativity 6

1.4.1 Stress-Energy Tensor 7

1.4.2 Friedmann-Robertson-Walker Metric 7

1.4.3 Friedmann Equations 8

1.4.4 Critical Density 8

1.4.5 Equations of motion 9

1.4.6 Newtonian Limit 10

1.5 Inflation 10

1.5.1 Fluctuation type 12

1.6 Matter and Energy 12

1.6.1 'Normal' Matter 12
1.6.2 Dark Matter 12
1.6.3 Dark Energy 13
1.6.4 Evolution of matter/energy densities 13
1.7 Testing the concordance cosmological model 14

I Cosmic Microwave Background 16

2 Introduction to the CMB 17
2.1 Standard Concordance Cosmological model 17
2.1.1 Nature of the CMB Temperature Fluctuations 18
2.1.2 Evolution of the CMB Temperature Fluctuations 19
2.1.3 Statistics of the CMB Temperature Fluctuations 20
2.2 CMB Discoveries 21
2.2.1 Black Body Temperature 23
2.2.2 Temperature Anisotropies 24
2.2.3 CMB Power Spectrum 25
2.3 Ongoing and Future CMB Studies 25
2.3.1 Wavelets and the Cold Spot 26
2.3.2 Phase Analysis 27
2.3.3 Multipole Vectors 28
2.3.4 Polyspectra 28

3 Zonal Modes of CMB Maps 30
3.1 Defining Zonal Modes 31
3.1.1 Spherical Harmonics 31
3.1.2 Zonal and Sectoral Modes 32
3.1.3 Non-Gaussianity of Zonal Modes 33
3.2 Statistical Analysis with Zonal Modes 34
3.2.1 WMAP CMB Maps 35
3.2.2 Extremal Mode Counts 37

3.2.3 Variance of Amplitudes for Zonal Modes 42
3.2.4 Zonal CMB Maps 43

4 CMB Maps from Anisotropic Cosmologies 47
4.1 Phase correlations of Bianchi CMB Maps 51
4.1.1 More about Spherical Harmonics 51
4.1.2 Visualising Bianchi Phase Correlations 53
4.1.3 Rotating maps and adding noise 59
4.1.4 Application to WMAP 5 Year Data 64
4.2 Cross Correlation of Bianchi CMB Maps 67
4.2.1 Defining the Spherical Harmonic Cross Correlation 68
4.2.2 Application to Bianchi Simulations of the CMB 69
4.2.3 Applications to Gaussian CMB Maps with Localised Spots 79
4.2.4 Application to ILC processed Gaussian simulations 79
4.2.5 Application to WMAP CMB Data 81

5 Conclusions for CMB Analyses 84

II Galaxy Clustering 90

6 Introduction to Galaxy Clustering 91
6.1 Standard Concordance Cosmological model 92
6.1.1 Background cosmology 93
6.1.2 Initial fluctuation spectrum 93
6.1.3 Dark Matter 93
6.1.4 Growth law 93
6.1.5 Transfer function 95
6.1.6 Growth of linear fluctuations 96
6.1.7 Non-linear evolution 98
6.1.8 Formation of galaxies 98
6.2 Measures of Galaxy Clustering 100
6.2.1 Modeling methodologies 100

6.2.2 Observational Data 102
6.2.3 Statistics of Galaxy Clustering 102
6.2.4 Cross correlation 104

7 Analytical Angular Correlation models 106
7.1 Linear galaxy angular correlation function 107
7.1.1 Initial linear matter power spectrum 107
7.1.2 Transfer function 107
7.1.3 Growing mode of linear perturbations 110
7.1.4 Linear dark matter power spectrum 110
7.1.5 Linear dark matter angular correlation function 111
7.1.6 Linear galaxy angular correlation function 112
7.2 Non-linear galaxy angular correlation (Fitting function) 114
7.3 Non-linear galaxy angular correlation (Halo model) 115
7.3.1 Limits and normalisation factors 116
7.3.2 Normalised Fourier transform of the halo density profile 117
7.3.3 Mass function 119
7.3.4 Scale dependent bias 119
7.3.5 Halo Occupation Distribution model 121
7.3.6 Non-linear Galaxy angular correlation 121

8 *Herschel* observations of galaxy clustering 123
8.1 *Herschel* Space Observatory 123
8.1.1 *Herschel* multi-tiered extragalactic survey 125
8.1.2 *Herschel* astrophysical terahertz large area survey 125
8.2 Model parameters 126
8.3 Analytical models compared to HerMES observations 128
8.3.1 Observations compared to a power law 130
8.3.2 Observations compared to a linear model 130
8.3.3 Observations compared to a non-linear fitting function 133
8.3.4 Observations compared to a non-linear Halo model 133

8.3.5 Estimates of the minimum halo mass 135
8.4 Analytical models compared to H-ATLAS observations 135

9 Investigating Galaxy Clustering Models 138
9.1 Investigating effect of variations in $\mathcal{N}(z)$ 138
9.2 Investigating effect of variations in M_{min} 142
9.2.1 Smooth cutoff at M_{min} 142

10 Conclusions for Galaxy Clustering 149

11 Discussion 154

List of Figures

1.1 3-D examples of closed, open, and flat geometries. 9

1.2 Evolution of the Universe's radiation, matter and dark energy densities since the Big Bang. 14

2.1 Intensity plot of the CMB as measured by COBE. 24

2.2 Temperature maps of the CMB as measured by COBE and WMAP. 25

2.3 The CMB temperature power spectrum from WMAP. 26

3.1 Illustrative examples of spherical harmonics for individual modes. 33

3.2 Illustrative examples of zonal and sectoral modes. 33

3.3 CMB maps generated via different cleaning methods from WMAP observations. 36

3.4 Reconstructed CMB maps using only the zonal ($m = 0$) modes. . 44

4.1 Simulated maps of the the CMB temperature using Bianchi type cosmologies. 50

4.2 Illustrative examples of spherical harmonics for ranges of modes. 52

4.3 Illustrative example of spherical harmonics when $\ell = 0 - 20$, $m = 2$, and $\Delta\Phi_{\ell m} = \pi/8$. 53

4.4 Example of the spherical harmonic phases ($\Phi_{\ell m}$) for the concordance model where ℓ, $m \in [0, 20]$. 54

4.5 Phases of the spherical harmonic coefficients ($\Phi_{\ell m}$) for Bianchi types V, VII_0 and VII_h. 56

4.6 Phases differences of the spherical harmonic coefficients ($\Delta\Phi_{\ell m}$) for Bianchi types V, VII$_0$ and VII$_h$. 57
4.7 Simulated maps of the the CMB temperature which are rotated with fluctuation noise added. 61
4.8 $\Delta\Phi_{\ell m}$ for Bianchi type V map at $z = 0$ with white, WMAP, and concordance model fluctuations noise maps. 63
4.9 $\Delta\Phi_{\ell m}$ for the 5 year ILC map with the axis of evil aligned with the preferred axis compared with random phases. 65
4.10 Cross correlation, $K_{kj}(\ell)$, for each of the Bianchi models at redshift $z = 0$. 70
4.11 Cross correlation, $K_{kj}(\ell)$, for the Bianchi VII$_h$ model (tilted to align with the cold spot). 71
4.12 Average K_{10} for each of the Bianchi models combined with Gaussianly distributed fluctuations at redshift $z = 0$. 73
4.13 Average K_{10} for the Bianchi V model combined with Gaussianly distributed fluctuations at redshifts $z = 2$, 6, and 30. 75
4.14 Average K_{10} for the Bianchi VII$_0$ model combined with Gaussianly distributed fluctuations at redshifts $z = 2$, 6, and 30. . . . 76
4.15 Average K_{10} for the Bianchi VII$_h$ model combined with Gaussianly distributed fluctuations at redshifts $z = 2$, 6, and 30. . . . 77
4.16 Average power spectrum of Bianchi CMB maps, with Gaussianly distributed fluctuations, at $z = 0$. 78
4.17 Example CMB map of a concordance model simulation with a cold spot added in the bottom-right. 79
4.18 Average normalised cross correlation, $K_{kj}(\ell)$, for 10000 concordance model simulations with a cold spot added. 80
4.19 Average cross correlation, $K_{kj}(\ell)$, for 10000 Gaussian simulations with ILC processing. 81
4.20 The cross correlation, $K_{kj}(\ell)$, for the 5 year WMAP ILC map. . 82

4.21 The cross correlation, K_{10}, for an individual realisation of a Bianchi VII$_h$ model combined with Gaussianly distributed fluctuations at redshift $z = 0$. 83

6.1 Observed distribution of galaxies from the 2dF Galaxy Redshift Survey. 91
6.2 ILC temperature map of the CMB as measured by WMAP. . . . 92
6.3 Snapshots from the Millennium (numerical) simulation of the matter density at different redshifts. 97

7.1 The differential number of halos as a function of mass using the Sheth-Tormen approximation. 120

8.1 An example of a field of galaxies as imaged by SPIRE (on-board *Herschel*) at 250, 350 and 500 μm. 125
8.2 The normalised distribution of galaxies as a function of redshift for data sets from HerMES and H-ATLAS. 127
8.3 The angular correlation function $\omega_{\rm obs}(\theta)$ from the HerMES survey compared to the Halo model (Cooray et al., 2010). 129
8.4 The angular correlation function $\omega_{\rm obs}(\theta)$ from the HerMES survey compared to a power law. 131
8.5 The angular correlation function $\omega_{\rm obs}(\theta)$ from the HerMES survey compared to a linear approximation. 132
8.6 The angular correlation function $\omega_{\rm obs}(\theta)$ from the HerMES survey compared to two non-linear approximations. 134
8.7 The angular correlation function $\omega_{\rm obs}(\theta)$ from the H-ATLAS survey compared to the fitting function method. 137

9.1 $\omega_{\rm obs}(\theta)$ for a selection of Gaussianly distributed $\mathcal{N}(z)$ with varying μ. 140
9.2 $\omega_{\rm obs}(\theta)$ for a selection of Gaussianly distributed $\mathcal{N}(z)$ with varying σ. 141

9.3 $\omega_{\mathrm{obs}}(\theta)$ for a selection of fixed M_{min} 143
9.4 $\omega_{\mathrm{obs}}(\theta)$ for a Gaussianly distributed M_{min} with $\mu = 13.4$ and $\sigma = 0.5$. 145
9.5 $\omega_{\mathrm{obs}}(\theta)$ for a Gaussianly distributed M_{min} with $\mu = 13.4$ and $\sigma = 1.0$. 146
9.6 The effective bias for a range of (fixed) M_{min} 147

List of Tables

3.1 Number of times for which the minimum (or maximum) amplitude is a zonal mode or a sectoral mode. 38

3.2 Monte Carlo estimates of the probabilities of the extremal mode counts for the various CMB maps in Galactic coordinates. 40

3.3 Monte Carlo estimates of the probabilities of the extremal mode counts for the various CMB maps in Ecliptic coordinates. 41

3.4 Percentage of simulated CMB maps which have a variance of the $x_{\ell m}$ which is greater than the specified map. 43

3.5 Probability of the observed maxima and minima in zonal maps, as derived from Monte Carlo simulations. 45

4.1 Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the Bianchi CMB maps with a random distribution of phases. 58

4.2 Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the Bianchi CMB maps, rotated by $\theta = \pi/8$, with a random distribution of phase differences. 60

4.3 Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the Bianchi CMB maps rotated by $\theta = \pi/8$ with white, WMAP, and concordance model noise maps. 62

4.4 Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the WMAP ILC map with a random distribution of phase differences 64

7.1 Cosmological parameters used as inputs to the angular correlation models, along with current estimate from WMAP. 111

8.1 Best fit parameters for the Halo and fitting function models for the HerMES results. 135

Acknowledgements

Firstly I would like to thank my supervisor Prof. Peter Coles to whom I am very grateful for his support throughout my PhD. I would also like to thank Prof. Pavel Naselsky, Prof. Steve Eales, Dr Rockhee Sung, Dr Steve Maddox and Dr Loretta Dunne for interesting and helpful discussions. I am appreciative of the financial support I have received from the Science & Technology Facilities Council (STFC) doctoral training grant. Finally I would like to thank family and friends for their continued encouragement.

Abstract

This thesis concerns the statistical characterisation of the large scale properties of our Universe. Two complementary data sets are considered: all-sky maps of the cosmic microwave background (CMB) temperature fluctuations from the Wilkinson Microwave Anisotropy Probe (WMAP); and large area maps of galaxies detected through the sub-millimetre electromagnetic emission using the *Herschel* Space Observatory.

The CMB temperature field has been extensively studied; one remaining task is to check that the distribution of temperature fluctuations is Gaussian, homogeneous and isotropic as predicted by the standard model. Since these perturbations could deviate from the standard model in many different ways, a number of complementary approaches to investigate the data are required.

All-sky maps of the CMB are often decomposed into spherical harmonic modes. Any modes aligned with the Galactic plane are particularly interesting because any anomalous behaviour in them could indicate errors in the subtraction of Galactic foreground. Here a simple statistical analysis of these modes is tested and shown to be a useful diagnostic of possible foreground subtraction systematics.

In addition, two methods for characterising patterns in all sky CMB maps resulting from anisotropic universes are discussed. The methods are tested against simulated CMB maps from anisotropic cosmologies and both show promise as effective diagnostics for identifying global asymmetry in the Universe.

Whilst many anomalous features have already been identified with WMAP data, they have only been found at confidence levels below 3σ (equivalent to 99%). Soon the next generation of CMB data from the recently launched Planck spacecraft will be available. Testing descriptors against this new data will finally allow analyses to probe higher statistical significances with which to hopefully confirm or reject the various anomalous claims.

The second part of this thesis concerns analytical models of the correlation function for the distribution of galaxies. The standard 'Halo' model is comprehensive, in that it considers separately all the components expected to contribute to the clustering of galaxies, but it is also complex and rather unwieldy.

Here we promote a simpler alternative based on fitting functions found from numerical simulations. Both models are shown to compare well to each other, and the observational data, thereby showing that the fitting function method can be a quick and easy option. Also, we show that some of the apparently 'key' Halo model assumptions, such as intra-halo correlations, are not required to produce a good fit.

We summarise by discussing the different approximations used in the current galaxy clustering models, and the limits of the currently available data. We consider the components which have the most impact on the resulting accuracy, such as uncertainty surrounding the redshift distribution of the sources. Finally we look at future areas of development.

Chapter 1

Introduction to Cosmology

Science is the systematic study of the structure and behaviour of processes and phenomena. Scientists propose hypotheses, based on the limited available evidence, which are subsequently investigated and tested against observable quantities and via experimentation. This methodical process of posing questions, suggesting potential solutions, and testing them against the available data has enabled successive generations to achieve an increasingly comprehensive and reliable understanding of the world in which we live.

Cosmology is the science of the origin and evolution of our Universe. Experimentation in cosmology can be difficult to implement. Some experiments can be done on the very small scales, but clearly not on large scales. Simulations are often used to try to bridge this gap, but they still need to be tested against data. Observations of our Universe, on all scales, are therefore crucial. Nevertheless, this also has its drawbacks. We live in, and can therefore only observe, one Universe. This can makes things difficult when you are trying to derive statistical conclusions about cosmology on Universal scales.

The current standard cosmological model is known as the *concordance* model, because it consists of many pieces of evidence from different sources (e.g. the cosmic microwave background, supernovae, and large scale structure) that all independently point to one set of assumptions being true. To provide a background to the work in this thesis, in this introduction we discuss the theoretical framework for the concordance model including highlighting some of the assumptions and a selection of the corresponding sets of evidence.

1.1 Cosmological Principles

The *Copernican principle* says that 'we' are not in a special place. In other words, given that there is no evidence to the contrary, the simplest assumption we can make is that our local Galactic neighbourhood is a good representation of our Universe as a whole.

The *Cosmological principle* follows on from this to say that on large scales our Universe is *homogeneous* (invariant to translation) and *isotropic* (invariant to rotation).

Both of these statements are 'principles' meaning they are not fact but reasonable assumptions given that we need to start somewhere and there is currently no strong evidence to the contrary.

To quote another principle, *Occam's razor* advocates choosing the hypothesis which makes the least new assumptions when the competing hypotheses are otherwise equivalent. However this still requires research into potential alternatives to assess their viability. For example later in this work we will consider a universe which is not isotropic and therefore breaks this Cosmological principle.

1.2 Hubble's Law

We know from just looking up into the night sky that our Universe is not totally 'smooth'. There are very dense areas, such as stars and galaxies, and in between there is rarefied gas. So our assumption that the Universe is homogeneous and isotropic must only be valid on large scales i.e. scales much larger than the size of a galaxy.

However, if our Universe is the same everywhere then there must infinitely many stars. If this were the case, then whatever direction we look in the night sky there should eventually be a star meaning that the sky would look white not black. This is *Olbers' paradox* and is one piece of evidence that points towards a *non-static universe* i.e. our Universe must be evolving.

The discovery of the expansion of our Universe is credited to Edwin Hubble (Hubble, 1929). He used observations of *variable*[1] stars in nearby galaxies to show that on average all these objects have a global motion away from us. *Hubble's Law* states that the velocity (v) at which an object is receding from us is proportional to their distance away from us (r) i.e.

$$v = H(z)r. \tag{1.1}$$

The Hubble Space Telescope Key project (Freedman et al., 2001) made a much more comprehensive study of these variable stars to extract a current value for the *Hubble constant* of $H_0 = 72 \pm 8$ km s^{-1} Mpc^{-1}.

Initially it was assumed that the expansion of the Universe was slowing down due to the action of its self gravity i.e. $H(z)$ was decreasing over time. However in 1998, two teams studying distant (type Ia) supernovae (Supernova Cosmology Project and the High-z SN Search Team) presented independent evidence that the expansion of the Universe is speeding up (Riess et al., 1998; Perlmutter et al., 1999).

1.2.1 Redshift

If an object is moving away from an observer then the wavelength of the photons it emits are stretched so that they appear more red. This is known as *Doppler redshift*. Alternatively, objects moving towards an observer appear more blue (known as *blue-shift*).

Similarly, *gravitational redshift* occurs when photons lose energy to climb out of a gravitational field, and *cosmological redshift* is caused as photons traverse

---

[1] Cepheid variable stars have regular variations in their luminosity which are directly related to the period of this pulsation.

expanding space. The amount of redshift (blue-shift) is calculated as,

$$z = \frac{\lambda_{\text{obsv}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}}. \tag{1.2}$$

1.2.2 Comoving coordinates

Because the Universe is expanding, the physical (*proper*) distances (x) between objects are always changing. It is therefore useful to introduce a *comoving coordinate* (χ) system to describe cosmological distances. This comoving coordinate remains constant in time (t), and a *scale factor* (a) is introduced to accommodate the actual physical change in distance,

$$x(t) = a(t)\chi. \tag{1.3}$$

So applying this to the cosmological redshift on the wavelength of a photon then,

$$\lambda_{\text{emit}} = \Delta x(t_{\text{emit}}) = a(t_{\text{emit}})\Delta\chi,$$

and

$$\lambda_{\text{obsv}} = \Delta x(t_{\text{obsv}}) = a(t_{\text{obsv}})\Delta\chi.$$

These expressions can then be substituted into Equation 1.2 to describe the scale factor in terms of a redshift i.e.

$$z = \frac{a(t_{\text{obsv}})}{a(t_{\text{emit}})} - 1,$$

or more generally $a(t_{\text{obsv}})$ or a_0 is normalised to 1, so this is written as,

$$a(t) = \frac{1}{1+z}. \tag{1.4}$$

This relates the observable quantity, i.e. the redshift, to the key parameter used in the theoretical General Relativity equations which will be discussed later in Section 1.4.

1.3 Equation of state

The contents of the Universe are thought of as a *fluid* which is characterised by a number of key parameters that describe the state of the fluid for example, pressure, density, temperature, or internal energy. The *equation of state* describes the relationship between all these key parameters.

1.3.1 Perfect Fluid

To simplify the model, and subsequent calculations, this cosmological fluid is often assumed to be a *perfect fluid* which is a fluid that can be completely described by just its energy density ρ and its pressure p i.e. it has no viscosity, and it does not conduct heat. Whilst this sounds like a rather idealised approximation, in certain situations it can be quite realistic. For example, the mean free path between particles is generally much smaller than the large scales considered in cosmology meaning that on these scales viscosity is negligible and conduction of heat is poor as required.

However this still leaves a number of different solutions for $p(\rho)$. We assume here, as is generally done, that the fluid is also an *ideal gas* meaning that the relationship between the pressure and the density is linear i.e.

$$p = p(\rho) = \omega\rho, \tag{1.5}$$

where ω is a constant.

1.3.2 Relativistic fluid

We also consider the special case where the fluid consists of radiation (or any ultra-relativistic particle) which will be useful to consider later in the chapter (see Section 1.6.4). To calculate the equation of state, consider a relativistic fluid where the particles have a kinetic energy (ϵ), a momentum (P), and an average velocity (v). Pressure is defined as the rate of momentum transfer in a given direction through a unit area per unit time. Since the direction of the momentum is randomly distributed in three dimensions, the pressure p is given by,

$$p = \frac{1}{3} \int_0^\infty P(\epsilon) v(\epsilon) n(\epsilon) d\epsilon, \tag{1.6}$$

where $n(\epsilon)d\epsilon$ is the number of particles with energy between ϵ and $\epsilon + d\epsilon$ in a unit volume. Given the fluid is relativistic, then $v \simeq c$ and therefore $P \simeq \epsilon/c$ to give,

$$p \simeq \frac{1}{3} \int_0^\infty \epsilon n(\epsilon) d\epsilon = \frac{1}{3}\rho c^2 \tag{1.7}$$

where ρ is the total energy density.

1.4 Einstein's General Relativity

A further major component of the concordance model is Einstein's theory of *General Relativity.* It provides a description of the space-time geometry, the action of gravity, and the properties of matter/energy. This relation is specified in *Einstein's Field Equations*[2] (Coles and Lucchin, 2002),

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.8}$$

where G is the *gravitational constant*, and c is the *speed of light.* The *Ricci curvature tensor* $R_{\mu\nu}$ and the *scalar curvature* R describe gravity, the *stress-energy tensor* $T_{\mu\nu}$ describes the properties of matter/energy and the *metric tensor* $g_{\mu\nu}$ describes the space-time. Note that ν and μ indices run from 0 to 3, where the first index is a time coordinate and the other three are spacial coordinates.

---

[2] For zero *cosmological constant.*

1.4.1 Stress-Energy Tensor

The *stress-energy tensor* $T_{\mu\nu}$ describes the properties of matter/energy. The components of the stress energy tensor can be given physical interpretations,

$$T_{\mu\nu} = \begin{array}{cc} & \begin{array}{cc} \nu = 0 & \nu = 1, 2, 3 \end{array} \\ \begin{array}{c} \mu = 0 \\ \mu = 1, 2, 3 \end{array} & \begin{pmatrix} \text{energy density} & \text{energy flux} \\ \text{momentum density} & \text{momentum flux} \end{pmatrix} \end{array}. \tag{1.9}$$

For an ideal perfect fluid this is,

$$T_{\mu\nu} = (\rho c^2 + p) u_\mu u_\nu - g_{\mu\nu} p, \tag{1.10}$$

where u_ν is the four-velocity. This can be further simplified in the rest frame ($u_\alpha = 0$ and $u_0 = 1$) to,

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \tag{1.11}$$

1.4.2 Friedmann-Robertson-Walker Metric

The standard metric used to describe the geometry of space-time is the *Friedmann-Robertson-Walker* (FRW) metric which assumes homogeneity and isotropy. The metric tensor takes the form (Coles and Lucchin, 2002),

$$g_{\mu\nu} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -\frac{a(t)^2}{1-kr^2} & 0 & 0 \\ 0 & 0 & -a(t)^2 r^2 & 0 \\ 0 & 0 & 0 & -a(t)^2 r^2 \sin^2\theta \end{pmatrix}, \tag{1.12}$$

where r, θ and ϕ are co-moving spherical polar coordinates, t is cosmological proper time, a is the cosmic scale factor, and k is the spatial curvature which in the later work in this thesis will be assumed to be flat (discussed further in Section 1.4.4).

1.4.3 Friedmann Equations

The solutions to the field equations for a Universe containing a homogeneous and isotropic perfect fluid, with the FRW metric, are known as the *Friedmann equations* (Coles and Lucchin, 2002),

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2} = \frac{8\pi G}{3}\rho, \tag{1.13}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right), \tag{1.14}$$

where the dot represents differentiation with respect to t, p is the pressure and ρ is the energy density.

1.4.4 Critical Density

A very simplified way of understanding the critical density (for flat universes) is that it determines the boundary value between a continually expanding and contracting universe. In other words, a universe with ρ_c has a density which is just sufficient to stop the expansion of space after infinite cosmic time has elapsed.

The critical density can be derived from the first of the Friedman equations (Equation 1.13) as,

$$\rho_c = \frac{3H^2}{8\pi G}, \tag{1.15}$$

where $H = \dot{a}/a$ and $k = 0$. More generically than the explanation above, the *critical density* (ρ_c) describes the overall geometry, or *curvature*, of the Universe. We can see this by rearranging Equation 1.13 to give,

$$k = \frac{\dot{a}^2}{c^2}\left[\frac{\rho}{\rho_c} - 1\right]. \tag{1.16}$$

So if the density is more than ρ_c, the geometry of space is called *closed* ($k = 1$), and if the density is less than ρ_c then the geometry is called *open* ($k = -1$). If the density is equal to ρ_c then the geometry is *flat* ($k = 0$) (see Figure 1.1).

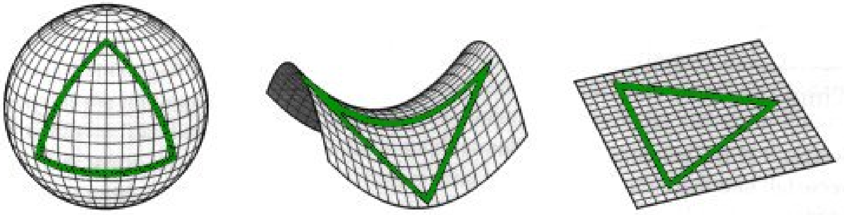

Figure 1.1: *Examples of 3-dimensional equivalents to the three types of 4-dimensional geometry discussed in Section 1.4.4: closed (left), open (centre), and flat (right).*

1.4.5 Equations of motion

The *equations of motion* describe the evolution of the cosmological matter density where energy and momentum are conserved. The stress energy tensor (Section 1.4.1) describes the properties of the matter/energy density and therefore the conservation law requires the derivative of this to be zero. Since it is a tensor we use a *covariant derivative*,

$$T_{\mu\nu;\beta} = T_{\mu\nu,\beta} - \Gamma^{\alpha}_{\mu\beta} T_{\alpha\nu} - \Gamma^{\alpha}_{\nu\beta} T_{\mu\alpha} = 0, \tag{1.17}$$

where $T_{\mu\nu,\beta} = \partial T_{\mu\nu}/\partial \xi_{\beta}$ is an ordinary partial coordinate derivative and the Γ^{d}_{bc} are called *Christoffel symbols* which are related to the metric as,

$$\Gamma^{\alpha}_{\mu\nu} = \frac{1}{2} g^{\beta\alpha} \Big[g_{\mu\beta,\nu} + g_{\beta\nu,\mu} - g_{\nu\mu,\beta} \Big]. \tag{1.18}$$

By considering the different components of the expression in Equation 1.17 we can derive the different equations of motion (Peacock, 1999).

The *continuity equation*, which describes the conservation of energy, can be derived from the $\mu = 0$ component of Equation 1.17 to give,

$$\frac{\partial \rho}{\partial t} + \rho \nabla . \mathbf{v} = 0. \tag{1.19}$$

where $\mathbf{v}$ is the classic 3-dimensional velocity and ∇ denotes the vector differential operator.

The *Euler equation*, which describes the conservation of momentum, can be derived from the $\mu = 1, 2, 3$ component of Equation 1.17 to give,

$$\frac{D\mathbf{v}}{Dt} = -\frac{\nabla p}{\rho} - \nabla\Phi, \tag{1.20}$$

where the convective derivative $\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v}.\nabla$, p is pressure, ρ is density and Φ is the gravitational potential. Note these equations are for the unperturbed case.

1.4.6 Newtonian Limit

If we consider Einstein's equations in the weak gravitational field limit (i.e. $\Phi/c^2 << 1$) then we can retrieve Newton's description of gravity,

$$\nabla^2\Phi = 4\pi G\rho. \tag{1.21}$$

This is *Poisson's equation* and describes the action of gravity for the Newtonian approximation (Peacock, 1999).

1.5 Inflation

Given our Universe is expanding could lead us to surmise that it started from a single point i.e. a *singularity*[3] - the beginning of not just the Universe but space and time itself. This singularity is estimated to have occurred approximately 13.7 billion years ago and has commonly become known as the *Big Bang*. However there are a number of puzzles that are not explained by this Big Bang model. For example, the energy density of the Universe is very close to the *critical density* (Section 1.4.4) which is very unlikely given that by the Copernican principle it could take any value. This is known as the '*flatness problem*'.

To address this, and other issues, *cosmic inflation* was introduced (Guth, 1981; Linde, 2008). Inflation is the accelerating expansion (i.e. $\ddot{a} > 0$) of the early

---

[3] Actually, this is not strictly true; an additional assumption of an *equation of state* for a *perfect fluid* is required else a singularity can be avoided. Also current theories of physics break down at scales below the Planck scale so we can't say much about the Universe at really early times, just that above these scales the Universe was rapidly expanding.

Universe thought to be driven by a vacuum energy density. This solves the 'flatness problem' because the accelerated expansion causes the total matter/energy density Ω_t to tend to the critical density whatever it started out as i.e. $|\Omega_t - 1|$ tends to zero. Another interesting point is that if you apply the condition $\ddot{a} > 0$ to Equation 1.14 then this requires $\left(\rho + \frac{3p}{c^2}\right) < 0$ and since for a perfect ideal fluid ($p = \omega\rho$) this must mean that the pressure is actually negative.

Inflation is generally interpreted as an isotropic *scalar field*[4], ϕ. By comparing the stress-energy tensor of the scalar field with that of a perfect fluid we can get expressions that describe the energy density and pressure of the field,

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi), \tag{1.22}$$

and

$$\rho_\phi c^2 = \frac{1}{2}\dot{\phi}^2 + V(\phi), \tag{1.23}$$

where the first term can be thought as a kinetic energy, and the second is a potential energy. These terms can be substituted into the Friedman equation for $\ddot{a}$ (Equation 1.14) gives,

$$\frac{\ddot{a}}{a} = -\frac{8\pi G}{3c^2}\left(\dot{\phi}^2 - V(\phi)\right). \tag{1.24}$$

So we see that for accelerating expansion (i.e. $\ddot{a} > 0$) this requires $\dot{\phi}^2 < V$. Therefore $\dot{\phi}$ is generally assumed to be small (known as the *slow roll condition*), which makes solving the Friedmann equations for a given form of $V(\phi)$ much simpler.

Inflation also explains the generation of perturbations in the matter density. By Heisenbergs Uncertainty Principle there would be quantum fluctuations in the early Universe which during the inflationary phase which would be magnified to a cosmic scale. The concordance model assumes these fluctuations are independent of scale, or in other words they have a Gaussian statistical distribution. This is predicted by many inflationary theories, for example Guth (1981), but even in the absence of definitive physical evidence for a specific theory this choice

---

[4] A scalar field is just a function that assigns a scalar value to every point in space.

is preferred because it makes the minimum amount of assumptions (Peacock, 1999). These fluctuations are observed in the cosmic microwave background (CMB - see Part I) , and are thought to be the seeds of large scale structure.

1.5.1 Fluctuation type

Adiabatic perturbations are density variations where all forms of matter have equal fractional over/under densities, whilst the total energy density does not necessarily remain constant. This is the other way around for *isocurvature* perturbations; density fluctuations for one form do not necessarily correspond to density variations in others but the total energy density is preserved. The concordance model assumes predominately adiabatic fluctuations. This is evidenced by the large anti-correlation seen in the cross power spectrum between the temperature and the polarisation in the CMB (Peiris et al., 2003).

1.6 Matter and Energy

1.6.1 'Normal' Matter

Normal matter, often referred to as *baryonic matter* which is the main constituent, includes baryons such as protons and neutrons, and leptonic matter such as electrons and neutrinos. At the beginning of the last century this material, seen in planets, stars and galaxies, was thought to constitute all the matter in the Universe. However it became apparent that this observed matter was not nearly sufficient to account for the total density of the Universe extracted from, for example the cosmic microwave background data.

1.6.2 Dark Matter

To account for the missing matter density, and other anomalous observations such as the rotation curves of galaxies, a form of matter that only interacts gravitationally was proposed called *dark matter*. This dark matter is very important for the formation of large scale structure as will be discussed further in Part II.

1.6.3 Dark Energy

In Section 1.2 we discussed that the expansion of the Universe is accelerating. The reason for this is not known, but cosmologists have circumvented the problem by proposing a hypothetical substance called *Dark energy* responsible for this expansion. In its most simple form it can be included as a simple *cosmological constant* on the right-hand side of the Friedmann equations (Section 1.4.3). It is often though of as a scalar field, similar to inflation, with a negative pressure which is causing the accelerated expansion of the Universe. Dark energy is a wide topic and involves a large body of research which, although interesting, has limited relevance to the main focus of this thesis and will therefore only be touched on here.

1.6.4 Evolution of matter/energy densities

To get a rough idea of how the density ρ varies as a function of time t, the Friedmann equations (Equations 1.13 and 1.14) can be combined with the equation of state (Equation 1.5), to give,

$$\rho \propto a^{-3(\omega+1)}. \tag{1.25}$$

Since the scale factor a is related to the redshift z as $\frac{a_0}{a} = 1 + z$ we get

$$\rho \propto (1+z)^{3(\omega+1)}. \tag{1.26}$$

For a pressureless material such as dust, $\omega = 0$ so,

$$\rho \propto (1+z)^{3}. \tag{1.27}$$

In Section 1.3.2 we saw that for relativistic fluids, $\omega = \frac{1}{3}$ so,

$$\rho \propto (1+z)^{4}. \tag{1.28}$$

By the nature of the dark energy term the density remains constant, which therefore results in a negative pressure. Therefore, as demonstrated in Figure 1.2 there have been three distinct periods in the history of the Universe; initially the energy density of the Universe was dominated by radiation until about 10^4 years after the Big Bang. Until recently it was dominated by matter, but now this mysterious dark energy appears to have become dominant.

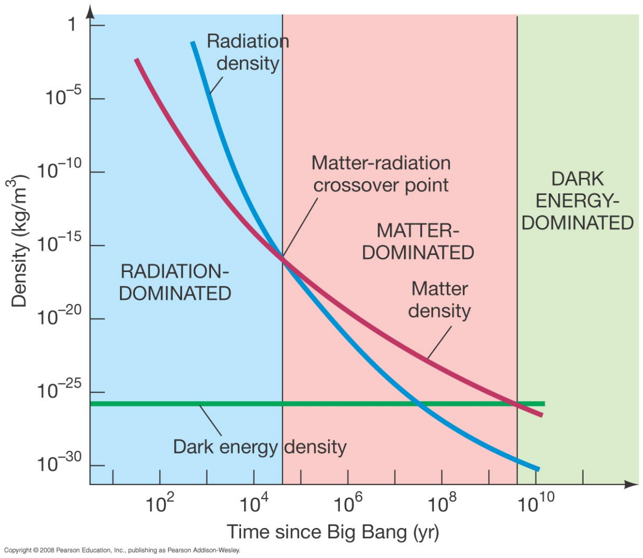

Figure 1.2: *Evolution of the Universe's radiation, matter and dark energy densities since the Big Bang.*

1.7 Testing the concordance cosmological model

So to summarise, the standard concordance cosmological picture is the current best fit model. It is supported by a collection of complementary sets of information from supernovae to the cosmic microwave background. However this model still has outstanding questions that need to be explained; two key ones include the nature of the dark matter and dark energy. So the model is still being developed as improved theories and analysis techniques continue to evolve, and increased volumes and quality of data continue to become available.

In this work we consider two complimentary sets of data. Firstly we chose the cosmic microwave background, which is the oldest light in the Universe, because it provides a large scale look at the Universe very shortly after the Big Bang and before the influence of gravity can disrupt any imprinted evidence of primordial conditions. Secondly, we take a look at the large scale distribution of galaxies which has also been affected by the primordial distribution of matter as well as being greatly influenced by gravity and the other processes of galaxy formation. Both these data sets, and a summary of the current understanding of the fields, are discussed in the two separate parts comprising this thesis.

Part I

Cosmic Microwave Background

Chapter 2

Introduction to the Cosmic Microwave Background

2.1 Standard Concordance Cosmological model

The standard *concordance* cosmological model (see Chapter 1) describes the very early Universe as containing a hot plasma of charged particles, including protons and electrons. Photons had small mean free paths due to Thomson scattering[1] and therefore the Universe was opaque.

As the Universe expanded, the temperature decreased sufficiently for ionised atoms and electrons to combine to form neutral particles ($\sim$ few thousand Kelvin) which is known as *recombination.* Therefore the mean free path of the photons suddenly grew large as their progress was no longer impeded by ionized atoms, and the Universe subsequently became transparent. These photons are seen today in the *Cosmic Microwave Background* (CMB). It is the furthest back in time that can ever be observed (at least via electromagnetic radiation) and is therefore known as the *last scattering surface.*

Another result of the Thomson scattering at the last scattering surface is that

---

[1] The elastic scattering of electromagnetic radiation by a free charged particle.

the CMB radiation is partially linearly polarized. This polarization signal is at least an order of magnitude weaker than the temperature anisotropy. Nevertheless it provides an important additional test of the concordance model.

2.1.1 Nature of the CMB Temperature Fluctuations

The distribution of matter in the early Universe is thought to have been homogeneous and isotropic on large scales (see Section 1.1), but there were fluctuations of the order $\delta\rho/\rho \sim 10^{-5}$. These are thought to have been generated via phase transitions during *inflation* (see Section 1.5). For the simplest theories of inflation, these fluctuations are thought to have had a Gaussian statistical distribution (Guth and Pi, 1982; Bardeen et al., 1986).

These matter fluctuations are thought to be closely related to the fluctuations in the CMB radiation. There are different ways in which these fluctuations in the CMB temperature would be generated. Here are just some.

- Firstly, if matter and radiation were coupled at the last scattering surface (i.e. *adiabatic* fluctuations) then areas of higher matter density would be intrinsically hotter, i.e.

  $$\frac{\delta T}{T} = \frac{1}{3}\frac{\delta\rho}{\rho}, \tag{2.1}$$

  where ρ is the matter density and T is the temperature.

- In addition to this, matter at the last scattering surface would have been moving around i.e. random motions. This would appear as a temperature fluctuations as the radiation would be *red* or *blue-shifted* depending on whether the matter was moving away from or towards the observer, i.e.

  $$\frac{\delta T}{T} \sim \frac{v}{c}, \tag{2.2}$$

  where v is the speed of the random motions and c is the speed of light.

- Lastly, the *ordinary Sachs-Wolfe effect* (Sachs and Wolfe, 1967) describes how CMB photons were gravitationally red-shifted (or blue-shifted) by over-densities (or under-densities) in the matter distribution at the last scattering surface,

$$\frac{\delta T}{T} = \frac{1}{3}\frac{\delta \Phi}{c^2}, \tag{2.3}$$

where Φ is the gravitational potential.

These different effects contribute to the CMB fluctuations on different scales. The first two processes occur at recombination and therefore dominate on intermediate scales around the size of the horizon at that time. The Sachs-Wolfe effect, whilst occurring at recombination, is due to gravity associated with primordial fluctuations and therefore dominates on the large scales.

2.1.2 Evolution of the CMB Temperature Fluctuations

The CMB temperature fluctuations observed today are also affected by processes that occurred since the surface of last scattering.

- The *early integrated Sachs-Wolfe effect* occurred immediately after the ordinary Sachs-Wolfe effect. The CMB photons were gravitationally red/blue-shifted due to over/under-densities in the radiation density, which likely dominated the energy density at that time.

- The *late integrated Sachs-Wolfe effect* caused CMB photons traveling through large-scale over-dense (under-dense) areas to get a residual boost (loss) of energy because of the comparatively recent acceleration of the expansion of the Universe due to dark energy.

- Significant contamination of the CMB has also been generated via the *Sunyaev-Zeldowich effect.* This is caused when CMB photons pass through areas of high energy electrons; the CMB photons gained energy when they *inverse Compton scattered* off the electrons. This depletion in lower energy photons, and an excess in higher energy photons, has the effect of making the temperature look lower.

2.1.3 Statistics of the CMB Temperature Fluctuations

Here we discuss some of the standard statistics used to describe a set of density fluctuations,

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}. \tag{2.4}$$

The *two point (auto) correlation function* of a set of density fluctuations is defined as,

$$\xi(\mathbf{r}) =< \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) >, \tag{2.5}$$

where $<>$ represents the normalised average over a given volume. It describes that probability of finding two objects separated by a distance **r** above what you would expect for a random distribution. However, rather than considering statistics in physical coordinates, it is frequently more convenient to consider fluctuations as a superposition of modes. The problem with considering say separate volumes in physical space is that over time they will no longer be independent as they interact with each other via gravity. The advantage of modes is that they will evolve independently unless effected by a very strong gravitational conditions (i.e. a non-linear regime). The fluctuations are expressible as a Fourier transform,

$$\delta(\mathbf{x}) = \sum_{\text{all } \mathbf{k}} \delta(\mathbf{k}) \exp{(\mathbf{ik}.\mathbf{x})}, \tag{2.6}$$

where **k** is the wave vector. The *power spectrum* is then defined as,

$$P(k) =< |\delta(k)|^2 >, \tag{2.7}$$

where $k = |\mathbf{k}|$ is the wave mode. So the power spectrum, which is essentially the Fourier transform of the two point correlation function, is a measure of the strength of the clustering as a function of scale. If the distribution of fluctuations is Gaussian there is no dependence on direction or position, so all the information contained in the fluctuations can be condensed into the power spectrum.

The CMB power spectrum provides much information about the early Universe, from its amplitude and the peaks that are observed in it. The peaks are known as *Doppler peaks.* As previously discussed, the density of the early Universe was incredibly smooth, bar some small fluctuations. Gravity acted upon over-densities to compress them, however radiation pressure counter acted this causing acoustic oscillations to form. At recombination, the radiation was released and therefore these oscillations ceased oscillating. Sound waves that had reached the extreme of their oscillation at that time would have enhanced CMB temperature fluctuations; this is seen as the peaks observed in the power spectrum.

These Doppler peaks are very sensitive to the different cosmological parameters. For example the first peak corresponds to the physical length scale of the horizon at recombination, which is equivalent to $\ell \sim 200$ in a flat universe. This in itself is not very sensitive to the cosmological parameters, but the geometry of the Universe affects the apparent size of this length scale to the observer. In a universe with positive curvature, the angle which subtends this length scale appears larger, and therefore the the associated ℓ calculated is small (vice versa for negative curvature).

As a bit of an aside, these Doppler peaks are also visible in the galaxy matter distribution (Percival et al., 2007) but at a much smaller contrast. This is because the 'normal' matter seen in galaxies comprises only a small fraction of the total matter, and the matter power spectrum has evolved significantly since the last scattering surface.

2.2 CMB Discoveries

The CMB was first discovered by Penzias and Wilson (1965) whilst they were looking for the source of some 'noise' in their radio detector. They were investigating the properties of atmospheric noise when they found a signal at microwave frequencies that they were unable to explain. With the help of Dicke and Peebles they soon realized that this 'noise' was, in fact, the CMB radiation

(Dicke et al., 1965). Following its discovery there have been many experiments to measure the CMB. Barreiro (2010) provides a good summary of recent and future experiments. Just of few of them are discussed here.

Cosmic Background Explorer COBE (Smoot et al., 1992) was one of the first space-based experiment to measure the CMB. COBE was based in space so it could map the whole sky without being subject to atmospheric absorption/noise. It carried three instruments; the Differential Microwave Radiometer (DMR) which measured the CMB temperate fluctuations, the Far-InfraRed Absolute Spectrophotometer (FIRAS) which measured the spectrum of the CMB, and the Diffuse InfraRed Background Experiment (DIRBE) which was used to map dust emission.

Millimeter Anisotropy eXperiment Imaging Array Spaced based experiments are very expensive, but the atmosphere is very restrictive to ground based telescopes when observing in the microwave bands. Many projects circumvent these problems by using balloon based instruments. MAXIMA (Hanany et al., 2000) was one such experiment, first launched in 1995, to investigate the Doppler peaks in the CMB power spectrum.

Balloon Observations Of Millimetic Extragalactic Radiation and Geophysics Another balloon based experiment, BOOMERanG (de Bernardis et al., 2000), had two flights around Antarctica in 1998 and 2003 with similar objectives as MAXIMA. It made large, high fidelity images of the CMB temperature anisotropies, over 40 times more detailed than the COBE map.

Wilkinson Microwave Anisotropy Probe WMAP (Bennett et al., 2003) was another space-based experiment which measured the small scale temperature anisotropies of the CMB with much greater precision yielding strong constraints on the cosmological parameters down to a few per cent. Rather than measure the temperature of the CMB, WMAP measured the difference in temperature between different points on the sky. It made observations in 5 wavebands; K-band (13 mm), Ka-band (9.1 mm), Q-band (7.3 mm), V-band (4.9 mm), and W-band (3.2 mm). The probe was launched in 2001 and has had data releases at 1, 3, 5 and 7 years.

Quest (Q and U Extra-Galactic Sub-mm Telescope) at DASI (Degree Angular Scale Interferometer) QuaD (Church et al., 2003) was a ground based instrument designed to measure the spatial distribution of polarization of the CMB. The CMB polarization is an important complementary set of information to the CMB temperature to extract information about cosmological parameters, origin of temperature anisotropies in the CMB and differentiating between different inflationary models.

Planck Space Observatory Planck (Tauber et al., 2010) was launched in 2009 with the preliminary results expected in 2012. The aim of this new space-based survey is to record the CMB to even smaller angular resolution with the aim of identifying more clearly areas of contamination by foreground structure. This will allow both study of the foreground and a more precise cleaning of the full sky CMB maps.

The key results from these instruments are discussed below.

2.2.1 Black Body Temperature

The discovery and subsequent study of the CMB has been pivotal in providing evidence for the concordance model. COBE found that the temperature distribution of the CMB was incredibly smooth (isotropic) over the whole sky which provided strong evidence that the radiation came from an extragalactic source. Also the spectrum from this background was incredibly close to that of a black body as shown in Figure 2.1. This provided strong evidence for the Big Bang model over its competitor at the time, the *Steady State* theory, which could not account for these observations.

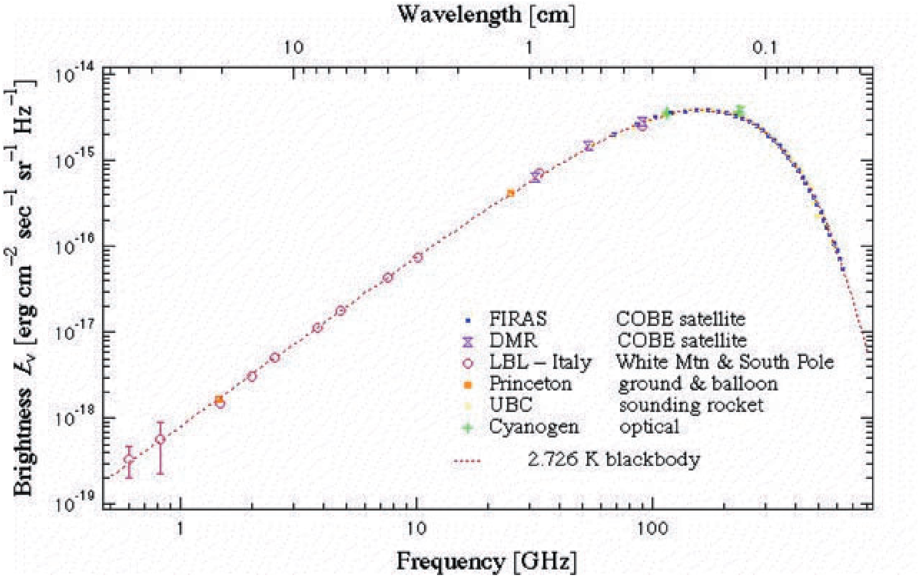

Figure 2.1: *Intensity plot of the CMB as measured by COBE and other ground based experiments (Smoot and Scott, 1996). The data points show excellent agreement with the theoretical intensity curve for a black body.*

2.2.2 Temperature Anisotropies

COBE found that whilst the CMB temperature is incredibly smooth over the sky there are some very small fluctuations at the level of around one part in one hundred thousand (see Figure 2.2). These temperature fluctuations are important because they directly relate to the matter density at the time the CMB formed (as discussed in Section 2.1.1), so this map actually provides a fingerprint of the matter distribution at that early time.

The fluctuations in the matter density are thought to be the seeds from which structure has since formed and so much effort has gone in to studying them. If these fluctuations are the seeds of structure formation it is important to study them in the CMB to understand their properties before they have been influenced by non-linear gravitational effects. These fluctuations hold information about the conditions and processes at the time of their formation and can tell us much about the very early Universe.

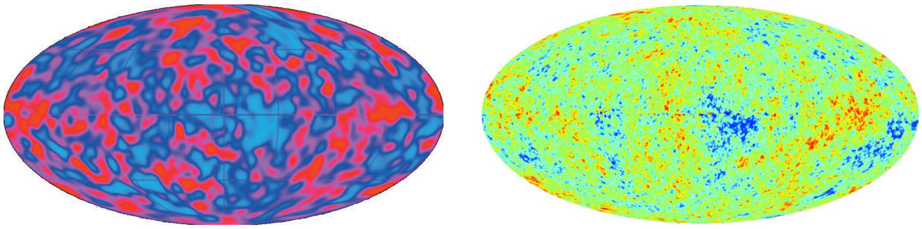

Figure 2.2: *Temperature maps of the CMB as measured by COBE (Smoot et al., 1992, left) and WMAP (Jarosik et al., 2011, right). The pink spots on the COBE map represent the hotter areas whereas the blue spots the colder areas. The fluctuations in the COBE map are of the same variance as the instrumental noise. The WMAP image shown is the 7 year Internal Linear Combination map. The temperature scale is of the order of milliKelvin (mK).*

2.2.3 CMB Power Spectrum

As discussed in Section 2.1.3, a power spectrum describes the difference between the local density and the mean density as a function of scale. Figure 2.3 shows the CMB power spectrum as derived from WMAP data compared to the best fit simulation using the concordance model. The errors are due to instrumental noise; the grey band represents cosmic variance, which is essentially statistical uncertainty given there is only one realization of the Universe.

The *Doppler* peaks in the power spectrum are of great interest because their positions and amplitudes are very sensitive to the different cosmological parameters such as the different matter/energy densities. For example the first peak is sensitive to the geometry of the Universe, which has been found to be very close to flat (de Bernardis et al., 2000).

2.3 Ongoing and Future CMB Studies

As well as trying to further constrain the cosmological parameters, many groups studying the CMB are interested in developing methods to look for deviations from the concordance cosmological model. One way in which this can be done is by considering the distribution of temperature fluctuations. The standard

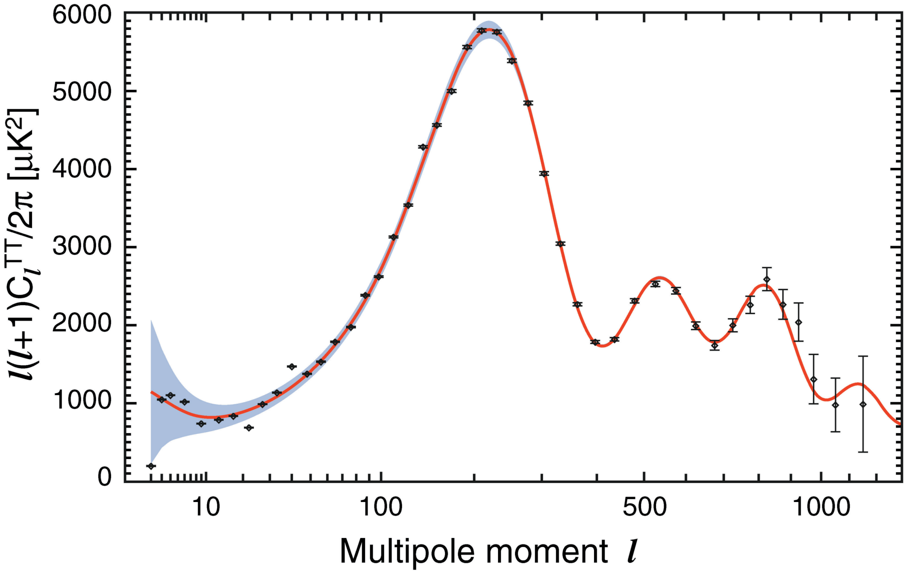

Figure 2.3: *The 7-year temperature power spectrum from WMAP as a function of multipole moment. The multipole moment is a Fourier parameter which is inversely proportional to scale. The curve is the current best fit cosmological model, which is the concordance model (see Chapter 1). The plotted errors are due to instrumental noise; the grey band represents cosmic variance.*

model predicts that these fluctuations should follow a Gaussian distribution (see Section 1.5) so many groups are currently looking to detect *non-Gaussian* anisotropies. Currently there are no definitive detections of non-Gaussianity although there have been some quite significant results. Martinez-Gonzalez (2008) provides a good summary. A few examples are discussed below.

2.3.1 Wavelets and the Cold Spot

Wavelets are a common technique used in signal processing, and have recently been employed in the analyses of CMB maps. They are small oscillations of various shapes that are used to decompose a signal over different scales. By analysing the distribution of wavelet coefficients at specific scales, it is possible to detect weaker non-Gaussianity than would be detectable in just pixel values.

The distribution of pixel temperature values can be insensitive non-Gaussianity because of the central limit theorem which says a combination of non-Gaussian features on different scales tends to a Gaussian as the number of features increases. Utilizing wavelets means that the signal can be analysed at specific scales, thereby avoiding this restriction (Hobson et al., 1999)

Wavelets were used to identify the *cold spot*, which is a cold area in the CMB ~ 10 deg in size at Galactic coordinates $b = -57$, $l = 209$. It was first reported by Vielva et al. (2004) and Cruz et al. (2005) using the wavelet analysis on a map generated from the combined signal in the Q+V+W bands from the 1 year WMAP data.

To account for this cold spot there have been many proposed explanations; two of these are a *cosmic texture* (Cruz et al., 2007) or a *large void* (Inoue and Silk, 2006). A texture is a type of topological defect which creates a time-varying gravitational potential. This causes CMB photons passing through to be red or blue-shifted. A void, or under-density of galaxies/clusters in the local universe, would cause CMB photons to appear bluer by the *integrated Sachs Wolfe effect* which causes photons traveling through large-scale under-dense areas to loose energy because of the comparatively recent acceleration of the expansion of the Universe. These predictions can be tested by their implications on up and coming lensing data sets (Masina and Notari, 2009; Das and Spergel, 2009). Of course another alternative is simply *cosmic variance* - that whilst unlikely, a cold spot of this size and magnitude is not implausible given the current standard model.

2.3.2 Phase Analysis

The concordance model says that the CMB fluctuations are well fit by a statistically homogeneous Gaussian random field, but this also means that the phases are independent and uniformly random (Watts and Coles, 2003). Therefore the sum of phases is also random (Coles et al., 2004) and so is the difference between phases (Chiang and Coles, 2000). The first step towards quantifying phase information was introduced by Coles and Chiang (2000) who applied a colour representation method, in this case to N-body simulations of large scale structure formation. This technique was later developed for application to the CMB by Chiang et al. (2003, 2004).

2.3.3 Multipole Vectors

Multipole vectors were first introduced to CMB non-Gaussianity studies by Copi et al. (2004) and Katz and Weeks (2004). They are constructed from spherical harmonics coefficients derived from CMB maps, but are useful because they give results in real (i.e. pixel) space. Land and Magueijo (2005) first identified an alignment of a number of multipoles in the 1 year WMAP data. This preferred direction became known as the *Axis of Evil*. A number of subsequent studies have further investigated and confirmed this claim: many different suggestions of its origin have been proposed but no definitive answers have yet been found. This isn't the only 'non-Gaussianity' to have been identified using multipole vectors. As well as identifying that the quadrupole (ℓ=2) and octopole (ℓ=3) are closely aligned, de Oliveira-Costa et al. (2004) found a lack of power in the quadrupole. Recently though work by Francis and Peacock (2010) may have potentially explained this anomaly; using photometric redshift surveys they estimated the integrated Sachs-Wolfe (see Section 2.1.2) effect on the CMB anisotropies and applied this to the WMAP data. This resulted in a reduction of the significance of a few reported anomalies (i.e. the low quadrupole power and the alignment of the quadrupole and octopole).

2.3.4 Polyspectra

As discussed in Section 2.2.3, the power spectrum describes the CMB temperature distribution compared to the mean temperature as a function of scale. It is the Fourier transform of the two point correlation function which describes the probability of finding two random galaxies separated by a given distance. One key observation made using the power spectrum by (Eriksen et al., 2004b; Park, 2004; Eriksen et al., 2007; Hoftuft et al., 2009; Hansen et al., 2009) is that there is an asymmetry in the power spectrum between the Northern and Southern hemispheres.

If the distribution of temperature fluctuations is Gaussian then all the information they hold can be contained in the power spectrum alone, with no correlation in the higher order polyspectra. The absence of correlation in the higher order polyspectra is therefore one way of looking for non-Gaussianities. Both the bispectrum and trispectrum (Fourier transforms of the three and four point correlation functions, respectively) have been used to test for deviations from the predicted Gaussianity of the temperature fluctuations.

A phenomenological way of parameterizing the level of non-Gaussianity in cosmological perturbations is to introduce a non-linearity parameter f_{nl} through the gravitational potential,

$$\varphi_g = \varphi + f_{nl}(\varphi^2 - < \varphi^2 >), \tag{2.8}$$

where φ represents the gravitational potential. Yadav and Wandelt (2008) found a non-Gaussian value of f_{nl} at a high significance of 99.5% using the bispectrum.

There have been many different attempts to investigate the Gaussianity of the CMB temperature fluctuations, with varied results and conclusions. So more detailed data and analyses methods are required to come to more conclusive results. In this work, a selection of new analyses methods are considered. In Chapter 3, and in Short and Coles (2010), a new simple diagnostic analysis is proposed and tested, which offers the possibility of identifying foreground-related biases and systematics in all-sky maps of the CMB. In Chapter 4, the behaviour of a selection of anisotropic models is studied so as to identify characteristics of the radiation fields they produce and test/develop methods that can be used to identify more general forms of anisotropy. Two statistical measures of anisotropy are studied in some detail. Finally, Chapter 5 summarises the conclusions.

Chapter 3

Zonal Modes of CMB Temperature Maps

As discussed in Chapter 2, there has recently been much interest in analysing full sky maps of the temperature anisotropies in the cosmic microwave background (CMB), particularly those from the Wilkinson Microwave Anisotropy Probe (Bennett et al., 2003, WMAP). More specifically much work has gone into identifying deviations from the standard *concordance* cosmological model which predicts Gaussianly distributed temperature fluctuations (see Section 1.5). This *non-Gaussianity* could either be from a *primordial* or *secondary* source; the analysis in this chapter is particularly aimed at identifying whether emission from local Galactic foregrounds has been cleaned sufficiently from a selection of full sky CMB maps. One of the greatest barriers to the detection of non-Gaussianity, or other departures from the framework of the concordance model, is the presence of residual foreground contamination or other systematic errors. Since our own Galaxy emits radiation at microwave frequencies, the emission from local foregrounds must be carefully cleaned out before a map can be obtained that is suitable for analysis. One way of avoiding this problem is to cut out regions of the map near the Galactic plane where contamination is particularly severe, but this negates the benefit of having full sky coverage.

All-sky maps of the CMB temperature fluctuations are frequently represented by a spherical harmonic decomposition (which will be explained later in more detail in Section 3.1.1). Briefly though, spherical harmonics involve modes labeled by ℓ and m (where $-\ell \leq m \leq +\ell$). The *zonal modes* (i.e those with $m = 0$) are of particular interest because they vary only with Galactic latitude; any anomalous behaviour in them might therefore be an indication of erroneous foreground substraction. In this chapter, a simple statistical analysis is performed on the zonal modes with low ℓ, for sky maps derived via different cleaning procedures from WMAP. This work shows that the zonal modes provide a useful diagnostic of possible systematics in CMB all-sky maps and was recently published in Short and Coles (2010).

3.1 Defining Zonal Modes

This section details what is meant by spherical harmonics, zonal and sectoral modes with some visual examples. We also describe how zonal modes can be used to distinguish between a Gaussian or non-Gaussian distribution of temperature fluctuations in the CMB.

3.1.1 Spherical Harmonics

All-sky maps of the CMB temperature are frequently represented by a spherical harmonic decomposition. Spherical harmonics are the angular portion of the solution to Laplace's equation in spherical coordinates. The statistical variation of the CMB temperature, $T(\theta, \varphi)$, over the celestial sphere can be conveniently decomposed as,

$$T(\theta, \varphi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta, \varphi), \tag{3.1}$$

where the $Y_{\ell m}(\theta, \varphi)$ are spherical harmonic functions, defined in terms of the Legendre polynomials, $P_{\ell m}$, using,

$$Y_{\ell m}(\theta, \varphi) = (-1)^m \sqrt{\frac{(2\ell + 1)(\ell - m)!}{4\pi(\ell + m)!}} P_{\ell m}(\cos\theta) e^{im\varphi}, \tag{3.2}$$

and the $a_{\ell m}$ are complex coefficients which can be expressed as,

$$a_{\ell m} = |a_{\ell m}| \exp(i\Phi_{\ell m}), \tag{3.3}$$

where $\Phi_{\ell m}$ are the phases and $|a_{\ell m}|$ are the amplitudes. Note, the Condon-Shortly phase definition is used here; this is the $(-1)^m$ factor which is not always included.

3.1.2 Zonal and Sectoral Modes

The spherical harmonics functions, $Y_{\ell m}$, can be visualised on a sphere by considering their *nodal lines*, i.e. the set of points (θ, φ) where $Y_{\ell m}(\theta, \varphi) = 0$. Nodal lines are therefore circles, either in the latitude or longitude direction with respect to the coordinate system being used; in the case of CMB maps this is usually the Galactic coordinate system. The number of nodal lines of each type is determined by the number of zeros in $Y_{\ell m}$ in the latitudinal and longitudinal directions; there are $\ell - |m|$ zeros in the latitude direction from the Legendre polynomial component and $2|m|$ zeros in the longitude direction from the exponential component (see Figure 3.1).

In cases where $0 < m < \ell$, there are zero-crossings in both directions, giving rise to a patchwork appearance; these are called *tesseral* modes (Figure 3.1). Two specific m are of particular interest in the context of this work: these denote the *zonal* modes, with $m = 0$, and the *sectoral* modes, with $m = \ell$. In the former case there are no zero-crossings in the longitude direction, so contours of equal temperature run parallel to latitude lines; in the latter the contours run parallel to longitude lines (see Figure 3.2). The analysis in this chapter will concentrate on the zonal modes because they vary only with Galactic latitude, therefore any anomalous behaviour in them may suggest some inconsistency

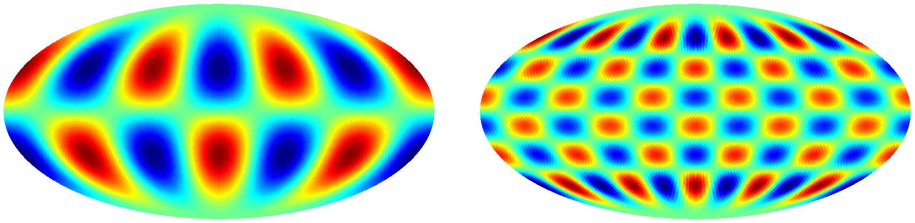

Figure 3.1: *Illustrative examples of spherical harmonics from individual ℓ and m mode combinations. The example to the left is for $\ell = 4$ and $m = 3$; the right hand example is for $\ell = 10$ and $m = 5$. Positive values of $Y_{\ell m}$ are shown in red and negative values in blue. The resulting patterns show an expected $\ell - |m|$ zero crossings (shown in green) in the latitudinal direction, and $2|m|$ in the longitudinal direction.*

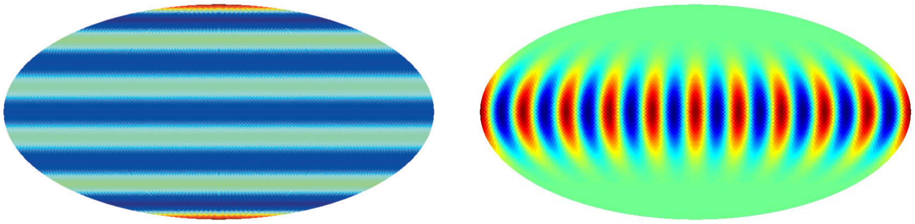

Figure 3.2: *Illustrative examples of zonal and sectoral modes for $\ell = 10$; the first example is the* zonal *mode with $m = 0$ and the second is the* sectoral *mode with $m = \ell = 10$. In the zonal case the contours of equal temperature run parallel to latitude lines; in the sectoral case the contours run parallel to longitude lines.*

with the subtraction of local Galactic foreground emission from full sky CMB maps.

3.1.3 Non-Gaussianity of Zonal Modes

This section considers the behaviour of the zonal modes in the null hypothesis i.e. for the concordance model. Statistically isotropic and Gaussianly distributed CMB temperature fluctuations, of the type that result from the simplest versions of the inflationary paradigm, possess spherical harmonic coefficients ($a_{\ell m}$) whose real and imaginary parts are both mutually independent and Gaussianly distributed (Bond and Efstathiou, 1987; Coles et al., 2004). The statistical properties of the fluctuations are therefore completely specified by the angular power spectrum, C_ℓ, (see Section 2.1.3) where,

$$C_\ell = \langle |a_{\ell m}|^2 \rangle = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} a_{\ell m}. \tag{3.4}$$

The spherical harmonics can be written as a sum of their real and imaginary parts,

$$a_{\ell m} = x_{\ell m} + i y_{\ell m}. \tag{3.5}$$

Note that for $m > 0$, the variances of the real and imaginary parts of $a_{\ell m}$ are equal,

$$\sigma^2(x_{\ell m}) = \sigma^2(y_{\ell m}) \equiv \sigma_\ell^2 = \frac{1}{2} C_\ell. \tag{3.6}$$

The distributions of $x_{\ell m}$ and $y_{\ell m}$ are independent and Gaussian with the same variances and zero mean when $0 < |m| \leq \ell$. Therefore the amplitudes $|a_{\ell m}|$ for these modes have a Rayleigh distribution, and the phases $\Phi_{\ell m}$ are random (Bond and Efstathiou, 1987; Stannard and Coles, 2005).

In the other case where $m = 0$, the imaginary part of $a_{\ell m}$ must be zero, so this mode always has zero phase $\Phi_{\ell m}$. This is because the phase relates to the variation around the polar axis only. Therefore the distribution of the amplitudes $|a_{\ell m}|$ for $m = 0$ is equal to that of the modulus of the real part $|x_{\ell m}|$ at a given ℓ, namely a Gaussian distribution with zero mean and stated variance.

3.2 Statistical Analysis with Zonal Modes

Statistical analysis of the spherical harmonics of CMB maps usually involves using all modes in an equivalent manner. However, since we are using the Galactic coordinate system this means the zonal modes are parallel to the Galactic plane. It is therefore worth looking at the properties of the zonal modes independently to see whether they hold any clues to possible residual contamination aligned with the Galactic plane.

In order to construct a test which involved the smallest possible number of assumptions, and in particular avoided the need to make estimates of the power-spectrum C_ℓ along the way, this analysis focused on the modes with maximum or minimum amplitude at a given ℓ. This way the test does not take into account the actual magnitude of the C_ℓ for which we would have had to make assumptions such as for the values of the cosmological parameters. The amplitudes for a given ℓ, which make up the the C_ℓ, are expected to be distributed randomly over the different m given Gaussian fluctuations regardless any other additional assumptions about the actual values of the C_ℓ. A preference for modes with $m = 0$ to display the maximum (or minimum) amplitude might therefore be plausibly interpreted as evidence that the zonal modes are either contaminated with residual foreground, or that foregrounds have been excessively subtracted. Both of these possibilities are supported by other evidence (Lopez-Corredoira, 2007; Chiang et al., 2007a; Naselsky and Verkhodanov, 2008; Chiang et al., 2009).

To complement this the variances of the zonal amplitudes of WMAP CMB maps, and the extrema of the remaining temperature map after removing all but the zonal modes, are also tested.

3.2.1 WMAP CMB Maps

The maps analysed in this chapter, and the next, are the 1-year (Bennett et al., 2003), 3-year (Jarosik et al., 2007), and 5-year (Hinshaw et al., 2009) Internal Linear Combination (ILC) maps from the WMAP team, the 1 and 3 year maps from Tegmark et al. (2003) and the 5 year harmonic ILC map by Kim et al. (2008) (hereafter ILC1, ILC3, ILC5, TOH1, TOH3 and HILC). They are all shown in Figure 3.3 at the resolution used i.e. $\ell = [0, 10]$. The maps show some variation from each other which is mainly localised around the Galactic plane.

The Internal Linear Combination (ILC) method takes smoothed temperature maps from each of the five WMAP wavebands and calculates a weighted linear combination that minimises the variance of the temperature values over the map

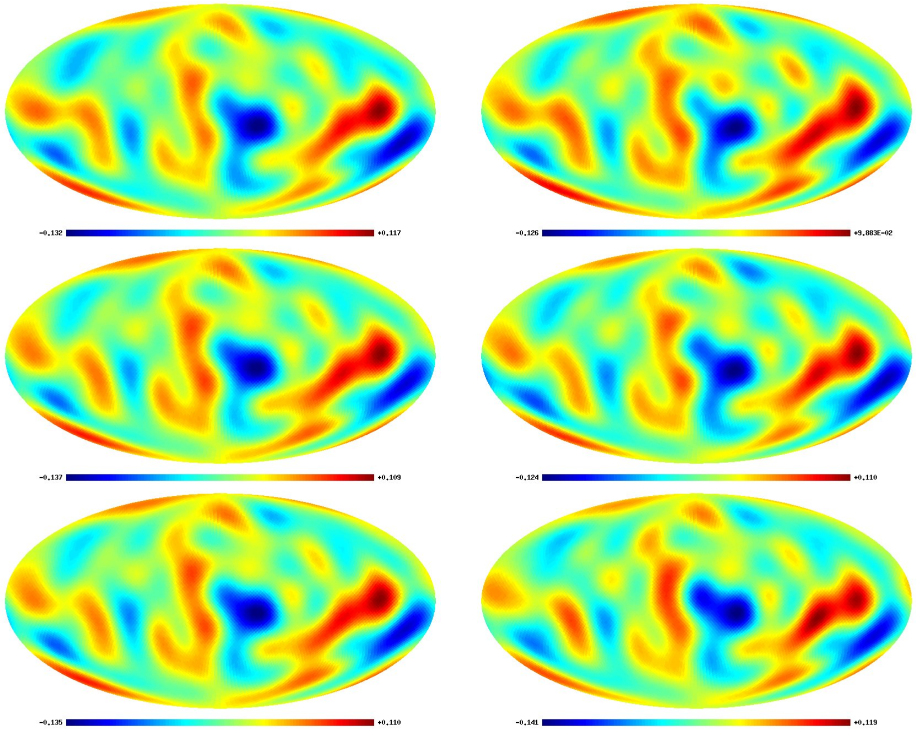

Figure 3.3: *CMB maps generated via different cleaning methods from observations using WMAP (resolution is for* $0 \leq \ell \leq 10$*). The maps are ILC1 (top left), ILC3 (middle left), ILC5 (bottom left), TOH1 (top right), TOH3 (middle right), and HILC (bottom right). The colour scale is marked in* μK*. The maps show some variation which is mainly localised around the Galactic plane.*

in real space (as opposed to Fourier space). To account for the spatial variation of the different amounts and sources of foreground contamination across the sky, this ILC method is applied to the CMB in twelve different regions, the majority of which cover the Galactic plane. The WMAP team are confident that on scales greater than ~ 10 degrees, or for $\ell \leq 10$, their ILC map gives a reliable estimate of the CMB signal (Limon, 2009). Hence this work uses only the smallest of modes up to $\ell = 10$.

However there have been some of criticisms of the ILC method. For example, the use of the different regions means that there are discontinuities in the maps.

Also, there is a cross correlation term between the foregrounds and CMB, which is not taken into account and could introduce bias. To overcome these two issues, Kim et al. (2008) proposed the harmonic ILC method (HILC) which works in a similar way to the ILC method but it minimises the variance in Fourier space as opposed to real space. Completing the variance minimisation in spherical harmonic space overcomes the first of the issues with discontinuities and they consider the cross term separately via an iterative method. Finally the method used to generate the maps by Tegmark et al. (2003) involves a combination of the two previous methods. The five different WMAP waveband maps are weighted together in such a way that depends both on both the angular scale and the distance to the Galactic plane.

3.2.2 Extremal Mode Counts

This first test considers the value of the m associated with the largest (or smallest) amplitude $|a_{\ell m}|$ for a given ℓ in the range $[0, 10]$. The spherical harmonics were extracted from the CMB maps discussed above using the anafast function in HEALPIX[1] (Gorski et al., 2005). The null hypothesis prediction, discussed earlier in Section 3.1.3, says that for all m where $0 < m \leq \ell$ the corresponding spherical harmonic amplitudes $a_{\ell m}$ should have the same variance; $a_{\ell m}$ for $m = 0$ should have half the variance. These predictions can then be used to calculate whether the minimum or maximum value of $|a_{\ell m}|$ in WMAP observations occur preferentially at any particular value of m compared to what would be expected. A very straightforward method is used to establish whether this is the case.

For each map the number of separate ℓ, within the range analysed, for which the minimum or maximum value amplitude occurs at $m = 0$ (the zonal mode) or $m = \ell$ (the sectoral mode) were counted. In other words, for each ℓ, the values of the m which have the minimum (or maximum) amplitude have been identified and then counted when either $m = 0$ or $m = \ell$. The resulting statis-

---

[1] http://healpix.jpl.nasa.gov

tics are recorded in Table 3.1. If the minimum is at $m = 0$ then this contributes to the count in Column 2 of the table; if the maximum is $m = 0$ it contributes to the count in Column 3; likewise occurrences of extrema at $m = \ell$ contribute to the counts in the following two columns. Given $\ell - 10$, the counts are out of a total of 11. As discussed in Section 3.2.1, this analysis is restricted to low ℓ modes because as previously discussed the high ℓ modes are not necessarily clean anyway so it would not reveal anything interesting to find zonal or sectoral anomalies amongst them.

Map	$m_{\text{min}} = 0$	$m_{\text{max}} = 0$	$m_{\text{min}} = \ell$	$m_{\text{max}} = \ell$
ILC1	6	3	3	4
ILC3	7	2	2	4
ILC5	7	2	2	4
TOH1	6	2	3	4
TOH3	8	2	2	4
HILC	8	2	1	5

Table 3.1: *This table shows, for the various cleaned all sky maps described in the text, the number of times (i.e. number of values of* ℓ*) for which the minimum (or maximum) amplitude* $|a_{\ell m}|$ *is at* $m = 0$ *(zonal mode) or* $m = \ell$ *(sectoral mode). For example, the* $m_{\text{min}} = 0$ *result for TOH3 means that 8 out of 11 different values of the* $\ell \in [0, 10]$ *have the lowest value of the amplitude* $|a_{\ell m}|$ *at* $m = 0$*.*

It is not a trivial matter to calculate significance levels analytically for this test because the number of available m increases with ℓ. It is therefore much more probable that the minimum amplitude is at $m = 0$ for $\ell = 2$ than for $\ell = 10$ under the null hypothesis. Assessing the significance of the number of occurrences of zonal or sectoral extrema involves a messy exercise in combinatorics. However, this difficulty is finessed by instead comparing the actual maps with simulations constructed to have Gaussianly distributed fluctuations as assumed. In addition to this, it is possible that the cleaning process used to remove foreground contamination from the raw observations might introduce some sort of bias into the statistical distribution of amplitudes or induce correlations between different modes. To circumvent this difficulty, as well as the one noted in the previous paragraph, the confidence levels in this test are based on a set of simulations performed by Eriksen et al. (2005). These take ‘raw’ simulated sky maps, generated assuming Gaussianly distributed fluctuations, add simulated foregrounds and then recover the signal using the ILC methodology. Therefore the simulated maps used already take into account any ‘artificial’ correlations that the ILC process may generate. As result any anomalies found are above the level known to be introduced by the ILC cleaning process.

In order to calculate significance levels of the results in Table 3.1, an ensemble of ten thousand independent Monte Carlo realisations of CMB skies with Gaussianly distributed fluctuations, contaminated with foreground and then cleaned according to the ILC prescription as described above, are used. These simulations are used to construct empirical distributions of the count statistics displayed in the previous table and from these the empirical significance levels are computed (see Table 3.2).

In the case of the $m_{\rm min} = 0$ result for TOH3 described in the caption for Table 3.2, the probability that a CMB sky has genuinely Gaussianly distributed fluctuations, processed in the way described above, would produce up to the number of zonal minima observed is only $(100 - 98.2) = 1.8\%$. The $m_{\rm min} = 0$ result for the HILC map is also of the same significance, which is pretty high, being just below 3σ. The result shows that the amplitudes of the zonal modes are smaller than the amplitudes of the other modes for a given ℓ more often than would be expected in the standard model. So, assuming that the amplitudes of the other modes do not deviate from the standard model which we confirm when we look at the variance in Section 3.2.3, this suggests that the amplitude of the zonal modes are smaller than would be expected for the standard model i.e. there is a deficit in the amplitudes of the zonal modes for these maps.

Note that since $N = 10000$ independent simulations are used, the results are expected to be affected by Poisson fluctuations at the level of order $\sqrt{N}$, which

Map	$m_{\text{min}} = 0$	$m_{\text{max}} = 0$	$m_{\text{min}} = \ell$	$m_{\text{max}} = \ell$
ILC1	75.2	15.4	64.2	80.4
ILC3	91.7	0.0	25.5	80.4
ILC5	91.7	0.0	25.5	80.4
TOH1	75.2	0.0	64.2	80.4
TOH3	98.2	0.0	25.5	80.4
HILC	98.2	0.0	0.0	94.9

Table 3.2: *Monte Carlo estimates of the probabilities of the extremal mode counts, i.e. occurrences of* $m_{\text{ext}} = 0$ *or* ℓ*, for the various CMB maps shown in Table 3.1. These are computed by forming the empirical distribution of the counts over an ensemble of simulated skies and counting what fraction of the ensemble gives the results obtained for the real maps. For example, in the case of the* $m_{\text{min}} = 0$ *result for TOH3, 9816 out of 10000 simulations have* less than *8 (from Table 3.1) occurrences of minimum amplitudes at* $m = 0$*. Given the probable sampling accuracy of around one percent, the results have been rounded. Note that these are discrete distributions, so the zero percentages do not (necessarily) indicate cases of exceptional significance, just that it is not possible to have less than the observed number of occurrences. In other words, they should be treated as a one-sided statistical test.*

means the probabilities in Table 3.2 will only to be accurate to approximately 1%.

The results for the sectoral modes (i.e. $m = \ell$) show no significant anomalies above 2σ (i.e. the 95% level); they are shown here for comparison and are not discussed further. For the case of zonal maxima (i.e. $m_{\text{max}} = 0$), again it is clear that there is no significant results.

We have seen the significant results are for the zonal minima ($m_{\text{min}} = 0$) in the HILC and TOH3 maps. It is also interesting to note that the TOH3 map gives a higher significance level than the TOH1 map, as does the ILC5 map compared to the corresponding 1 year map ILC1. This is an interesting result because it suggests that the anomaly in the zonal modes has actually become more prominent over time. This is counter intuitive to what we initially expected; the improving techniques to remove foreground contamination from CMB maps should mean that any anomalies become less significant over time.

These calculations are all performed in the Galactic coordinate system. However a number of anomalies have been identified in CMB maps aligned with the Ecliptic plane (Eriksen et al., 2004a; Copi et al., 2006; Diego et al., 2009). Therefore, for completeness the same calculation as described above is repeated for the same maps but in Ecliptic coordinates i.e. the maps are rotated by $8\pi/15$ degrees. The results are shown in Table 3.3. All of these results are much less significant than previously observed in the maps when considered in Galactic coordinates. Therefore no evidence is found here for anomalies from the standard concordance model for the CMB maps in Ecliptic coordinates.

Map	$m_{\min} = 0$	$m_{\max} = 0$	$m_{\min} = \ell$	$m_{\max} = \ell$
ILC1	19.5	49.1	88.5	0.0
ILC3	19.5	49.1	64.2	0.0
ILC5	19.5	49.1	88.5	0.0
TOH1	19.5	49.1	88.5	0.0
TOH3	4.1	49.1	25.5	0.0
HILC	48.0	15.4	64.2	0.0

Table 3.3: *Monte Carlo estimates of the probabilities of the extremal mode counts, i.e. occurrences of* $m_{\mathrm{ext}} = 0$ *or* ℓ*, for the various CMB maps in Ecliptic coordinates. These are computed by forming the empirical distribution of the counts over an ensemble of simulated skies and counting what fraction of the ensemble gives the results obtained for the real maps. Given the probable sampling accuracy of around one percent, the results have been rounded. Note that these are discrete distributions, so the zero percentages do not (necessarily) indicate cases of exceptional significance, just that it is not possible to have less than the observed number of occurrences. In other words, this should be treated this as a one-sided statistical test.*

3.2.3 Variance of Amplitudes for Zonal Modes

To further investigate the source of the apparent deficit in the zonal modes for the TOH3 and HILC maps reported in Section 3.2.2, this section considers the variance of the spherical harmonic amplitudes of the zonal modes. The results in Section 3.2.2 suggest that the spherical harmonic amplitudes when $m = 0$ are low in the 'real' CMB maps compared to the simulated CMB maps, so this test considers whether this is because the variance of the amplitudes is lower for $m = 0$ than for all other m. Since there is no imaginary part to the amplitude when $m = 0$, just the real parts of the $a_{\ell m}$ (i.e. the $x_{\ell m}$) are considered.

For this test the variance of the $|x_{\ell 0}|$ for each of the six WMAP CMB maps was calculated for $0 \leq \ell \leq 10$. The corresponding probabilities were calculated by comparing these values to the variances of the $|x_{\ell 0}|$ from the simulated CMB maps with Gaussianly distributed fluctuations, again from Eriksen et al. (2005). The resulting probabilities are detailed in Table 3.4; they show that the variance of the $|x_{\ell 0}|$ in the WMAP CMB maps are somewhat smaller than in the simulations. For all maps other than the ILC1, the variance of the $|x_{\ell 0}|$ is low at a significance above 2σ (i.e. the 95% level).

For comparison, the same calculation is also applied to the spherical harmonic amplitudes '*for all m*' (i.e. $|x_{\ell m}|$). The results '*for all m*', shown in the third column of Table 3.4, show that the variance in the WMAP CMB maps is lower than in the simulations with Gaussianly distributed fluctuations. This is consistent with the previously reported low variance in the combined Q+V+W map by Monteserin et al. (2008). However the result is less significant for the '*for all m*' than it was for the zonal modes; only the ILC5 map shows a result above 95%.

This leads to the further question of whether the zonal mode amplitudes are entirely responsible for this low variance. So finally the variance of the spherical harmonic amplitudes was also found '*for all m not equal to 0*'. The probabilities of the resulting values derived from the concordance model simulations are

Map	$m = 0$	$\forall m$	$m \neq 0$
ILC1	89.8	87.4	63.9
ILC3	97.1	93.0	69.8
ILC5	96.4	92.1	68.7
TOH1	95.2	95.5	80.1
TOH3	96.9	93.5	70.5
HILC	97.8	87.4	53.1

Table 3.4: *Percentage of simulated CMB maps which have a variance of the* $x_{\ell m}$ *which is greater than the specified WMAP CMB map, for given selection of* m*. For example, the TOH3 result for* $m = 0$ *means that 9691 out of 10000 simulations had a variance of the* $x_{\ell 0}$ *greater that the variance of the* $x_{\ell 0}$ *in the TOH3 map. Since the Poisson fluctuations are of order 1% the resulting percentage has been rounded.*

recorded in the last column of Table 3.4. They show that by removing the $m = 0$ amplitudes, the number of simulated maps with a variance greater than the test map falls dramatically, suggesting that indeed the low variance of the $m = 0$ amplitudes does have a notable affect on the overall variance of the amplitudes.

3.2.4 Zonal CMB Maps

It is interesting to reconstruct what the CMB sky would look like if it only contained the zonal modes as these are the modes that appear to have anomalous properties, as discussed in Sections 3.2.2 and 3.2.3. This was done by setting all amplitudes in the original CMB maps to zero bar the $|a_{\ell 0}|$; the resulting maps generated using the synfast function in HEALPIX are shown in Figure 3.4. The hottest regions tend to be near the poles and along the Galactic plane; the coldest regions are above and below the Galactic plane. This third section tests whether the minima and maxima pixel values of these *zonal maps* from WMAP are of abnormal amplitude compared to the simulated maps; note this test is done in pixel-space rather than using the $a_{\ell m}$. Because the resolution of the

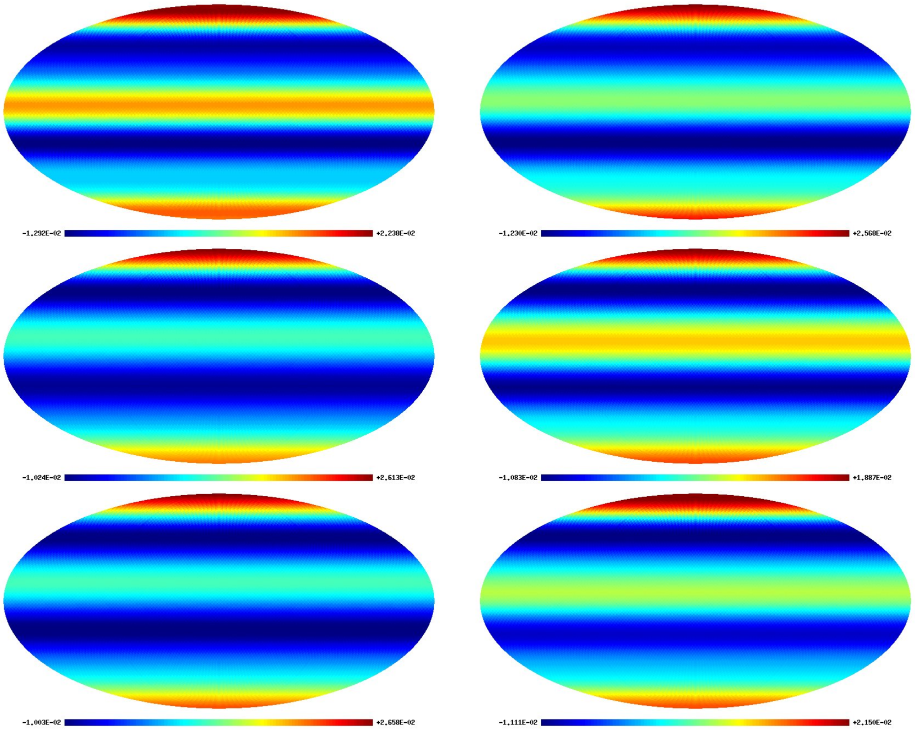

Figure 3.4: *Reconstructed CMB maps using only the zonal ($m = 0$) modes i.e. this map is constructed from an original CMB map which has had all amplitudes but the $|a_{\ell 0}|$ set to zero. The maps are ILC1 (top left), ILC3 (middle left), ILC5 (bottom left), TOH1 (top right), TOH3 (middle right), and HILC (bottom right). The colour scale is marked in μK. The hottest regions tend to be near the poles and along the Galactic plane; the coldest regions are above and below the Galactic plane.*

maps is quite low, the analysis was restricted to the parts of the map that are well-sampled (i.e. the Galactic pole areas were neglected).

For this test the maximum and minimum pixel values from each of the six different WMAP CMB zonal maps were recorded, as can be seen in Table 3.5. To calculate the corresponding probabilities, Monte Carlo simulations by Eriksen et al. (2005) were again used. Zonal maps were generated from each of the 10000 simulated CMB maps, and the maximum and minima pixel values were found in the same way as with the WMAP CMB maps. These Monte Carlo

Map	Min. value on the zonal map	Probability of the min. value	Max. value on the zonal map	Probability of the max. value
ILC1	-12.9	96.3	17.4	90.4
ILC3	-10.2	98.9	16.8	91.5
ILC5	-10.0	99.0	16.8	91.6
TOH1	-12.2	97.1	16.6	92.1
TOH3	-10.8	98.7	11.1	99.0
HILC	-11.1	98.4	16.3	92.6

Table 3.5: *Probability of the observed maxima and minima in zonal maps, as derived from Monte Carlo simulations. For example take the results for ILC5: out of 10000 simulations, 9903 have a minimum 'temperature' in the zonal* ($m = 0$) *map which is* less than *that observed in the ILC5 zonal* ($m = 0$) *map (see Figure 3.4). Similarly, out of 10000 simulations, 9159 have a maximum 'temperature' in the zonal* ($m = 0$) *map which is* greater than *that observed in the ILC5 zonal* ($m = 0$) *map. The 'temperature' is recorded in mK.*

simulations were then used to estimate the probability of each of the maxima or minima pixel values observed in the WMAP CMB maps. For example, if 9903 out of the 10000 simulations had a minimum pixel value in the zonal map which was *less than* that observed in the WMAP zonal map then the probability recorded would be 99.0%. The resulting probabilities are shown in Table 3.5. They show that the maxima (and minima) of the zonal maps are not as high (and low) as expected compared to the simulations. The significance of these results is higher in some WMAP CMB maps than others; the result for the minima values is more significant than than for the maxima. For the case of the minimum values, all maps give a result greater than 95% significance; only the TOH3 maps gives a result at that significance for the maximum case. This is consistent with work by Larson and Wandelt (2004) who found that the hot and cold spots of the separate WMAP frequency maps (for all m) are not hot and cold enough.

Also note that the anomalous results increase in significance as the maps supposedly become cleaner over time with improving techniques, corroborating the evidence in Sections 3.2.2 and 3.2.3.

Chapter 4

Statistical Characterisation of Temperature Patterns in Anisotropic Cosmologies

The previous chapter tested observed maps of the CMB against the standard *concordance* cosmological model, which predicts temperature fluctuations with a Gaussian distribution (Section 2.1.1). Observations of the CMB provide compelling support for the concordance model (Section 2.1.2). However, there is some evidence for non-Gaussianity, including a hemispherical power asymmetry (Eriksen et al., 2004b; Park, 2004) and a cold spot has been identified (Vielva et al., 2004; Cruz et al., 2005). In other words there is some evidence of an anisotropic universe; i.e. either a universe in which the background cosmology is not be described by the standard Friedman-Robertson-Walker (FRW) metric, or a universe with a non-standard topology.

This chapter considers a non-standard alternative to the concordance model, and looks at methods for characterising the coherent large-scale patterns from CMB temperature maps in globally anisotropic cosmologies. The methods investigated are reasonably general; the particular models which the methods are tested on are the homogeneous but anisotropic relativistic cosmologies described by the Bianchi classification. Although the temperature variations produced in these models are not stochastic, they give rise to a 'non-Gaussian' distribution of temperature fluctuations over the sky.

The Bianchi classification provides a complete characterisation of all the known homogeneous but anisotropic exact solutions to General Relativity. The classification was first proposed by Bianchi (1897), to describe the different symmetries of Lie groups, and was later applied to General Relativity by Ellis and MacCallum (1969). Initial studies used the lack of large-scale asymmetry in the CMB temperature to put strong constraints on the possible Bianchi models (Barrow et al., 1985; Bunn et al., 1996; Kogut et al., 1997). However, simulations of the CMB from Bianchi universes not only show a preferred direction, but models with negative spatial curvature (such as the types V and VII_h) can produce localised features (Barrow et al., 1985). So more recently attention has shifted to reproducing a cold spot such as that claimed to exist in the WMAP data (see Section 2.3.1). Initially, the Bianchi type VII_h was the favoured model to best reproduce the anomaly (Jaffe et al., 2005, 2006a,b); this has subsequently been investigated quite thoroughly (McEwen et al., 2006; Bridges et al., 2008). However Pontzen and Challinor (2007) and Pontzen (2009) found that whilst the VII_h type produces a localised feature in the temperature field, it generates large amounts of polarisation which is not observed. Therefore more recent work has also looked at the Bianchi type V which also produces localised features without the excess in the polarisation (Sung and Coles, 2009).

In the concordance model, the temperature fluctuation field is produced by stochastic fluctuations which are Gaussianly distributed and statistically stationary over the celestial sphere. As discussed in Section 3.1.3 this means that distribution of the real and imaginary parts of the spherical harmonic coefficients ($a_{\ell m}$) are independent and Gaussian with equal variance. If this assumption holds, then the amplitudes of the spherical harmonic coefficients ($|a_{\ell m}|$) are Rayleigh distributed and the phases of the spherical harmonics ($\Phi_{\ell m}$) are independent and uniformly random on the interval $[0,\, 2\pi]$ (Coles et al., 2004).

If instead the temperature pattern on the sky is produced by a Bianchi geometry then the $a_{\ell m}$ are no longer stochastically generated but can be directly calculated from parameters of the model. An analytical result for the CMB temperature fluctuations induced by the Bianchi models was first derived by Barrow et al. (1985). However this preliminary analysis did not include a dark energy component and so the work was later extended by Jaffe et al. (2006b). More recently, work by Pontzen and Challinor (2007); Sung and Coles (2011) considered a more comprehensive approach with a greater selection of free parameters and a more sophisticated treatment of recombination and reionsiation. As well as producing temperature results they also produced predictions for the polarisation fields. In this chapter, two methods are tested on simulations of the radiation field for anisotropic universes by Sung and Coles (2011). The first method is also discussed in Sung et al. (2010) and the second in Hansen et al. (2011).

Of all the models which have the FRW universe as a limit, the most interesting range of anisotropic structures are produced in Bianchi types VII$_h$, VII$_0$ and V. These different Bianchi types have the effect of focusing and/or twisting the initial quadrupole over time (see Figure 4.1). In this chapter, the behaviour of these Bianchi models is studied so as to identify characteristics of the radiation fields they produce and develop methods that can be used to identify more general forms of anisotropy. Understanding the characteristics identified in these particular cases will hopefully help to find better and more systematic ways of constraining the level of global asymmetry present in the real Universe. In this work, just characteristics observable in the CMB temperature map are considered; not the polarised component. Please note all the Bianchi simulations are courtesy of Rockhee Sung, and were generated as described in Sung and Coles (2011).

Two methods are explored for quantifying non-Gaussian and/or non-stationary fluctuation fields in order to see how they respond to the Bianchi models. Both methods involve the Fourier decomposition of the maps; these spherical harmonics are extracted from the simulations of the Bianchi CMB temperature maps

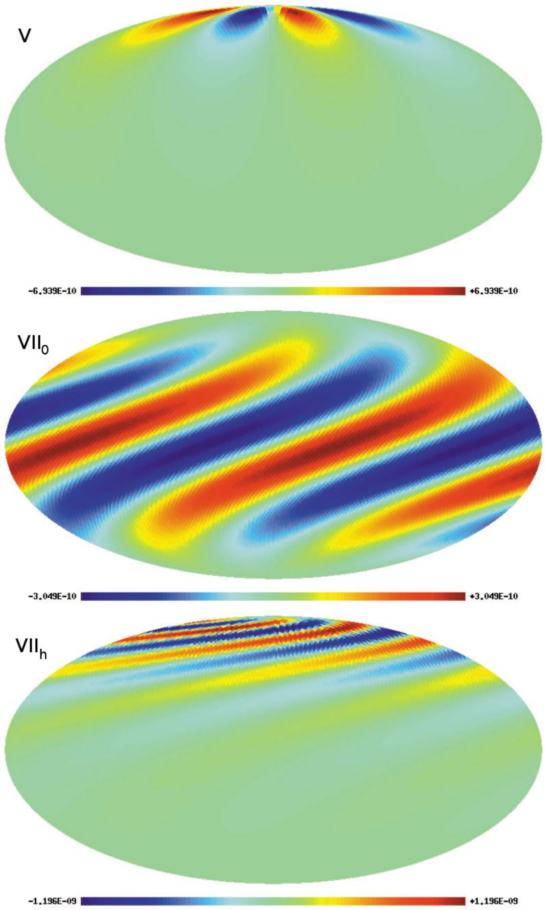

Figure 4.1: *Simulated maps of the the CMB temperature, at redshift* $z = 0$*, using Bianchi type cosmologies. The Bianchi types are: V, VII*$_0$ *and VII*$_h$ *(top, middle, and bottom respectively). The colour scale is marked in milliKelvin but please note that this and the redshift scale are arbitrary and can be scaled as required. All the maps started as a quadrupole at* $z = 500$*. The Bianchi V map shows a focused feature, the Bianchi VII*$_0$ *map has a twisted feature and the Bianchi VII*$_h$ *map has both focusing and twisting in the resulting temperature pattern.*

using Healpix . The first method investigates the behaviour of the phases of the (complex) coefficients obtained from the spherical harmonic decomposition of all-sky CMB maps. The second method examines the (normalised) cross-correlation of adjacent spherical harmonic coefficients. They both give clear signals of the presence of non-Gaussianity when applied to the selected Bianchi models, suggesting that they have some promise as diagnostics of the presence of global asymmetry in the Universe.

4.1 Phase correlations of Bianchi CMB Maps

This section considers the phases of the spherical harmonic coefficients ($\Phi_{\ell m}$) from simulated Bianchi CMB temperature maps to see if they have the potential to be a good descriptor to identify anisotropic universes. The phases $\Phi_{\ell m}$ were first used to analyse CMB temperature maps for non-Gaussianity by Coles et al. (2004) who tested them on isotropic and homogeneous but non-Gaussian fields. Since they have mainly been used to analyse observations (e.g. Naselsky and Verkhodanov 2008) rather than quantify expected deviations from Gaussianity for different types of models as they will be used here. The advantages of using the phases $\Phi_{\ell m}$ is their simplicity i.e. the assumption that they should be randomly distributed over the interval $[0, 2\pi]$ follows directly from the essential definition of the statistically homogeneous and isotropic Gaussian random field. The disadvantage of this is that they are somewhat generic in that they identify all types of 'non-Gaussianity'. They therefore may not be particularly optimised for efficiently finding the temperature patterns resulting from these anisotropic universes. But rather than make further assumptions about the data set, thereby making the descriptor less robust, the phases $\Phi_{\ell m}$ are tested here using a very simple statistical test, as described further below.

4.1.1 More about Spherical Harmonics

In Section 3.1.1, spherical harmonics were introduced with some visualisations of individual combinations of the ℓ and m. We recall some examples of visualisations of spherical harmonics on a sphere for individual values of ℓ and m in Figures 3.1 and 3.2; they all have $\ell - |m|$ rows and $2m$ columns. To further understand spherical harmonics, it is useful to see how $a_{\ell m}$ are visualised on a sphere for different *ranges* of ℓ and m, as demonstrated in Figure 4.2. The left plot shows that for ranges in ℓ, the initial quadrupole is disrupted so that the peaks/troughs fade towards the bottom of the map. This is because the number of rows in each individual ℓ map increases with ℓ. So when the individual ℓ maps are combined over the range, peaks and troughs near the origin always

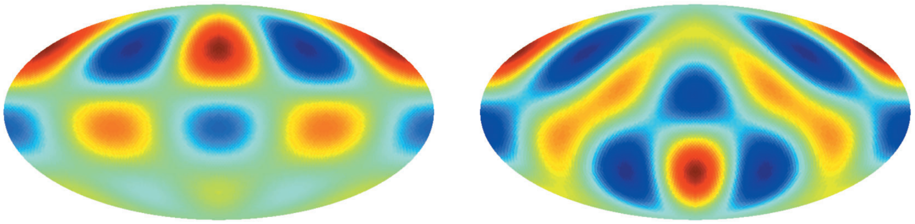

Figure 4.2: *Illustrative examples of spherical harmonics for different ℓ and m mode combinations. The first example is for ℓ = 3 - 4 and m = 2 (left); the second example is for ℓ = 4 and m = 2 - 3 (right). In other words the first example is $Y_{3\,2} + Y_{4\,2}$ and the second example is $Y_{4\,2} + Y_{4\,3}$. The first plot shows that for ranges in ℓ the quadrupole is disrupted so that the peaks/troughs fade towards the bottom of the map. For the range in m the signal is intensified in a 'v' shape.*

align, whereas the lower rows become increasingly out of sequence and hence annihilate. As the range in ℓ is increased, the quadrupole is focused towards the origin. This looks similar to the Bianchi type V map (Figure 4.1 - top); the number of columns in the Bianchi map suggests that $m = 2$ is dominant and the columns are much more focused suggesting a larger range of ℓ than in this example. For ranges in m (Figure 4.2 - right) the signal is intensified in a 'v' shape, which is similarly due to the fact that the number of rows and columns vary with m.

So a focused quadrupole can be expressed using spherical harmonics; the other feature to explain in the Bianchi maps in Figure 4.2 is the twisting seen in both the VII$_0$ and VII$_h$ type maps. If a small rotation was applied between each ℓ map in the previous example then it would make sense that a spiral might start to form. The rotation of a map is described by the phases ($\Phi_{\ell m}$) of the $a_{\ell m}$. Therefore $\Phi_{\ell m}$ should increase with ℓ to generate a spiral whose direction will depend on the sign of the phase difference ($\Delta\Phi_{\ell m}$) between each individual ℓ map. The example in Figure 4.3 demonstrates that this method does generate twisting and it is interesting because it looks similar to the map for the Bianchi type VII$_h$ map (Figure 4.1 - bottom). This will be useful to understanding the distribution of the phases of the Bianchi type VII$_h$ CMB map which we will discuss further in Section 4.1.2.

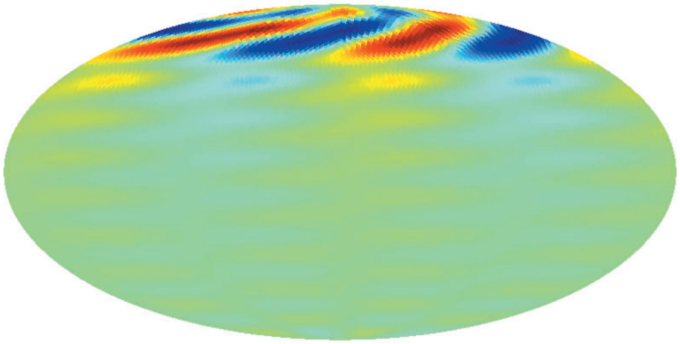

Figure 4.3: *Illustrative example of spherical harmonics when* $\ell = 0 - 20$*,* $m = 2$*, and* $\Delta\Phi_{\ell m} = \pi/8$.

One interesting point to note about Figure 4.3 is that the latitude of the spiral depends on the size of the $\Delta\Phi_{\ell m}$ since it depends on how much the maps are rotated by as to which bits annihilate and which intensify. Therefore to generate the Bianchi type VII$_0$ map (Figure 4.1 - middle) there must be a different $\Delta\Phi_{\ell m}$ to the Bianchi VII$_h$ map. This different $\Delta\Phi_{\ell m}$ is needed to align the spiral in the middle of the map. In addition to this, a smaller range of ℓ than used in Figure 4.3 is required to focus the twisted quadrupole less.

4.1.2 Visualising Bianchi Phase Correlations

A convenient way to visualise the information held in the phases of the spherical harmonic coefficients, $\Phi_{\ell m}$, of the Bianchi maps is to plot them over all ℓ and m. Rather than using a 3D plot, colour has been used to represent the $\Phi_{\ell m}$ following Coles and Chiang (2000). The colours equate to the angle on a colour wheel: red ($\Phi_{\ell m} = 0$), green ($\Phi_{\ell m} = \pi/2$), cyan ($\Phi_{\ell m} = \pi$), and purple ($\Phi_{\ell m} = 3\pi/2$).

To understand these plots, it is useful to first consider what would be expected in the case of an isotropic and homogeneous universe as predicted by the concordance model. This would be a uniform map (as at this point fluctuations are not yet being considered) but in spherical harmonics this only has power in one mode ($\ell = m = 0$), so there is no phase for the other modes. Better to consider a map with Gaussianly distributed fluctuations as later in this chapter fluctuations are added to the Bianchi CMB maps. Figure 4.4 shows the phases ($\Phi_{\ell m}$) for a homogeneous and isotropic map with Gaussianly distributed fluctuations. The phases are random over the space i.e. there are no visible patterns in the distribution of colours in the plot.

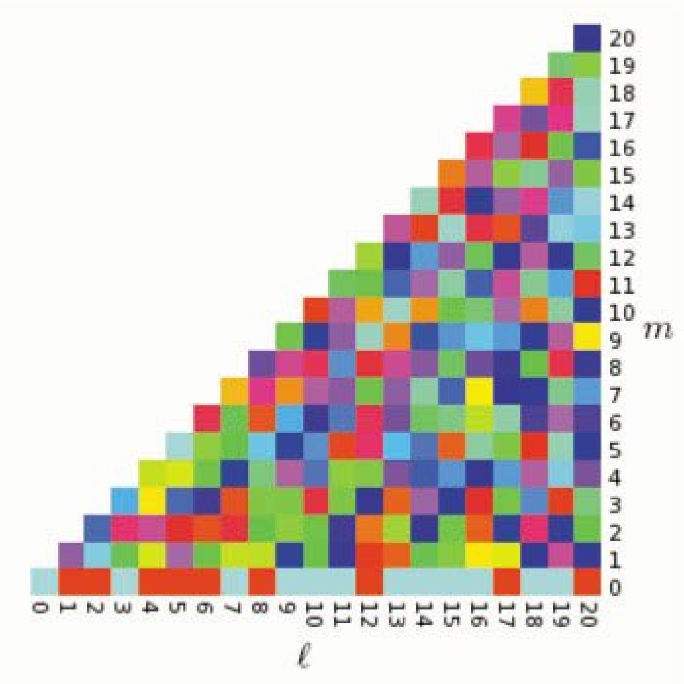

Figure 4.4: *Example of the spherical harmonic phases ($\Phi_{\ell m}$) for the concordance (isotropic) model where ℓ, $m \in [0, 20]$. The distribution of $\Phi_{\ell m}$ is random, apart from for $m = 0$, where the phases can only be 0 or π by definition. The colours represent the different values of the $\Phi_{\ell m}$; red ($\Phi_{\ell m} = 0$), green ($\Phi_{\ell m} = \pi/2$), cyan ($\Phi_{\ell m} = \pi$), and purple ($\Phi_{\ell m} = 3\pi/2$).*

Note that for all the maps, $\Phi_{\ell m}$ will equal either 0 or π for $m = 0$ because the $a_{\ell m}$ coefficients are defined so that $a_{\ell m} = a_{l,-m}$. Other than this, the distribution of $\Phi_{\ell m}$ is random. Also, note that for all $|m| > \ell$, $\Phi_{\ell m} = 0$.

In this analysis, the $\Phi_{\ell m}$ from each of the Bianchi maps were extracted using HEALPIX and plotted in the same way as Figure 4.4; the results are shown in Figure 4.5. The plots show that the $\Phi_{\ell m}$ are not random but have patterns, i.e. the harmonic modes manifest some form of phase correlation. For all the Bianchi types, $\Phi_{\ell m} = 0$ for all odd m. For the VII_0 and V types, all the $\Phi_{\ell m}$ are orthogonal i.e. they are either 0, $\pi/2$, π, or $3\pi/2$. Both the VII_0 and VII_h types show sequences of increasing/decreasing phases, which are particularly prominent for $m = 2$.

While some patterns are apparent in these plots, an even better way to visualise the phase correlations is to look at the phase differences which are defined here as,

$$\Delta\Phi_{\ell m} = \Phi_{\ell m} - \Phi_{\ell-1,m}. \tag{4.1}$$

Since the phases are random, the difference between any two random phases is itself random (Chiang et al., 2007b). The phase differences are shown in Figure 4.6 and the correlations are much more apparent compared to the plots of $\Phi_{\ell m}$. All the $\Delta\Phi_{\ell m}$ for the V type are lined up, i.e. either 0 or π. The $\Delta\Phi_{\ell m}$ for the VII_0 type are again orthogonal, but whereas in the phases the distribution of 0, $\pi/2$, π, and $3\pi/2$ seemed some what random, in the phase differences similar values 'aggregate' together. Similarly, the sequences of colours in the type VII_h (see $m = 2$ for example) are now even more prominent.

So strong correlations are observed in the phases and phase differences of the simulated Bianchi CMB maps, but only over large angular scales where there are only a small number of independent data points. Therefore, even without noise, it is important to ask the question whether these correlations are likely to be statistically significant. One way to quantify this is to use a *Kolmogorov-Smirnov test*. This is a non-parametric statistical test which measures the maximum distance of a given distribution from a reference probability distribution, which in this case is a random set of $\Delta\Phi_{\ell m}$, i.e. a uniform distribution, which is predicted by the concordance model.

To calculate the Kolmogorov-Smirnov test statistic, a set of phase differences

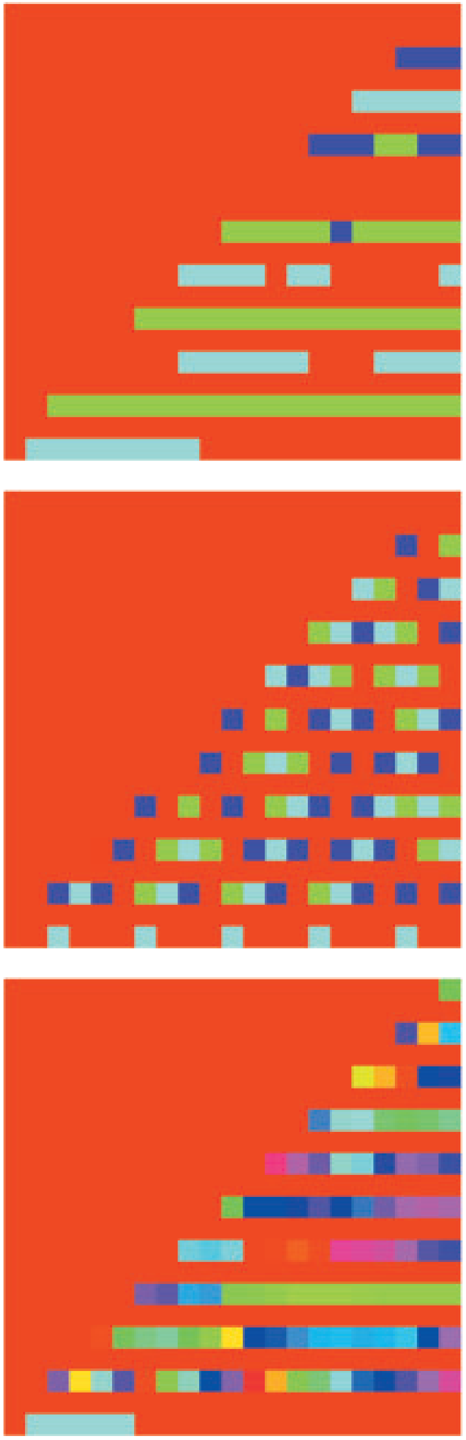

Figure 4.5: *Phases of the spherical harmonic coefficients ($\Phi_{\ell m}$) for Bianchi types V (top), VII*$_0$ *(middle) and VII*$_h$ *(bottom) where ℓ, $m \in$ [0, 20] and $z = 0$. Note that ℓ is plotted against the x axis, increasing from left to right, and m is plotted against the y axis, increasing from bottom to top. The distributions are not random (as in Figure 4.4) but exhibit some distinctive features. All the $\Phi_{\ell m}$ for the VII*$_0$ *and V types are orthogonal, and there are sequences of colours in the type VII*$_h$ *(see $m = 2$).*

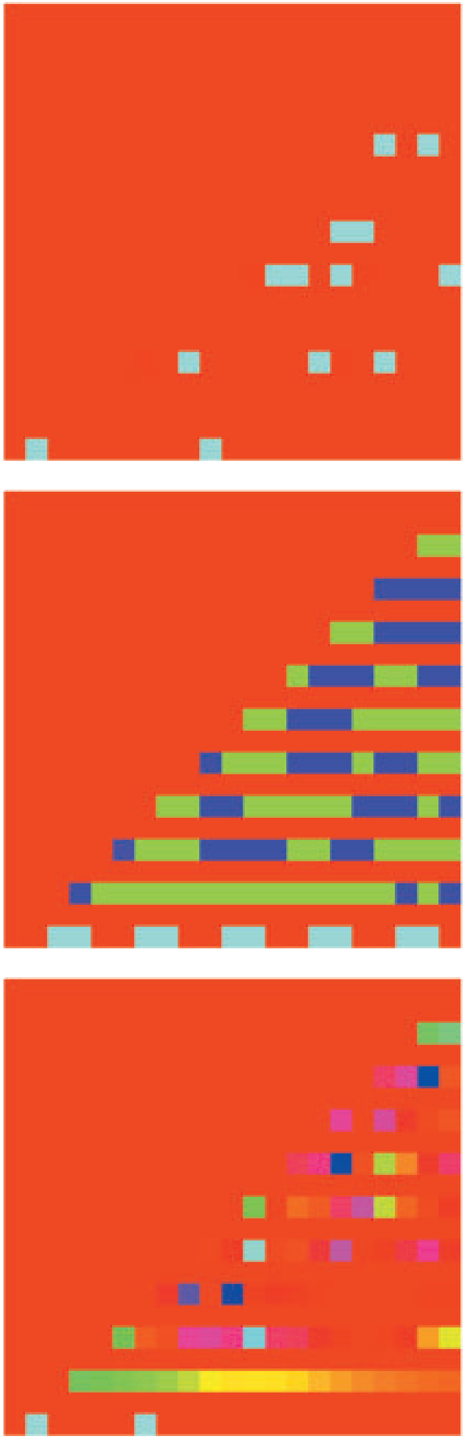

Figure 4.6: *$\Delta\Phi_{\ell m}$ for Bianchi types V (top), VII$_0$ (middle) and VII$_h$ (bottom) where ℓ, $m \in [0, 20]$ and $z = 0$. Note that ℓ is plotted against the x axis, increasing from left to right, and m is plotted against the y axis, increasing from bottom to top. Like the phases (Figure 4.5) the distributions are not random but exhibit some distinctive features. All the $\Delta\Phi_{\ell m}$ for the V type are either 0 or π. The $\Delta\Phi_{\ell m}$ for the VII$_0$ type are again orthogonal but in a more correlated way. Similarly, the sequences of colours in the type VII$_h$ are now even more prominent (see $m = 2$).*

$\Delta\Phi_{\ell m}$ are separated into ten bins of equal size between 0 and 2π. The number of $\Delta\Phi_{\ell m}$ which fall into each bin are counted and a cumulative distribution derived. If the distribution is uniform, as in the case of the reference probability distribution, then the number of $\Delta\Phi_{\ell m}$ in each of the bins should increase roughly linearly. The difference between both the sample and uniform cumulative distributions is found for each bin and the biggest difference is the Kolmogorov-Smirnov statistic **D**.

To deduce the significance of **D**, a set of ten thousand tests have been run to generate sets of random angles of equal size to the sample sets. **D** was found for each of these sets and this data was used to find the significance of **D** for the sample distributions from the Bianchi maps.

The Kolmogorov-Smirnov statistic **D**, and the derived probability of that statistic P(**D**), for all the Bianchi maps are detailed in Table 4.1. This table shows

Map	z	**D**	P(**D**) %
VII_h	500	0.11	94.5
VII_h	60	0.14	99.2
VII_h	3	0.27	>99.9
VII_h	1	0.38	>99.9
VII_h	0	0.27	>99.9
VII_0	0	0.28	>99.9
V	0	0.76	>99.9

Table 4.1: *Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the Bianchi CMB maps with a random distribution of phases as predicted by the concordance model.* ***D*** *is the Kolmogorov-Smirnov statistic; P(**D**) is the Monte Carlo estimate of the probability of getting the observed value of **D**, or less. These are computed by forming an empirical distribution of **D** from sets of random simulations and counting what fraction of the ensemble gives the results obtained for the Bianchi maps. For example, in the case of the P(**D**) for the* VII_h *map* ($z = 500$) *9447 out of 10000 simulations have a value of **D*** less than *0.11. Given the probable sampling accuracy of around one percent, the results have been rounded.*

that there is indeed a significant quantifiable deviation from a uniform distribution for the phase differences for all Bianchi types. Of the 10000 random sets of data, none showed a value for **D** as high as seen for the Bianchi cases.

The Bianchi VII$_h$ type was also considered at different redshifts [1] to see how the correlations changed with time. Table 4.1 shows that in general value of **D** gets more significant over time i.e. the correlations in the phase differences of the Bianchi maps become stronger over time.

4.1.3 Rotating maps and adding noise

In Section 4.1.2 the Kolmogorov-Smirnov test was applied to a 'clean' map that is perfectly aligned with the vertical axis. This section addresses how noise and rotation affect the identification of correlations in the phases of the spherical harmonics of CMB maps from Bianchi models.

First rotation is considered. Phases of spherical harmonic coefficients are not rotation-invariant. Rotating the coordinate system used to represent a CMB map in ϕ (which is equivalent to rotation around the z axis) would increase each of the spherical harmonic phases by ϕ, so the phase differences would remain the same. Therefore rotation in ϕ would have no effect on the value of the Kolmogorov-Smirnov statistic **D**. Rotation in θ is more complicated to express so an empirical approach is used to quantity the effect on **D**. The Bianchi CMB maps were rotated by a small angle, $\theta = \pi/8$, and then the spherical harmonic coefficients were derived and used to calculate **D**. The results in Table 4.2 show that the values of **D** for each of the maps are even higher than in maps that hadn't been rotated, indicating the presence of even stronger correlations. This

---

[1] Simulated CMB maps were generated for a selection of different redshifts. The temperature pattern at recombination ($z = 500$) was assumed to be a quadrupole. This quadrupole was required to generate interesting high order patterns in the radiation field (Sung et al., 2010). The radiation field was then evolved over time, and therefore smaller redshifts such as $z = 0$ represent how the patterns would look at later times. Whilst in this work we generally use the $z = 0$ example as an approximation for current time, the time scale was arbitrary hence we also considered a selection of other redshifts.

suggests that, at least for small rotations off the axis, the correlations are just as significant, if not more so.

Map	z	**D**	P(**D**) %
VII_h	0	0.46	>99.9
VII_0	0	0.38	>99.9
V	0	0.77	>99.9

Table 4.2: *Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the Bianchi CMB maps, rotated by $\theta = \pi/8$, with a random distribution of phase differences as predicted by the concordance model. $\boldsymbol{D}$ is the Kolmogorov-Smirnov statistic; $P(\boldsymbol{D})$ is the Monte Carlo estimate of the probability of getting the value of $\boldsymbol{D}$, or less. These are computed by forming an empirical distribution of $\boldsymbol{D}$ from sets of random simulations and counting what fraction of the ensemble gives the results obtained for the Bianchi maps. For example, in the case of the $P(\boldsymbol{D})$ for the VII_h map all of 10000 simulations have a value of $\boldsymbol{D}$* less than *0.46.*

As an aside, the colour plots of the phase differences for Bianchi maps rotated by a number of different θ in the range 0 to 2π were generated. These plots have been condensed together into movies[2] which show that the correlations in the VII_0 and V maps are visible across all θ and for the VII_h map are visible within about $\pi/3$ of the preferred axis.

Now to investigate the effect of noise, three different types of noise are considered. Firstly the simplest form was tried by just adding white noise to the Bianchi map. A map of random Gaussian noise (white noise) was generated and using HEALPIX the spherical mode resolution was reduced to $\ell \leq 20$. Then the 'noise' map was modified to have zero mean and variance the same as the Bianchi map. The second 'noise' map was derived from a product available on the WMAP LAMBDA[3] website which provides the effective number of observations per pixel. A map of random Gaussian noise was again generated.

---

[2] see http://www.astro.cardiff.ac.uk/research/theoreticalcosmology/?page=research to view the movies of the Bianchi phase differences.

[3] http://lambda.gsfc.nasa.gov/

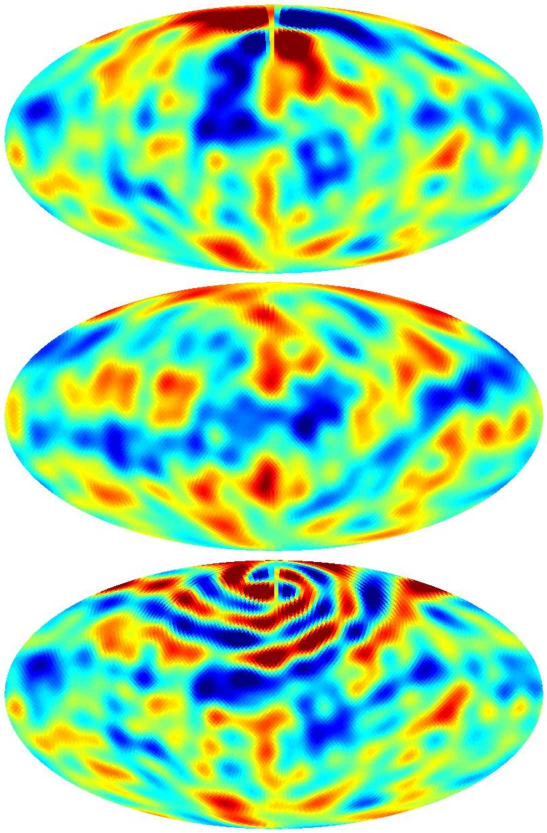

Figure 4.7: *Simulated maps of the the CMB temperature as shown in Figure 4.1, rotated off axis the same amount as the cold spot and with concordance fluctuations added. The Bianchi types are: V, VII_0 and VII_h (top, middle, and bottom respectively). The colour scale is again marked in milliKelvin but please note that this scale is arbitrary.*

The variance was modified per pixel so that it was inversely proportional to the square of the number of observations in that pixel. Using Healpix the spherical mode resolution was reduced to $\ell \leq 20$. Then the noise map was modified to have zero mean and variance the same as the Bianchi map. The final 'noise' map used a simulation of concordance fluctuations of the CMB as performed by Eriksen et al. (2005). Again the noise map was modified to reduce the spherical mode resolution to $\ell \leq 20$ and have the same variance of the Bianchi map.

Each of these 'noise' maps was added to each of the rotated Bianchi maps, and then the combined map was modified to return the overall variance and mean to the same as the original Bianchi map (examples are shown in Figure 4.7). The example in Figure 4.8 shows that the spherical harmonic coefficients derived still have visible correlations in the phases for the Bianchi V map. The results of the Kolmogorov-Smirnov test (see Table 4.3) show that the correlations are still detectable and significant for the Bianchi V and VII_h maps but not so well for the VII_0 maps[4]. So the method is better for detecting focused features that twisted features.

Map	z	$\mathbf{D}_r$	P($\mathbf{D}$)%	$\mathbf{D}_w$	P($\mathbf{D}$)%	$\mathbf{D}_c$	P($\mathbf{D}$)%
VII_h	0	0.11	94.5	0.11	94.5	0.16	99.8
VII_0	0	0.08	66.1	0.07	66.1	0.06	52.0
V	0	0.17	99.8	0.17	99.9	0.12	96.7

Table 4.3: *Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the Bianchi CMB maps rotated by $\theta = \pi/8$ with white (r), WMAP (w), and concordance model (c) noise maps ($z = 0$) . $\boldsymbol{D}$ is the Kolmogorov-Smirnov statistic found when considering the phase differences. $P(\boldsymbol{D})$ is the Monte Carlo estimate of the probability of getting the value of $\boldsymbol{D}$, or less, found for the Bianchi models, from a random selection of phase differences. These are computed by forming an empirical distribution of $\boldsymbol{D}$ from sets of random simulations and counting what fraction of the ensemble gives the results obtained for the Bianchi maps. For example, in the case of the $P(\boldsymbol{D})$ for the VII_h map with white noise 9447 out of 10000 simulations have a value of $\boldsymbol{D}$* less than *0.11.*

The effect of adding fluctuations here is not dissimilar to adding just Gaussian noise. The concordance model predicts fluctuations which are stationary and Gaussianly distributed, so although these fluctuations are correlated on the sky, they have random phases so are incoherent with respect to what our statistic measures.

The 'noise', or fluctuation, maps are added to the Bianchi maps so that the ratio

---

[4] Perhaps the most realistic noise model is a combination of 'WMAP' plus 'concordance model' noise. A test with equal contributions from each of these noise maps results in confidence levels which are consistent with the lowest values obtained when the noise maps were considered individually.

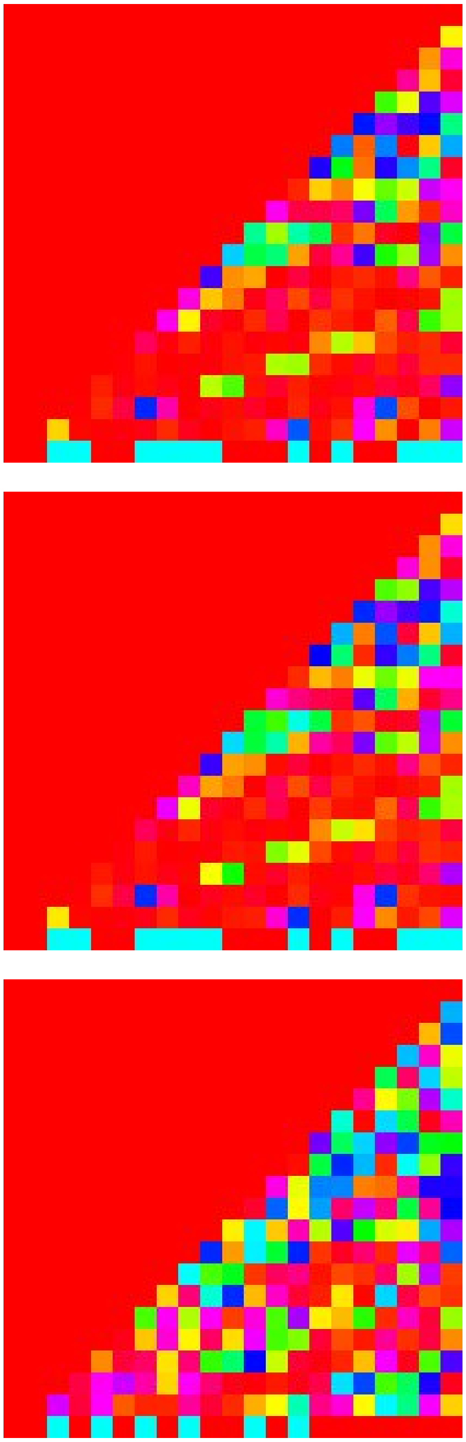

Figure 4.8: *$\Phi_{\ell m}$ for ℓ, $m \in$ [0, 20] (Bianchi type V map at $z = 0$ with white (top), WMAP (middle), and concordance model fluctuations (bottom) noise maps , $\theta = \pi/8$). Note that ℓ is plotted against the x axis, increasing from left to right, and m is plotted against the y axis, increasing from bottom to top. Correlations can still be observed; the colour of the plot overall has increasing red content.*

of the variances is of order unity. However, any ratio is possible; this specific choice is just for illustrative purposes to demonstrate the proposed methods. Nevertheless, for any amount of random-phase (Gaussian) signal that is superimposed on the Bianchi template, the phase coherence of the resulting map will still be degraded. If the Gaussian component is too large, the overall map will be indistinguishable from one with purely random phases. This analysis shows the method still functions well with this level of 'contamination', but if the noise variance is much higher than that of the Bianchi maps the method begins to reduce in effectiveness.

4.1.4 Application to WMAP 5 Year Data

For methodological interest, the approaches described in Section 4.1.2 are applied here to the WMAP 5 year ILC map (Hinshaw et al., 2009). The results of the Kolmogorov-Smirnov test on the ILC map in the Galactic coordinate system show very low significance correlations in $\Delta\Phi_{\ell m}$ (see Table 4.4), i.e. showing no evidence for an anisotropic Universe.

Map	Axis	**D**	P(**D**) %
ILC	Galactic	0.06	66.20
ILC	Evil	0.07	86.05

Table 4.4: *Results from the Kolmogorov-Smirnov test comparing the distribution of phase differences in the WMAP ILC map, rotated to align with either the Galactic axis or axis of evil, with a random distribution of phase differences as predicted by the concordance model.* ***D*** *is the Kolmogorov-Smirnov statistic;* $P(\boldsymbol{D})$ *is the Monte Carlo estimate of the probability of getting the value of* ***D****, or less, found for the Bianchi models, from a random selection of phase differences. These are computed by forming an empirical distribution of* ***D*** *from sets of random simulations and counting what fraction of the ensemble gives the results obtained for the Bianchi maps. For example, in the case of the P(****D****) for the ILC map in the Galactic plane, 6620 out of 10000 simulations have a value of* ***D*** less than *0.06.*

However Section 4.1.3 showed that to see correlations in the phases generated by the Bianchi cosmologies, the CMB maps needed to be rotated relatively close to the preferred axis. There have been studies that have found a preferred axis in the WMAP CMB maps, highlighted by the alignment of at least the quadrupole ($\ell = 2$) and octopole ($\ell = 3$). This preferred axis is known as the *axis of evil* (Land and Magueijo, 2005). Therefore here we rotate this WMAP CMB map to the axis of evil coordinate system to look to see if there is any evidence for the Bianchi cosmologies as any correlations should appear stronger close to the preferred axis. The methods from Section 4.1.2 are applied to the ILC map rotated so that the vertical axis aligns with the axis of evil. These $\Delta\Phi_{\ell m}$, plotted in Figure 4.9, do not show any visual correlations. For comparison the figure also includes a plot of the same ILC data but with the phases replaced with random angles (i.e. so as to not affect the magnitude of the amplitudes of the $a_{l,m}$).

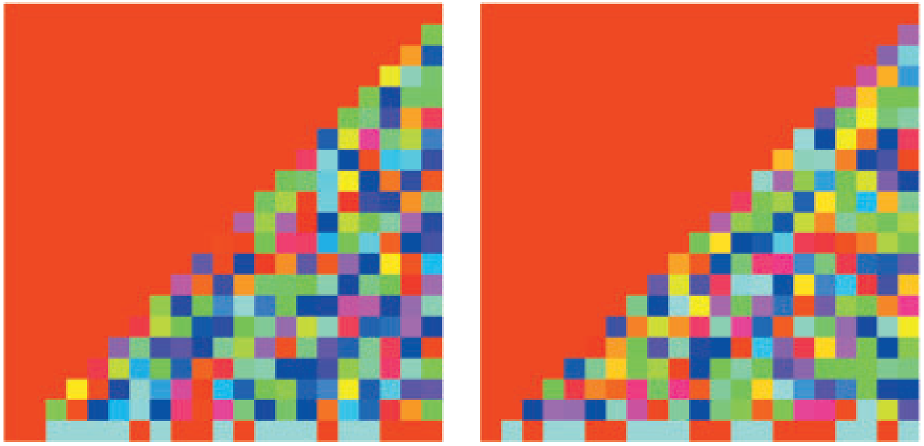

Figure 4.9: $\Delta\Phi_{\ell m}$ *for* ℓ, $m \in$ *[0, 20] for the 5 year ILC map with the axis of evil aligned with the preferred axis (left) and for the same map but with the phases replaced with random phases (right). Note that* ℓ *is plotted against the x axis, increasing from left to right, and m is plotted against the y axis, increasing from bottom to top. No correlations are visible in either plot.*

Whilst the Kolmogorov-Smirnov test (Table 4.4) finds higher significance results than when the map was in Galactic coordinates, the results are still at a low significance. The results show no significant detections, so if we do live in an anisotropic universe then any evidence for this hidden in the CMB must be obscured with considerable 'noise' (fluctuations). However the fact that the significance of the results does increase when the map is aligned with the axis of evil is intriguing; it would therefore be worthwhile investigating this method on future cleaner data sets such as that from Planck to see whether the anomaly still arises. A cleaner data set would allow the investigation of a wider range of modes; this would allow higher significances than considered here to be probed.

4.2 Cross Correlation of Bianchi CMB Maps

In this section, a second method for analysing the Bianchi CMB maps is considered. Hansen et al. (2011) proposed a new approach for identifying possible anisotropy in full sky CMB maps, using the cross correlation of the spherical harmonic coefficients ($a_{\ell m}$). The concordance model predicts the CMB temperature to be a Gaussian random field; this requires that the real and imaginary parts of the $a_{\ell m}$ should be independently and identically distributed. Therefore the amplitudes of the spherical harmonic coefficients, $|a_{\ell m}|$, should follow a Rayleigh distribution and the $\Phi_{\ell m}$ should be randomly distributed over $[0, 2\pi]$. The $a_{\ell m}$ are therefore not expected to be correlated and any cross correlation between them should be zero in the Gaussian case. Any couplings between different amplitudes would therefore indicate possible deviation from the statistical isotropy and Gaussianity of the CMB signal.

Hansen et al. (2011) showed that for dipole modulation anisotropy a cross correlation of the spherical harmonics is generated, and a selection of non-standard anisotropic models of the CMB are considered to classify the different features of the cross correlation. These included; a dipole modulation of the primordial signal, a primordial magnetic field, and anisotropic Bianchi cosmological models. This section discusses my contribution to the work which included investigating the expected signal from the anisotropic Bianchi universes using this descriptor. The results show no evidence for the dipole modulation or primordial magnetic fields in WMAP data, although these analyses do show evidence for residual foreground contaminating the maps. Therefore definitive results will require more accurate data such as that from Planck. For further discussion see Hansen et al. (2011).

4.2.1 Defining the Spherical Harmonic Cross Correlation

The normalised cross correlation of the spherical harmonic coefficients of full sky CMB maps is defined here as,

$$K_{kj}(\ell) = \frac{\sum_m a_{\ell m} a^*_{(\ell+k)(m+j)}}{\sqrt{\sum_m |a_{\ell m}|^2 \sum_{m'} |a_{(\ell+k)(m+j)}|^2}} = \kappa_{kj}(\ell) + i\chi_{kj}(\ell) \tag{4.2}$$

so that $\kappa_{kj}(\ell)$ and $\chi_{kj}(\ell)$ are given by,

$$\kappa_{kj}(\ell) = \gamma^{-1}(\ell) \sum_{m=-\ell}^{\ell} |a_{\ell m}||a_{(\ell+k)(m+j)}| \cos(\Phi_{\ell m} - \Phi_{(\ell+k)(m+j)}), \tag{4.3}$$

and

$$\chi_{kj}(\ell) = \gamma^{-1}(\ell) \sum_{m=-\ell}^{\ell} |a_{\ell m}||a_{(\ell+k)(m+j)}| \sin(\Phi_{\ell m} - \Phi_{(\ell+k)(m+j)}), \tag{4.4}$$

where

$$\gamma(\ell) = \left[\sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \sum_{m'} |a_{(\ell+k)(m+j)}|^2 \right]^{\frac{1}{2}}. \tag{4.5}$$

Note that in this calculation Φ_{l-m} was actually set to equal Φ_{lm}, since no additional information is held in the $\Phi_{\ell m}$ for $m < 0$ because a_{l-m} is the complex conjugate of $a_{\ell m}$, and else many of the correlations would otherwise be canceled out.

As is shown in Hansen et al. (2011), only a select few combinations of k and j will ever yield non-zero values for $K_{kj}(\ell)$. Another point to make here is that there are some choices that can be made about how to interpret Equation 4.2. Hansen et al. (2011) chose to use the convention where the sum over m' was from $-(\ell + k - |j|)$ to $(\ell + k - |j|)$ and the j in the subscript of $a_{(\ell+k)(m+j)}$ is set to 0. This effectively means that the only correlations investigated are for consecutive values of ℓ with no change in m which means that,

$$\kappa_{-1,0}(\ell) \simeq \kappa_{1,0}(\ell-1) \simeq \kappa_{-1,1}(\ell) \simeq \kappa_{1,-1}(\ell-1) \simeq \kappa_{-1,-1}(\ell) \simeq \kappa_{1,1}(\ell-1),$$

and,

$$\chi_{-1,0}(\ell) \simeq -\chi_{1,0}(\ell-1) \simeq \chi_{-1,1}(\ell) \simeq -\chi_{1,-1}(\ell-1) \simeq \chi_{-1,-1}(\ell) \simeq -\chi_{1,1}(\ell-1).$$

Here, we investigate $K_{kj}(\ell)$ further to see whether correlations were revealed between adjacent m values as well ℓ, and so in this work the sum over m' is taken to be from $-(\ell+k)$ to $(\ell+k)$ with the j in the $a_{(\ell+k)(m+j)}$ utilised. However many of the results for $K_{kj}(\ell)$ are still closely related i.e.

$$\kappa_{-1,0}(\ell) \simeq \kappa_{1,0}(\ell-1),$$

$$\kappa_{-1,1}(\ell) \simeq \kappa_{1,-1}(\ell-1) \simeq \kappa_{-1,-1}(\ell) \simeq \kappa_{1,1}(\ell-1),$$

$$\chi_{-1,0}(\ell) \simeq -\chi_{1,0}(\ell-1),$$

$$\chi_{-1,1}(\ell) \simeq -\chi_{1,-1}(\ell-1) \simeq \chi_{-1,-1}(\ell) \simeq -\chi_{1,1}(\ell-1).$$

So essentially the value of j has been defined differently here than in Hansen et al. (2011) which gives different results in some cases. To clarify, the results for $j = 0$ in this work match the results for all different values of j quoted in Hansen et al. (2011), and the plots for $|j| = 1$ in this work are new additional results. Therefore, for the rest of this chapter, only examples of $K_{kj}(\ell)$ for $k = 1$ and $j = [0, 1]$ are considered.

4.2.2 Application to Bianchi Simulations of the CMB

Here the normalised cross correlation of the spherical harmonics, $K_{kj}(\ell)$, is calculated for a selection of CMB maps simulated from Bianchi universes as described at the beginning of this chapter. The spherical harmonic amplitudes of the maps were found using HEALPIX and the cross correlation was subsequently calculated using Equations 4.2 to 4.5. The results shown in Figure 4.10 are for Bianchi maps with their main axis aligned with the coordinate system used to decompose the map. There are some strong correlations in $\kappa_{kj}(\ell)$ for Bianchi

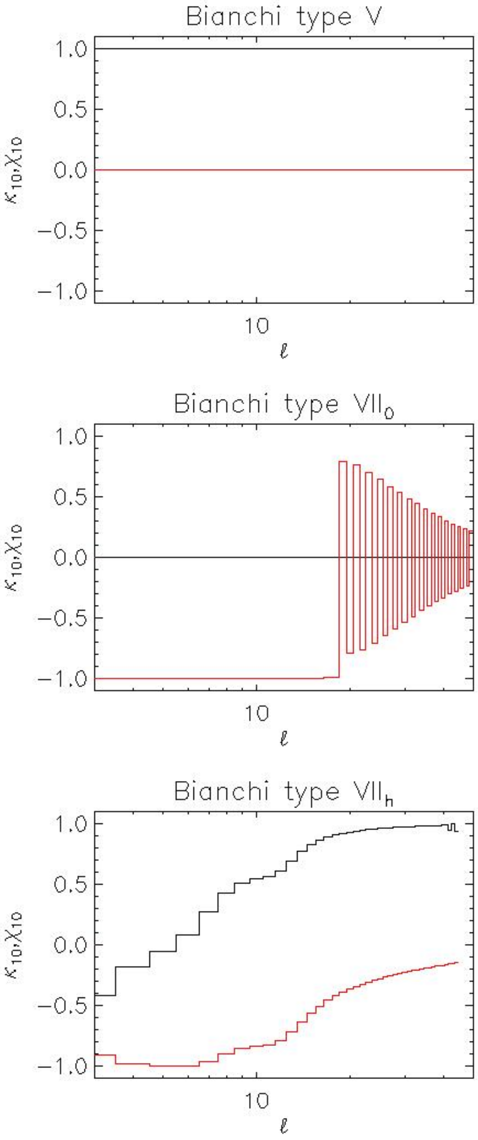

Figure 4.10: *Cross correlation, $K_{kj}(\ell) = \kappa_{kj}(\ell) + i\chi_{kj}(\ell)$, for each of the Bianchi models at redshift $z = 0$ (solid line). Results are for $k = 1$ and $j = 0$ only since for these models which are perfectly aligned with the axis there are no correlations for $j = 1$. $\kappa_{kj}(\ell)$ is shown in black and $\chi_{kj}(\ell)$ is shown in red. Strong correlations are observed in κ_{10} from the Bianchi types V and VII$_h$, whereas strong correlations in χ_{10} are only seen for types VII$_0$ and VII$_h$.*

type V, in $\chi_{kj}(\ell)$ for Bianchi type VII$_0$ and in both $\kappa_{kj}(\ell)$ and $\chi_{kj}(\ell)$ for Bianchi type VII$_h$. If the plots are not aligned with the coordinate axis, for example if they are tilted so as to align with the observed WMAP cold spot, then this results in the cross correlation for $j = 0$ being mixed into the cross correlation for $j = 1$ (see an example for Bianchi type VII$_h$ in Figure 4.11).

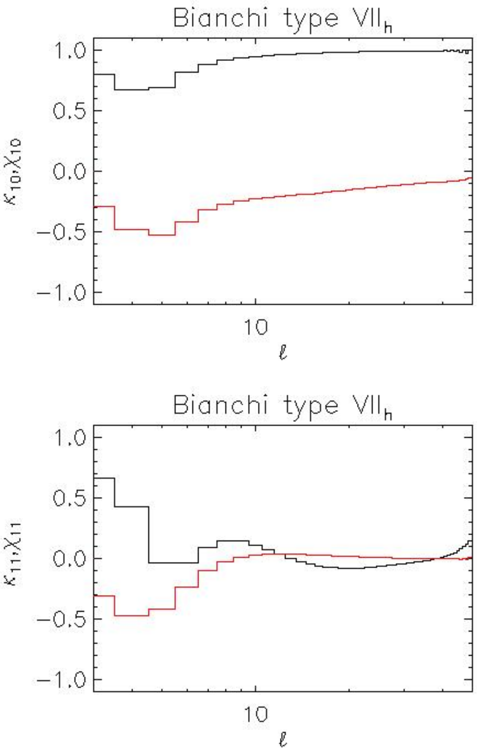

Figure 4.11: *Cross correlation, $K_{kj}(\ell) = \kappa_{kj}(\ell) + i\chi_{kj}(\ell)$, for the Bianchi VII$_h$ model (tilted to align with the cold spot), at redshift $z = 0$ (solid line). Results are for $k = 1$ and $j = [0, 1]$; $\kappa_{kj}(\ell)$ is shown in black and $\chi_{kj}(\ell)$ is shown in red. Strong correlations are still observed for $j = 0$ but now some weaker correlations are observed for $j = 1$ too.*

In other words cross correlations in the $j = 0$ example are being diluted and transfered to the $j = 1$ example. For the sake of keeping the number of plots to the minimum, the rest of the work shall use un-tilted versions of the Bianchi model, but the same principle as seen here applies in other maps if rotated off the axis.

Cross Correlation of Bianchi CMB Maps with Gaussian Fluctuations

The Bianchi CMB maps considered so far do not have stochastic fluctuations superimposed on the deterministic pattern so, whilst the modes are strongly correlated, the associated amplitudes may not be that significant. To understand the dominant correlations, this section considers the Bianchi CMB maps with Gaussianly distributed fluctuations added to them. The Gaussianly distributed fluctuations were taken from simulations by Eriksen et al. (2005) using the standard concordance model assumptions. So as to understand the general affect of adding fluctuations, the average results from 10000 simulations are found. The Bianchi CMB maps were combined with the Gaussianly distributed fluctuation maps at a signal-to-noise ratio of 1:1 as described in Section 4.1.3. Note that the original paper considered a signal-to-noise ratio of 1:3 by standard deviation to be consistent with the variance of the WMAP cold spot but here a signal-to-noise ratio of 1:1 is used to be consistent with the phase analysis in Section 4.1. The cross correlation of each of these combined maps was calculated as described previously; the average $K_{kj}(\ell)$ over all the 10000 combined maps are shown in Figure 4.12 for each of the Bianchi types. In addition, the one σ confidence limits that the cross correlation is consistent with the concordance model are plotted for comparison. These are again calculated from the Gaussian simulations by Eriksen et al. (2005). As in the case without fluctuations, the Bianchi V results show a characteristic deviation from the null hypothesis for κ_{10} but in this case less strongly, particularly at higher values of ℓ. The Bianchi VII$_0$ case shows a strong deviation in the χ_{10} variant, but again it is weaker now there are fluctuations and is mainly localised to small $\ell \leq 10$. The Bianchi VII$_h$ type shows now very weak correlations in χ_{10} but stronger correlations in κ_{10} for $10 \leq \ell \leq 50$.

Cross Correlation of the Bianchi Maps for different redshifts

It is useful to consider how features of the Bianchi models, characterised by $K_{kj}(\ell)$, vary as a function of redshift. The cross correlation was calculated as

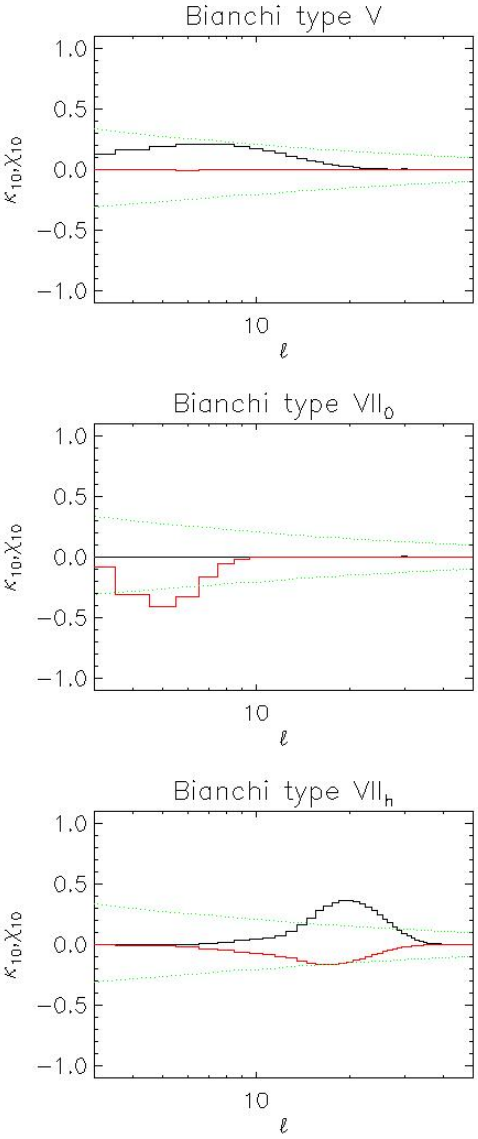

Figure 4.12: *Average* K_{10} *for each of the Bianchi models combined with Gaussianly distributed fluctuations at redshift* $z = 0$ *(solid line).* κ_{10} *is shown in black and* χ_{10} *is in red. The green dotted lines represent the one* σ *confidence limits that the cross correlation is consistent with the concordance model. Again correlations, albeit less strong than seen when there were no fluctuations, are observed in* κ_{10} *for the Bianchi types V and* VII_h*, and in* χ_{10} *for types* VII_0 *and* VII_h*.*

described above for each of the Bianchi maps (with Gaussianly distributed fluctuations) at a selection of different redshifts (see Figure 4.13 - 4.15). Please note that the redshift in these Bianchi simulations is chosen in the initial conditions of the simulations to be arbitrary, so these results are indicative of a trend not conclusive. For all the different Bianchi classifications considered, the characteristic peaks observed in the different variations of the cross correlation, $K_{kj}(\ell)$, all move to lower ℓ and decrease in intensity as redshift increases. For the Bianchi type V, the characteristic peak in $K_{kj}(\ell)$ at no point exceeds the one sigma limits of the concordance model expected variation; the test is unlikely to distinguish a Bianchi V cosmology from the standard concordance cosmology. However for both the VII$_h$ and VII$_0$ cases there are stronger features that could perhaps distinguish them from the standard cosmology. For the Bianchi type VII$_0$ the cross correlation shows a negative deviation from zero in χ_{10} around $\ell \leq 5$. For the Bianchi type VII$_h$, κ_{10} shows a positive deviation from zero around $\ell \leq 20$.

Power Spectrum of the Bianchi maps with Gaussian fluctuations

The power spectrum of the CMB temperature fluctuations is well known from observations such as those from WMAP. Therefore any non-Gaussianity generated by anisotropic universes needs to not affect the power spectrum else they would immediately be eliminated as a possibility. Here the power spectrum of the Bianchi CMB maps (with Gaussianly distributed fluctuations added as described in Section 4.2.2) is considered to check it is consistent with observations. Figure 4.16 shows the average power spectrum of each of the above mentioned Bianchi maps combined with Gaussianly distributed fluctuations. Here there are deviations from the standard model observed. However it is for a small number of ℓ modes. Considering the power spectrum of the CMB as measured by WMAP shown in Figure 2.3 in Section 2; whilst the majority of the data points fit the model well there are deviations from the model, for example at $\ell \sim 40$, $\ell \sim 2$ and $\ell \sim 12$. Therefore the power spectrum does not rule out

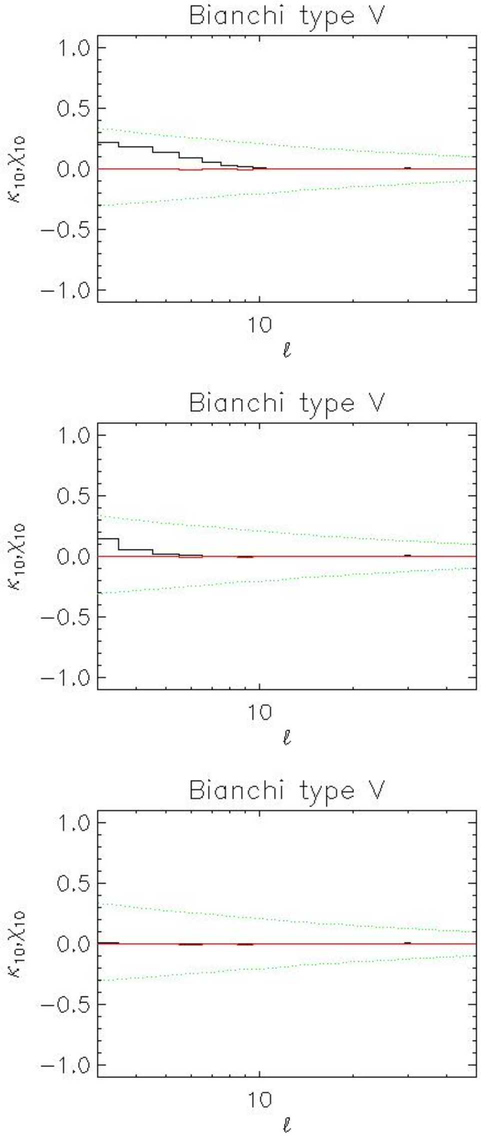

Figure 4.13: *Average* K_{10} *for the Bianchi V model combined with Gaussianly distributed fluctuations at redshifts* $z = 2$*, 6, and 30, top to bottom respectively (solid line).* κ_{10} *is shown in black;* χ_{10} *is in red. The characteristic features for all of the different Bianchi models move to lower* ℓ*, and decrease in intensity, as* z *increases.*

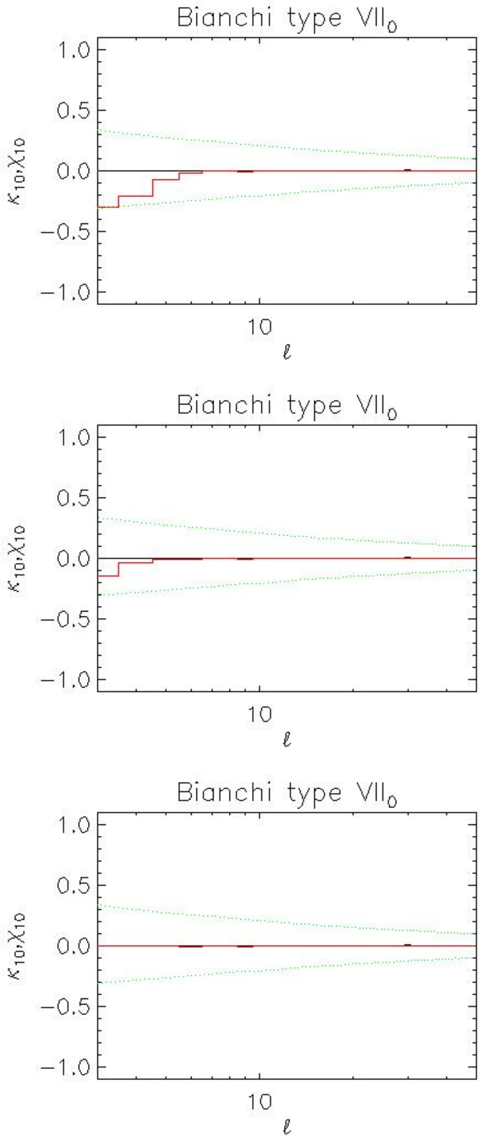

Figure 4.14: *Average* K_{10} *for the Bianchi VII*$_0$ *model combined with Gaussianly distributed fluctuations at redshifts* $z = 2$*, 6, and 30, top to bottom respectively (solid line).* κ_{10} *is shown in black;* χ_{10} *is in red. The characteristic features for all of the different Bianchi models move to lower* ℓ*, and decrease in intensity, as* z *increases.*

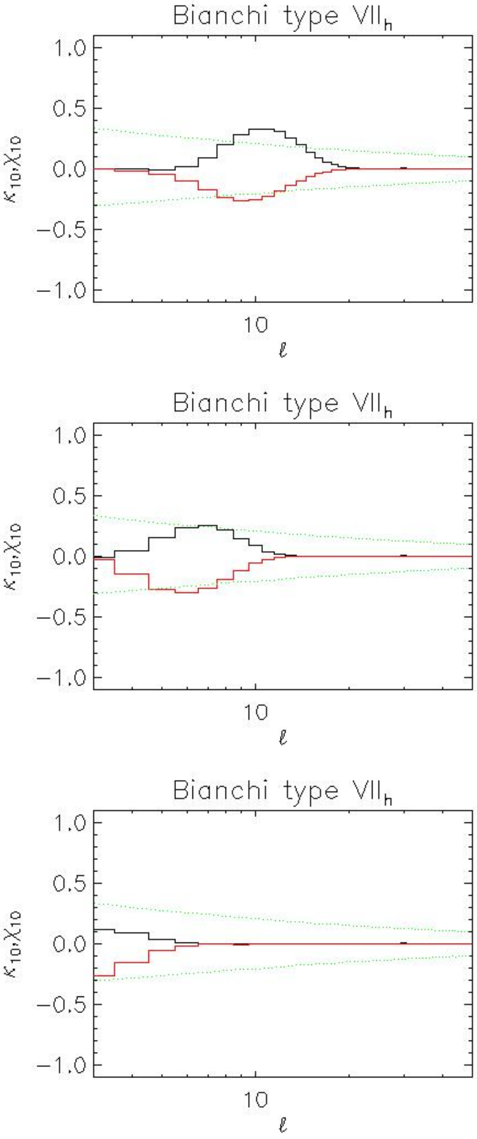

Figure 4.15: *Average* K_{10} *for the Bianchi VII*$_h$ *model combined with Gaussianly distributed fluctuations at redshifts* $z = 2$, *6, and 30, top to bottom respectively (solid line).* κ_{10} *is shown in black;* χ_{10} *is in red. The characteristic features for all of the different Bianchi models move to lower* ℓ, *and decrease in intensity, as* z *increases.*

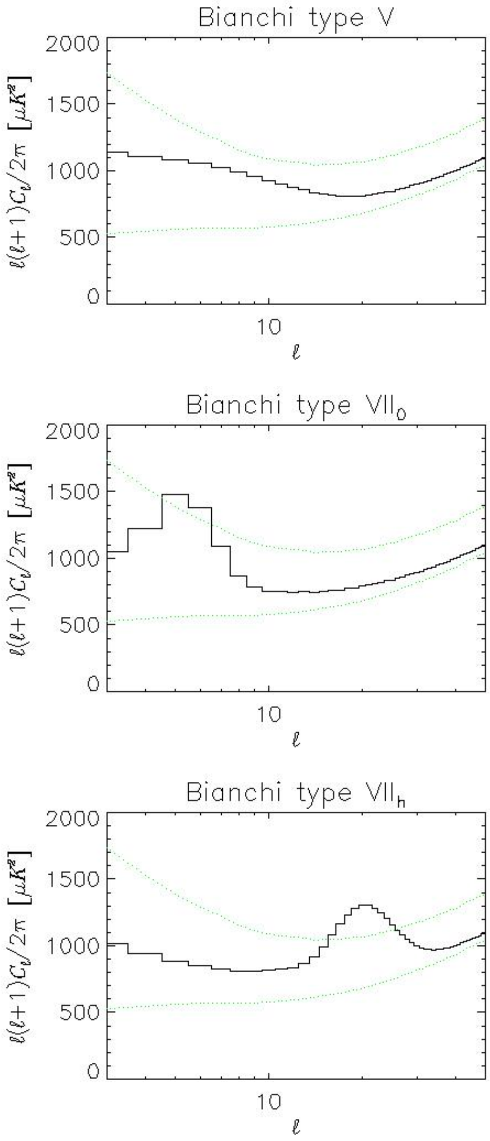

Figure 4.16: *Average power spectrum of Bianchi CMB maps, with Gaussianly distributed fluctuations, at $z = 0$. The green dotted lines represent the one σ confidence limits that the power spectrum is consistent with the concordance model.*

the Bianchi models but it does need to be considered in conjunction with the cross correlation. If there is a deviation in the cross correlation then the power spectrum should show corresponding deviations and vice versa. In provides a useful cross check in this case. The expected deviations from the standard model should occur at the same ℓ in the power spectrum as in the cross correlation.

4.2.3 Applications to Gaussian CMB Maps with Localised Spots

For a comparison to the Bianchi results, this section considers how manually adding a cold spot to Gaussianly distributed simulations would affect the cross correlation (for example see Figure 4.17).

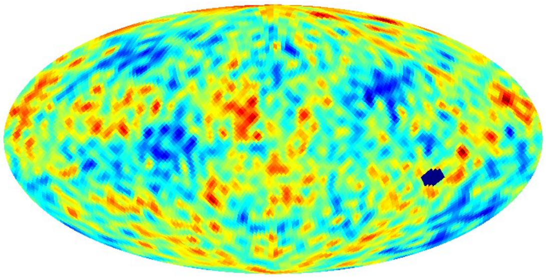

Figure 4.17: *Example CMB map of a concordance model simulation with a cold spot added in the bottom-right.*

The average $K_{kj}(\ell)$ calculated for these examples is shown in Figure 4.18; the results show much weaker correlations for $K_{kj}(\ell)$ compared to that for the Bianchi maps, in this case around $\ell \sim 20$ which corresponds to the size of the spot. Therefore this shows that the normalised cross correlation statistic would not be good at identifying localised small deviations from the standard model such as hot or cold spots.

4.2.4 Application to ILC processed Gaussian simulations

The expectation value of the cross correlation for Gaussianly distributed CMB temperature fluctuations is zero. But even if the primordial CMB signal was

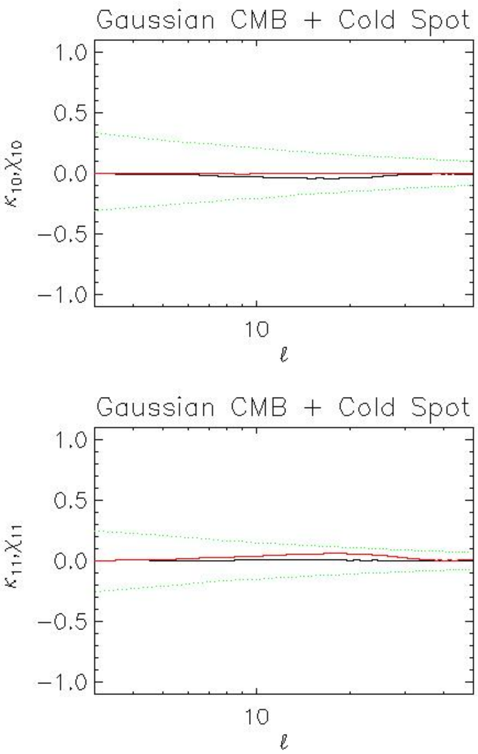

Figure 4.18: *Average normalised cross correlation, $K_{kj}(\ell) = \kappa_{kj}(\ell) + i\chi_{kj}(\ell)$, for 10000 concordance model simulations with a cold spot added (solid line). The results for $\kappa_{kj}(\ell)$ are shown in black and the results for $\chi_{kj}(\ell)$ are in red, where k = 1 and j = [0, 1]. The dotted lines represent the one σ confidence limits that $K_{kj}(\ell)$ is consistent with the concordance model. The correlations, around $\ell \sim 20$, are very weak.*

Gaussian, it has been suggested that the cleaning processes for generating full sky maps might induce correlations into the map (Chiang et al., 2009). To explore this for the cross correlation, it was calculated for the simulated Gaussian CMB maps from Eriksen et al. (2005) which have been processed using the ILC pipeline so as to replicate any systematics that might be generated in the observed maps. To establish an average effect, $K_{kj}(\ell)$ is found for 10000 simulations and the average results are shown in Figure 4.19. The results show that the ILC processing introduces some small correlations at low ℓ, but these correlations are much smaller than the cosmic variance. Therefore the ILC processing will not itself generate significant correlations in the observed ILC maps, although it will still be useful to use these simulations to compare the maps to for a more accurate significance.

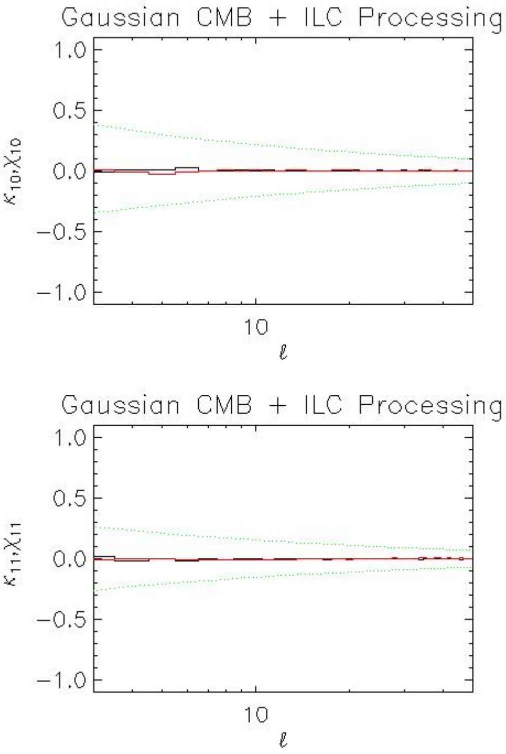

Figure 4.19: *Average cross correlation, $K_{kj}(\ell) = \kappa_{kj}(\ell) + i\chi_{kj}(\ell)$, for 10000 Gaussian simulations with ILC processing (solid line). The results for $\kappa_{kj}(\ell)$ are shown in black and the results for $\chi_{kj}(\ell)$ are in red, where k = 1 and j = [0, 1]. The dotted lines represent the one σ confidence limits that $K_{kj}(\ell)$ is consistent with the concordance model. Negligible deviations from the concordance model are observed.*

4.2.5 Application to WMAP CMB Data

This section is finished off by calculating the cross correlation of the spherical harmonic coefficients for actual observational data from the WMAP 5 year ILC map (Hinshaw et al., 2009). The results are plotted in Figure 4.20 as a function of ℓ, for the range $0 \leq \ell \leq 100$. For comparison we also include single realisations of the results from Figure 4.12 in Figure 4.21. From the plots in Figure 4.20 it is clear that the WMAP 5 year ILC map shows few significant correlations i.e. very few of the $K_{kj}(\ell)$ results lie outside the expected variation. There are some instances of $K_{kj}(\ell)$ going above the level of cosmic variance (e.g. at $\ell = 5, 9, 17$) although only for one ℓ at a time not for a grouping of ℓ's. To be significant at all we would expect these deviations to be consistent with deviations in the power spectrum. Whilst the power spectrum of the CMB as measured by WMAP and

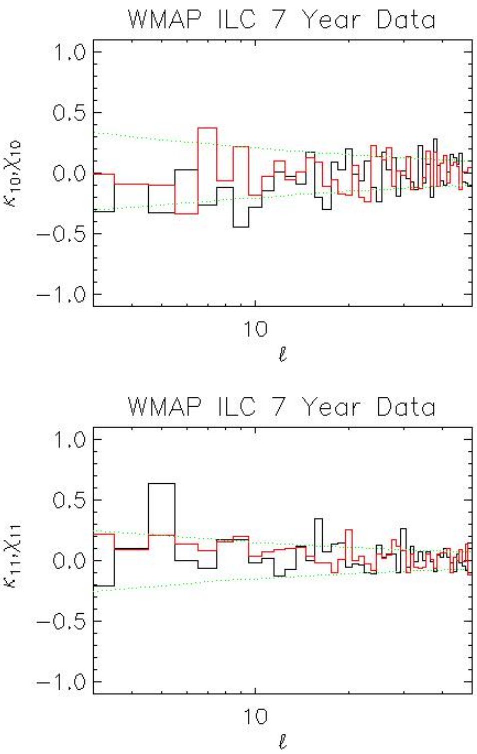

Figure 4.20: *The cross correlation,* $K_{kj}(\ell) = \kappa_{kj}(\ell) + i\chi_{kj}(\ell)$*, for the 5 year WMAP ILC map. The results for* $\kappa_{kj}(\ell)$ *are shown in black and the results for* $\chi_{kj}(\ell)$ *are in red, where* $k = 1$ *and* $j = [0, 1]$*. The dotted lines represent the one* σ *confidence limits that* $K_{kj}(\ell)$ *is consistent with the concordance model.*

other experiments is well known and shown to be in good agreement with the concordance model, there remain some anomalies. For example see Figure 2.3 in the Introduction which shows the power spectrum as measured by WMAP. In this we see there are positive deviations from the model for $\ell \sim 40$ and negative deviations for $\ell \sim 2$ and $\ell \sim 12$. The deviations observed in the power spectrum are at different ℓ to the deviations in the cross correlation. However this could be because the cross correlation is analysed at a different rotation from the preferred axis and hence the signal has been diluted.

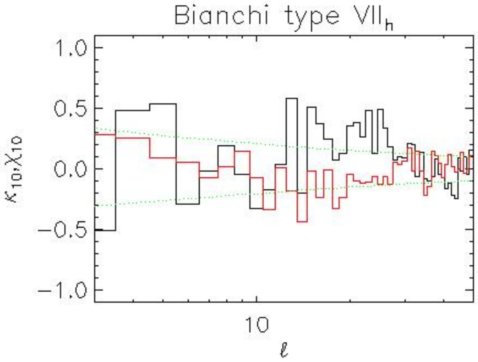

Figure 4.21: *The cross correlation, $K_{kj}(\ell) = \kappa_{kj}(\ell) + i\chi_{kj}(\ell)$, for an individual realisation of a Bianchi VII$_h$ model combined with Gaussianly distributed fluctuations at redshift $z = 0$. The results for $\kappa_{kj}(\ell)$ are shown in black and the results for $\chi_{kj}(\ell)$ are in red, where k = 1 and j = [0, 1]. The dotted lines represent the one σ confidence limits that $K_{kj}(\ell)$ is consistent with the concordance model.*

Chapter 5

Conclusions for CMB Analyses

Observations of the temperature anisotropies in the cosmic microwave background (CMB), particularly those from the Wilkinson Microwave Anisotropy Probe (Bennett et al., 2003, WMAP), provide good evidence for the very successful *concordance* cosmological model. An essential ingredient of this model is the assumption that the primordial density fluctuations that seeded the formation of galaxies and large-scale structure were statistically homogeneous, isotropic and Gaussianly distributed. Analysis of currently available WMAP data provide strong limits on the level of non-Gaussianity (Komatsu et al., 2003; Spergel et al., 2007). On the other hand, Yadav and Wandelt (2008) reported a detection of primordial non-Gaussianity at greater than 99.5% significance. Further detailed analyses of non-Gaussianity in the CMB are clearly necessary in order to reconcile and understand the various constraints and claimed detections.

Non-Gaussianity can be classified into two different types: primordial or secondary (i.e. present at the last scattering surface or occurring since then). In Chapter 3 we presented a simple statistical analysis looking for secondary anomalies due to correlations induced via the process of cleaning out Galactic foreground emission. This was based on properties of the *zonal* modes of full sky CMB maps, i.e. those modes aligned parallel to the Galactic plane. An application of the test to a selection of CMB maps generated via different cleaning methods gave interesting results. At the 95 per cent level no significant anomalies appear in the WMAP ILC maps (Bennett et al., 2003; Jarosik et al., 2007; Hinshaw et al., 2009), but there seems to be a significant tendency in some other maps (Tegmark et al., 2003; Kim et al., 2008) to have zonal modes with systematically lower amplitudes than would be expected in the concordance model. Intriguingly, the maps that provide the most significant departures from the behaviour expected under the null hypothesis are those based on later issues of the WMAP data.

An additional investigation considering the variance of the zonal amplitudes similarly found them to be significantly smaller than expected. It is worth highlighting that the variance of amplitudes over all modes is also low, although at less significance than the zonal modes, which is consistent with previous work by Monteserin et al. (2008), although they considered the combined WMAP Q+V+W map[1]. Further analysis shows that once the zonal modes are removed from the selection, the low variance of the amplitudes is no longer significant. This strongly suggests that the zonal modes have a notable contribution towards this low variance.

The final part of the analysis of the zonal modes considered the distribution of maxima and minima in zonal maps, which showed the extrema values are not as extreme as would be expected. Again, the earlier finding that the non-Gaussian anomalies increase with later releases of data are reinforced.

Of course the different CMB maps themselves are *not* statistically independent. Indeed, if the cleaning processes involved were perfect then they would all be identical. The varying results found for the different maps are attributable to the nature of the zonal modes and their extra sensitivity to structures associ-

---

[1] Note this map is slightly different to the ILC map which we used; it is a noise weighted average of the foreground cleaned maps from each of the Q, V and W bands (see Section 2.2).

ated with the Galactic plane. The appearance of significant anomalies in some maps rather than others is not a statistical fluke but is clear evidence that some cleaning methods leave artifacts in the distribution of mode amplitudes.

It must be noted that the probabilities quoted of around 98 to 99 percent are not overwhelming; the results obtained are indicative rather than decisive. This is not surprising, given the relatively small number of modes used. However, when the analysis was repeated for a coordinate system aligned with the Ecliptic, rather than Galactic plane, no significant results were found at all. This lends further credence to the interpretation of the outcome of this analysis in terms of an effect related to over-subtraction of Galactic emission consistent with work by Chiang et al. (2009) and Naselsky et al. (2008), which becomes increasingly pronounced with each data release. However, a more definitive result will have to wait until more detailed foreground subtraction can be attempted, such as will be the case with the *Planck* satellite.

The second half of this CMB part, in Chapter 4, explored a specific form of primordial non-Gaussianity; characterising the large-scale temperature patterns in CMB maps generated by anisotropic Bianchi type V, VII_h and VII_0 universes. The ultimate purpose of investigating this behaviour was to find ways of quantifying the global properties of the pattern produced in order to isolate the effect of anisotropy from that of non-Gaussianity. Note that the non-Gaussianity talked about here is not related to a stochastic field; there are no fluctuations in the Bianchi simulations.

The first method described in Section 4.1, a *phase analysis*, has not been used to quantify many alternative situations to the concordance model. The phases of the spherical harmonic coefficients provide a generic way of looking at correlations in harmonic space that could arise from anisotropy or non-Gaussianity. While it is a potential strength of the approach that phase correlations will not just be useful for identifying anisotropies specific to the Bianchi models but in theory any isotropy introduced to the CMB, it could also prove a weakness in that more general methods such as this may lack the power to discriminate very specific models. In this work though the phase correlations identified in the

Bianchi maps were much stronger than first expected; given the generic nature of this descriptor it was not expected to yield good results. In addition to this, the strong correlations for the Bianchi types V and VII$_h$ were found to be robust to both rotation and moderate noise. However using the same methods on the WMAP 5 year ILC map showed little evidence of non-Gaussianity when analysed in Galactic coordinates. Further investigation of the map rotated to the previously reported preferred axis showed a more significant result result, but not above the 95 percent level.

Whilst the phase analysis has been shown to be an effective method for identifying non-Gaussianities, it identified them in harmonic (as opposed to pixel) space. Therefore it is difficult to say whether any of the anomalies identified this way are due to anisotropy or more generic non-Gaussianity. This anticipated limit of the analytical method is something that was further considered in the original paper (Sung et al., 2010) and addressed by also using multipole vectors; these are based on spherical harmonic phases but give a resulting direction in pixel space.

The alternative method described in Section 4.2, uses the *cross correlation of spherical harmonic coefficients* to look for global anisotropy. This method is similar to the phase analysis in that it relies very simply on the assumption of a stationary Gaussian field generating no cross correlation and hence the result is very robust. This also has the advantage over the phase analysis that it contains additional information about the amplitudes and therefore should be less susceptible to noise (i.e. phases can be correlated but if their associated amplitudes are weak then they will easily be masked by any amount of noise). This analysis quantified the characteristic features of this measure specific to the anisotropic Bianchi cosmological models in contrast to the concordance model. The correlations found were not highly significant, but the cross correlation was effective for identifying the VII$_0$ and VII$_h$ models given a reasonable amount of noise and rotation. Analysis of the WMAP 5 year ILC map showed little evidence for these cosmological models although they are not ruled out.

The patterns in the CMB temperature generated by Bianchi models represent fairly extreme departures from the standard framework so it is no real surprise that they register strongly in the descriptors used. However, both analyses have involved only a relatively small number of modes, so the fact that quantifiable effects are seen emerging is very encouraging. Each of the measures considered are based on very simple assumptions and seem to be most sensitive to different aspects of non-Gaussianity; focused features, as seen in Bianchi types V and VII_h, are detected in the case of the phases and twisting features, as seen in Bianchi types VII_0 and VII_h, are detected in the case of the cross correlation. Therefore the two methods (phases analysis and cross correlation) considered here provide complementary information.

Both methods consider the Bianchi simulations with fluctuations, or in other words Gaussian ‘noise’, added to them at an arbitrary ratio. Any ratio is valid; the greater the amount of fluctuations the greater the degradation of the underlying correlations. Therefore a null detection does not preclude the possibility of a Bianchi universe. This investigation was more aimed at developing effective methods for identifying the Bianchi models, as well as other forms of anisotropy, and both methods have shown promise.

These analyses of the Bianchi simulations were restricted to large angular scales because this is where the resulting patterns in the temperature map are strongest. Therefore the significance of any potential non-Gaussianity found is always limited by the small number of modes. To increase the robustness of these tests, more information needs to be utilised. This work simply considers the temperature maps, but the Bianchi models also generate very characteristic signals in the polarisation maps (Pontzen, 2009; Sung and Coles, 2011). Further investigation of this polarisation data is required to characterise it and develop quantifiable measures. Therefore, in future, by using the temperature and polarisation results in conjunction hopefully a more robust test can be achieved.

These have been very preliminary analyses, aimed at establishing whether the diagnostics considered are *in principle* capable of uncovering evidence of underlying anomalies in CMB data. It remains the case that the standard cosmological model is a good fit to a huge range of observational data. Nevertheless, it is important that tools are developed that are sufficiently sensitive to hunt efficiently for possible anomalies in the next generation of observations. There are many ways that the CMB temperature pattern could be anomalous other than those which have been studied here. Just as there are many ways a distribution can be non-Gaussian, so are there also many ways a fluctuation field can be non-stationary. Testing for departures from the standard model will require not one but a battery of statistical techniques each sensitive to particular aspects of the distribution.

Part II

Galaxy Clustering

Chapter 6

Introduction to Galaxy Clustering

The large scale structure observed in the local Universe is a complex pattern of long filaments, made up from clusters of galaxies, intersected with large empty voids (see Figure 6.1). In contrast the early Universe had a matter distribution

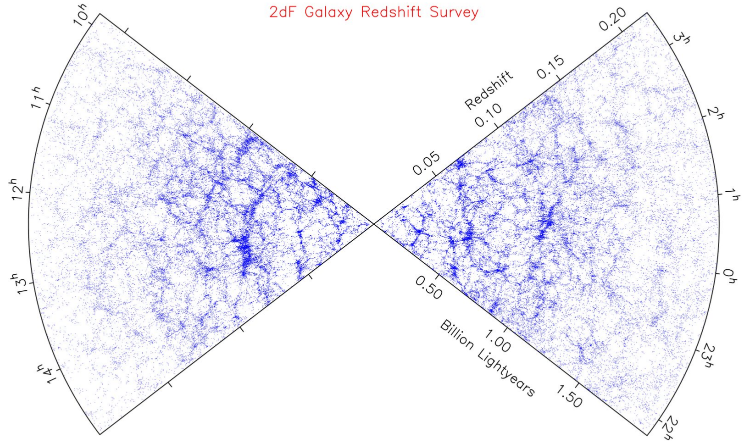

Figure 6.1: *Observed distribution of galaxies from the 2dF Galaxy Redshift Survey (Colless et al., 2001). Galaxies are displayed as a function of look-back time or redshift, and right ascension. Structure is clearly seen in filaments and voids.*

of relatively consistent density. However, very small temperature fluctuations are observed in the cosmic microwave background (CMB, see Part I) as shown in Figure 6.2.

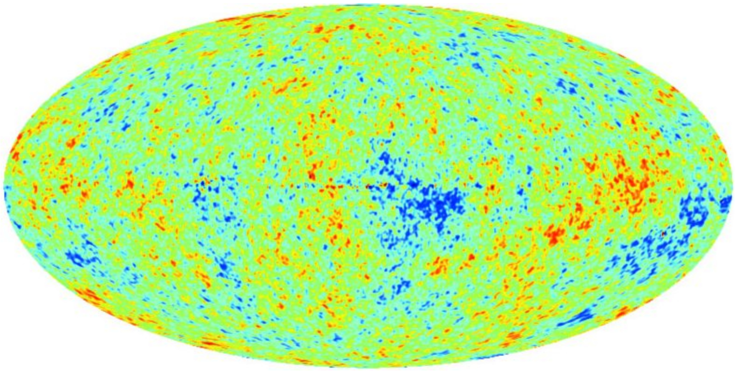

Figure 6.2: *ILC 7-year temperature map of the CMB as measured by WMAP (Jarosik et al., 2011). Note, the temperature has a direct linear relationship with the matter density, so this is effectively a map of the matter density variation over the full sky. The amplitude of the fluctuations is of the order* 10^{-5} *that of the mean temperature value.*

These temperature fluctuations suggest corresponding variations in the matter density in the early Universe. Models of galaxy formation provide strong evidence that these matter fluctuations are the seeds that ultimately led to the formation of large scale structure (Peebles, 1980). This chapter gives a brief overview of the current theories describing the processes governing the growth of this observed structure.

6.1 Standard Concordance Cosmological model

The current standard *concordance* cosmological model (see Chapter 1) is made up of a framework of different parts, all of which are ongoing areas of study. Together they form the current accepted model to describe how the total matter distribution in the Universe evolves over time.

6.1.1 Background cosmology

The background cosmology was discussed in detail in Chapter 1, but to summarise the following are assumed in the concordance model: Einstein's General Relativity (see Section 1.4), Friedman Robertson Walker metric (see Section 1.4.2), flat geometry (see Section1.4.4) and cosmological constant (see Section 1.6.3).

6.1.2 Initial fluctuation spectrum

The concordance model assumes *adiabatic* density fluctuations in the early Universe which are independent of scale. Therefore the initial linear power spectrum, $P_0(k)$, is assumed to be given by a featureless power law and the distribution of fluctuations is assumed to be *Gaussian*.

6.1.3 Dark Matter

So far we have talked about matter fluctuations but what do we mean by 'matter'? Prior to the 1930s, matter meant 'baryonic matter' i.e. the material seen in planets, stars and galaxies. However, by the 1970s there was mounting evidence for a huge amount of matter, both on local and large scales, which could not be accounted for simply by baryonic matter. In the 1980s an alternative solution was proposed in the form of *dark matter*; a form of matter that only interacts gravitationally. Nowadays, (cold) dark matter is commonly accepted to make up $\sim 83\%$ of our Universe's matter density, with the remaining matter being baryonic. Bergstrom (2009) gives a comprehensive review of the different candidate particles for dark matter, along with the current experimental status; we shall not go into any further detail here.

6.1.4 Growth law

The previous items summarised above form the 'initial conditions' for structure formation. Now we need a methodology for describing the subsequent evolution of these fluctuations. The growth of the perturbations due to the cosmology discussed above is described by the *growth law.*

To derive this we need to consider the equations of motions for a perturbation, where $\rho = \rho_0 + \delta\rho$, $p = p_0 + \delta p$, $\mathbf{v} = 0 + \delta\mathbf{v}$, $\Phi = \Phi_0 + \delta\Phi$ and $\nabla\Phi = 0$. Therefore ignoring higher orders of the small terms, Equations 1.19, 1.20 and 1.21 become,

$$\frac{\partial\delta\rho}{\partial t} + \rho_0\nabla.\delta\mathbf{v} = 0, \tag{6.1}$$

$$\frac{\partial\delta\mathbf{v}}{\partial t} = -\frac{\nabla\delta p}{\rho_0} - \nabla\delta\Phi, \tag{6.2}$$

and,

$$\nabla^2\delta\Phi = 4\pi G\delta\rho. \tag{6.3}$$

We also consider the equation of state $p = p(\rho)$ (Section 1.3) which becomes,

$$\delta p = p(\rho_0 + \delta\rho) = p(\rho_0) + \left(\frac{dp}{d\rho}\right)_{\rho=\rho_0} \delta\rho + ...,$$

and since $c_s^2 = \left(\frac{dp}{d\rho}\right)_{\rho=\rho_0}$ this gives,

$$\delta p \simeq c_s^2\delta\rho. \tag{6.4}$$

Taking the covariant derivative of Equation 6.2 and then substituting in Equation 6.4 and 6.3 gives,

$$\frac{\partial\nabla.\delta\mathbf{v}}{\partial t} = -\frac{c_s^2\nabla^2\delta\rho}{\rho_0} - 4\pi G\delta\rho.$$

Taking the time derivative of Equation 6.1, substituting this into the previous expression and rearranging this gives,

$$\left(\frac{\partial^2}{\partial t^2} - c_s^2\nabla^2 - 4\pi G\rho_0\right)\delta\rho = 0. \tag{6.5}$$

To solve this equation we look for solutions in the form of plane waves,

$$\delta\rho = \exp[i(\mathbf{k}.\mathbf{x} - \omega t)], \tag{6.6}$$

which substituted into Equation 6.5 gives,

$$\omega^2 - c_s^2k^2 + 4\pi G\rho_0 = 0. \tag{6.7}$$

This *dispersion relation* then has two types of solution depending on whether the wavelength ($\lambda = 2\pi/k$) is less than or greater than the *Jeans' length* (Jeans, 1902),

$$\lambda_J = c_s\sqrt{\frac{\pi}{G\rho}}. \tag{6.8}$$

If $\lambda < \lambda_J$ then the $\delta\rho$ solution from Equation 6.7 will oscillate. Else if $\lambda > \lambda_J$ then there are two solutions of exponential form, one of which decays and one of which grows. Overtime this decaying solution disappears and so is not considered later in the evolution of large scale structure.

The growing mode solution will have a dependence on t so to get a dependence on z you need to consider the background geometry. Later we see that this is approximated using a fitting function to numerical simulations. Whilst the amplitude of the modes is small then linear theory can be assumed and so the power spectrum grows proportionally to the growing mode i.e.

$$\frac{P(k,t)}{P(k,t_0)} = \frac{D_+^2(k,t)}{D_+^2(k,t_0)}. \tag{6.9}$$

6.1.5 Transfer function

However the growth of perturbations is affected by other phenomena such as dissipative and radiation related processes. These are accounted for by the *transfer function.*

The growth of perturbations can be inhibited by the presence of *radiation pressure*; pressure only opposes gravity effectively for scales smaller then the *Jeans' length* (Equation 6.8). For both baryonic and dark matter, λ_J is significant before recombination, but becomes small afterwards and so the growth of perturbations is then not inhibited by radiation pressure. However, the Jeans' length scale for massive neutrinos remains significant after recombination. Because the neutrino density does not trace the matter perturbations (due to its small interaction cross section) it has the effect of reducing the gravitational collapse of matter over-densities.

In addition to radiation pressure, dissipative processes affect the growth of small scale perturbations. The first of these is *free streaming* which means that random particle motions cause over-densities to disperse. In the early Universe this initially happens at all scales; as the Universe cools, and the particles slow down to non-relativistic speeds, this dispersion gradually becomes less effective at smaller and smaller scales. This process switches off when particles become non-relativistic. For massive neutrinos this occurs around recombination; for dark matter this occurs well before.

In the case of baryonic matter a similar process occurs known as *silk damping* (Silk, 1967) which causes fluctuations to be smoothed when matter and radiation are coupled. When the Universe was still hot enough for matter to be ionised, matter existed in a plasma coupled with radiation (i.e. adiabatic fluctuations). According to the *Jeans' Instability Theory*, fluctuations smaller than the *Jeans' length* will oscillate like acoustic waves. These oscillations continue because when the density is increased by a compression, the pressure generated by photons causes a restoring force. When the fluctuation size falls below the mean free path of a photon, then the photons are no longer providing this balancing pressure and just move out to other areas resulting in a smoothing effect.

6.1.6 Growth of linear fluctuations

The growth law and the transfer function combine to provide a *linear* description of the growth of fluctuations i.e. they describe the process well when the over/under densities are still small. This linear description provides a good approximation of the matter distribution over large angular scales, as demonstrated in Figure 6.3, which shows some images of the matter distribution from numerical simulations, as a function of time. The images start at the oldest image (earliest time) which shows primordial density fluctuations, to the youngest image which corresponds to the matter distribution at the present time which shows a clear structure of voids and filaments. We see that over large scales, the matter distribution stays the same i.e. we can see the brightest point of the

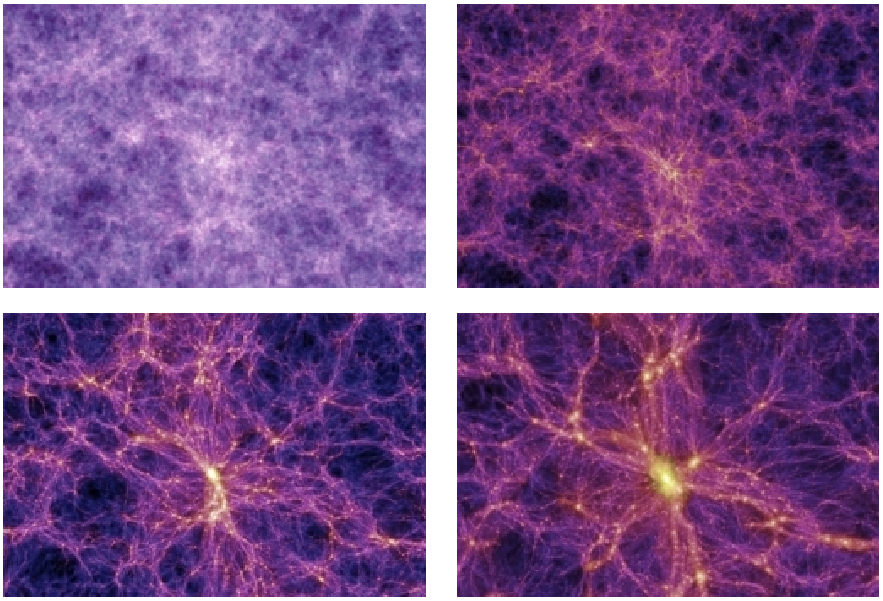

Figure 6.3: *The images show four snapshots from the Millennium (numerical) simulation of the matter density at different redshifts (Springel et al., 2005). Each of the images shows the same co-moving area. Higher matter densities are shown as bright white/yellow areas, and areas of under-density are shown by darker purple colours. The different redshifts of the images are 18.3 (top left), 5.7 (top right), 1.4 (bottom left) and 0.0 (bottom right).*

image at early times is roughly in the centre of the image. It stays the brightest spot through each of the following images. Similarly, the darker patches tend to stay dark. But on smaller scales it is a different story. The image is becoming more ‘focused’. Matter is being pulled into the areas which are denser, under the influence of gravity, and leaving empty voids in between. So the linear approximation is good at large scales, but to describe the non-linear process in areas experiencing strong gravitational effects we need further explanation.

6.1.7 Non-linear evolution

The linear regime of gravitational instability begins to break down when the magnitude of the density fluctuations reaches the same order as the background density. Interactions are at this point *weakly non-linear*; the amplitudes of the density perturbations may remain low, but gradually the distribution of fluctuations will become more and more non-Gaussian as perturbations become correlated. When perturbations become so massive that they separate from the Hubble expansion of the Universe to become bound structures this is known as the *strong non-linear regime.* These bound structures are known as *dark matter halos.*

The non-linear growth of fluctuations is too involved to derive a simple analytic expression as in the linear case. However there are currently many different methods being developed for approximating this non-linear evolution, both via modeling and analytical methods. These are discussed further later in Section 6.2.

6.1.8 Formation of galaxies

So far we have discussed the overall matter density; here the evolution of just the baryonic matter is considered i.e. the matter than can be directly observed. As the fluctuations in the dark matter start to collapse, they separate from the Hubble expansion. Initially the baryonic gas is coupled to the radiation and is prevented from collapsing with the dark matter by radiation pressure, but after recombination the normal matter starts to follow the dark matter density. Although the formation of the dark matter halos is now fairly well understood, the formation of galaxies remains more elusive. When these bound structures or halos form, the dark matter can no longer collapse as it cannot dissipate its energy. Dark matter differs from normal baryonic matter because it only interacts with other matter via gravity. It has a very small interaction cross section so it is effectively able to pass right through other matter. As the dark matter halo collapses, the motion of the baryonic matter separates from that of the dark matter, since the path of normal matter will be modified when it hits other baryonic matter.

This collapsing baryonic matter then starts to form galaxies, albeit via processes which are not yet well understood. There are different types of galaxies which can be roughly categorised in the three groups; *disks*, *ellipticals* and *irregulars*. It is thought that disks are 'naturally' formed via the collapse process; as the baryonic matter flows in to the centre of the halo, it rotates faster and faster. This rotation causes the matter to flatten into a disk.

These disk galaxies are still influenced by gravity, and this causes them to be pulled together causing them the merge and grow into larger galaxies. Mergers of galaxies are often associated with bursts of new star formation; these powerful ultra-luminous galaxies are known as *starburst* galaxies. The larger galaxies observed are likely to have formed via a number of mergers. It is thought that elliptical galaxies are likely to result from a large number of mergers and have therefore gone through so many bursts of star formation that the majority of the baryonic matter is no longer in gas but locked up in (older) stars.

That was a very brief background of the current view of galaxy formation. However, research in the field is ongoing and hence there in considerable literature on the subject. Much of this is not relevant to this analysis though and therefore we will not go into any further depth here. Benson (2010) provide a good review.

Dark matter is not directly observable; only the baryonic matter can be observed, which is mostly visible in galaxies. Given current theoretical models predict that galaxies form and evolve in cold dark matter (CDM) halos, the galaxies consequently tend to trace the distribution of the total mass, although the manner in which they do this may be biased (Kaiser, 1984; Coles, 1993; Mo and White, 1996). In principle, therefore, once the bias is allowed for, it is possible to use measurements of the clustering of galaxies to determine the clustering properties of the underlying total matter. Therefore observations of galaxies are of much interest in this work.

There are several different types of galaxies/objects that are commonly used for studies of large scale structure. Some of the objects that are common enough to yield good statistics, and bright enough to be seen at the large distances, are;

- *Lyman Break galaxies* are bright star-burst galaxies detected using the Lyman break. Radiation at wavelengths less than the Lyman limit at 91.2 nm is almost completely absorbed by the neutral gas found surrounding star-forming regions of galaxies. This break in amplitude is used as a selection technique to find star-forming galaxies at high redshift.
- *Sub-millimeter galaxies* are detected via strong emission in the submillimeter wavelength band (100 to 1000 μm). These are early, strongly star-forming galaxies; by their nature are obscured by dust. Therefore a high proportion of the radiation emitted from these galaxies is absorbed by this dust and re-emitted at longer wavelengths.
- *Quasars* are associated with the very bright centres of some galaxies. They are thought to be be driven by some very high energy processes, probably related to the accretion of matter onto a supermassive black hole.

6.2 Measures of Galaxy Clustering

Much work has gone into understanding the global properties of large scale structure. These are interesting firstly because, as described above, the matter distribution is sensitive to the cosmological parameters and therefore serves as an additional independent test of the standard cosmological model. Secondly, clustering is also sensitive to other more physical parameters such as the minimum mass a halo with one galaxy has (in your set of observable objects). Parameters like this are of interest because they can be used to better understand the processes involved in the formation of galaxies.

6.2.1 Modeling methodologies

Various different methods have been used to try to model the galaxy distribution, some of which are discussed here.

Numerical

Some groups have built large numerical simulations to simulate large scale structure formation, examples of which we will refer to later in Chapter 7. Initially it was only possible to simulate the distribution of dark matter on large scales, but this was key to comfirming the link between anisotropis in the CMB with the growth of large scale structure. However, recent advances in computational power have meant that numerical simulations are now increasingly used to simulate formation of galaxies. However these galaxy simulations are still limited because the processes involved in galaxy formation occur at scales much below that of the current resolution possible in these simulations.

Semi-Analytical

Because of the limitations of numerical simulations discussed above, semi-analytical methods are often employed which use numerical simulations to generate the dark matter distribution, then add analytic approximations for subsequent evolution of the baryonic matter (Springel et al., 2005). Statistics can be extracted from these simulations and compared to those from our own Universe.
These simulations can be used to characterise different scenarios from just the standard model e.g. by varying cosmological parameters or properties of dark matter. Also they can be used to investigate trends over different parameters (e.g. redshift, mass) where the parameters are well known unlike in observations. However these simulations cannot yet describe all the observations and also require a huge amount of computing power to generate, which can be limiting when trying to probe the smaller scales (Baugh, 2006).

Analytical

Analytical methods consider purely theoretical approximations for the overall statistical properties of the galaxy distribution. They combine a number of different components expected to contribute to galaxy clustering. This work considers just analytical methods and they will be discussed in further detail in

Chapter 7.

6.2.2 Observational Data

We have said that analytical, and semi-analytical, methods can be used to estimate the global statistics of galaxy clustering. These statistics can be compared to observations of our own Universe. A steadily increasing number of surveys of large-scale galaxy clustering are now becoming available. In the *optical* waveband there are projects such as the UKIDSS Ultra Deep Survey (Lawrence et al., 2007; Hartley et al., 2010) and the Sloan Digital Sky Survey (SDSS) Redshift Survey (York, 2000; Connolly et al., 2002; Ross and Brunner, 2009) which are being used to extract information on clustering over large areas and as a function of redshift. *X-ray* observations, which are used to locate quasars, have been performed by the ROSAT-NEP (Mullis et al., 2004; Miyaji et al., 2011) and subsequently the XMM-COSMOS (Hasinger et al., 2007; Gilli et al., 2009) surveys. *Radio* observations were made by the FIRST and NRAO VLA surveys (Cress et al., 1996; Blake and Wall, 2002); Infra-red observations have been performed by Blast (Viero et al., 2009), 2dF (Percival et al., 2001), and UKIDSS (Hartley et al., 2010). Most recently, sub-millimeter observations have been made by the *Herschel* Space Observatory (Pilbratt et al., 2010; Maddox et al., 2010; Cooray et al., 2010).

6.2.3 Statistics of Galaxy Clustering

Several different statistics are commonly used, which we discuss further below.

Projected Angular Correlation Function

One standard way to quantify the distribution of galaxies is using a *correlation function.* A *two point correlation function* effectively measures the likelihood of finding a pair of galaxies at a given physical separation, above what you would expect compared to a random distribution. It is the Fourier transform of the clustering power spectrum. However, measurements of distances to individual objects in many surveys are quite uncertain, so it is usual to use the *angular correlation function.* This describes the excess probability that a galaxy will be observed to be within a given angle of a randomly selected galaxy on the sky. Or in other words, it considers the increased likelihood of finding a pair of galaxies at a given angular separation compared to what you would expect if there were no clustering.

The correlation function can be measured from data (i.e. maps of 2-dimensional galaxy fields) by counting pairs of galaxies in the data as a function of the angular separation. These counts are then compared to counts of pairs of galaxies in a random distribution in a map of similar size and with similar selection effects. The two sets of counts can then be combined to estimate the angular galaxy correlation function, $\omega(\theta)$. There are several different methods for doing this; one commonly-used approach is the Landy & Szalay estimator (Landy and Szalay, 1993) which calculates the correlation function as,

$$\omega(\theta) = \frac{DD - 2DR + RR}{RR}, \tag{6.10}$$

where DD is the number of data-data pairs, DR is the number of data-random pairs and RR is the number of random-random pairs, all at a given angular separation θ.

Two point correlation function / power spectrum

However, with the increasing availability of observations with associated redshift data, work has begun to try to calculate the non-projected two point correlation function. The best redshift information comes from *spectroscopic redshifts* which compares well known lines in spectra in objects to accurately extract the shift in the spectrum. However spectroscopic redshifts are not always possible without good spectra from objects. Therefore in recent years *photometric redshifts* have been used much more prolifically - these use relatively wide filters to calculate a redshift.

Whilst extracting this ‘true’ two point correlation is more complex, the resulting correlation amplitude is stronger/clearer than the projected one since it is not diluted out over redshift. Many groups have been using this additional redshift information to try to investigate the redshift evolution of the clustering signal, for example that by Dodelson et al. (2002) who considered the power spectrum. Of course higher order correlation functions can be and are now being investigated e.g. Verde et al. (2000) and Marin (2010). These are useful as complementary investigations to the two point correlation function to find parameters such as the bias parameters. In addition the higher orders hold additional useful information which may provide evidence of non-Gaussianity (McBride et al., 2010).

6.2.4 Cross correlation

So far the correlations discussed here have been auto correlations i.e. the cross correlation of the galaxy distribution with itself. However, as has previously been mentioned, there are now ongoing galaxy surveys over various wavelengths. These different wavebands provide different views of the matter distribution. Newly forming, dusty galaxies emit radiation in the submillimeter, whereas starburst galaxies emit predominately in the optical range. Cross correlating the observations between the different wavebands will show how the clustering for the particular objects compares e.g. whether they are stronger/weaker or have different evolution. This will then yield clues on the different formation/evolution of the particular objects.

One project that is very relevant to this is the Galaxy and Mass Assembly (GAMA) survey (Driver et al., 2009) which will collate imaging of the same part of the sky in the optical, infrared, radio, ultraviolet and submillimeter wavelength ranges, along with some spectroscopic/photometric redshift observations. So far only the optical and sub-mm observations have been completed (using data from the SDSS and *Herschel* surveys) but the results so far look promising. For example, Guo et al. (2011) used the large optical galaxy catalog with well known spectroscopic redshifts from the SDSS to cross correlate with the smaller submillimeter galaxy catalog from *Herschel* to show a weaker clustering in the submillimeter. Of course there are other cross correlation studies other than those related to GAMA such as work by Krumpe et al. (2010) who cross correlated X-ray (ROSAT All-Sky Survey) and optical (SDSS) sources, and Jarvis et al. (2010) who cross correlated submillimeter (*Herschel*) and radio (NRAO VLA) observations.

The work in this second part of my thesis, and published in Short and Coles (2011), considers a selection of analytical models for calculating the projected two point angular auto correlation function which are described in Chapter 7. The results of these models are compared with each other and against recent observations from the *Herschel* Space Observatory in Chapter 8. Chapter 9 investigates the effect of varying the different free parameters on the galaxy angular correlation function. Chapter 10 summarises the findings and discusses potential future avenues for investigation.

Chapter 7

Analytical Galaxy Angular Correlation models

This chapter describes a selection of analytical models for the galaxy angular correlation function, which is a standard way of measuring galaxy clustering (see Section 6.2.3). In Chapter 8 the results from these models will be compared to observations from the *Herschel* Space Observatory.

The main focus of the work is to investigate a particular non-linear fitting function model, but for comparison, two additional models are also considered here. First, a *linear approximation* (see Section 6.1.4) is considered which essentially assumes that the matter distribution today remains the same as it was at early times, such as that observed in the CMB. This is a very simplified approach, so it is included to show what increase in accuracy the fitting function approach for the non-linear approximation gives.

The other model considered is a *non-linear approximation* (see Section 6.1.7) which takes into account strong gravitational effects on small scale clustering. The *Halo model* is a complex and comprehensive model that takes in the account all the components expected to contribute to the angular correlation function, such as the halo dark matter profile. This Halo model is included as a comparison at the other extreme of complexity to the linear model to see if the fitting function method can do anywhere near as well.

The *fitting function* version is also a non-linear method but uses a fairly simple approach based on work by Peacock and Dodds (1996) who found an approximation for the non-linear matter angular correlation combined with a bias approximation by Moscardini et al. (1998).

All the models start by finding the dark matter power spectrum and then the galaxy correlation function is subsequently found by effectively Fourier transforming that power spectrum.

7.1 Linear galaxy angular correlation function

7.1.1 Initial linear matter power spectrum

As described in Section 6.1.2, the initial linear power spectrum, $P_0(k)$, is assumed to be a featureless power law, i.e.

$$P_0(k) = Ak^n, \tag{7.1}$$

where k is the linear *wave number*, and n is known as the *spectral index*.

7.1.2 Transfer function

The *transfer function*, $T(k)$, describes how the initial linear power spectrum evolves due to radiative and dissipative processes (see Section 6.1.5). The processed power spectrum[1], $P(k)$, which results following these additional processes (not gravity), is calculated as;

$$P(k) = P_0(k)T^2(k) = AT^2(k)k^n. \tag{7.2}$$

There are several different versions of the transfer function which have been derived using fitting formulae to numerical simulations, for example those proposed by Bond and Efstathiou (1984) and Bardeen et al. (1986). The currently

---

[1] Note that this power spectrum for the cold dark matter and does not include Baryon Acoustic Oscillations.

favoured transfer function is that derived by Eisenstein and Hu (1999), which is more complex than the others and but gives a better fit to observations. This model takes into account cold dark matter, dark energy and neutrino densities, and is described in detail below.

The redshift at matter-radiation equality is given by,

$$z_{eq} = 2.5 \times 10^4 \Omega_m(0) h^2 \Theta_{\mathrm{cmb}}^{-4}, \tag{7.3}$$

where the temperature of the cosmic microwave background, T_{cmb}, is parametrised as $\Theta_{\mathrm{cmb}} = T_{\mathrm{cmb}}/2.7\mathrm{K}$. We also define the redshift where ionised atoms were released from their coupling to radiation at recombination (see Section 2.1) using a fitting formulae result by Hu and Sugiyama (1996) and later refined by Eisenstein and Hu (1998),

$$z_{\mathrm{drag}} = \frac{1291(\Omega_m(0)h^2)^{0.251}(1 + b_1(\Omega_m(0)h^2)^{b_2})}{1 + 0.659(\Omega_m(0)h^2)^{0.828}}, \tag{7.4}$$

where

$$b_1 = 0.313(\Omega_m(0)h^2)^{-0.419}(1 + 0.607(\Omega_m(0)h^2)^{0.674}), \tag{7.5}$$

and

$$b_2 = 0.238(\Omega_m(0)h^2)^{0.223}. \tag{7.6}$$

So the relative expansion between the matter-radiation equality and the point where baryons are released from the drag of photons at recombination can be expressed as,

$$y_{\mathrm{drag}} = \frac{1 + z_{\mathrm{eq}}}{1 + z_{\mathrm{drag}}}. \tag{7.7}$$

The sound horizon, which is defined here as the co-moving distance a sound wave can propergate before this drag at recombination, is approximated as (Eisenstein and Hu, 1998),

$$s = \frac{44.5 \log\left(\frac{9.83}{\Omega_m(0)h^2}\right)}{\sqrt{1 + 10(\Omega_m(0)h^2)^{0.75}}} \ \mathrm{Mpc}. \tag{7.8}$$

Following Eisenstein and Hu (1998) we define this factor to approximate the suppression of the growth rate by the presence of neutrinos,

$$p_c = \frac{1}{4}(5 - \sqrt{1 + 24f_c}), \tag{7.9}$$

where f_c is the ratio of the cold dark matter density to the total mass density i.e. $f_c = [\Omega_m(0) - \Omega_b(0) - \Omega_\nu(0)]/\Omega_m(0)$. A factor for the suppression of power is (Eisenstein and Hu, 1999),

$$\alpha_\nu = \frac{(1 - f_b)(5 - 2p_c)(1 - 0.553f_b + 0.126f_b^3)}{5(1 + y_{\rm drag})^{p_c}}\Big[1 + \frac{p_c(1 + \frac{1}{7(3-4p_c)})}{2(1 + y_{\rm drag})}\Big]. \tag{7.10}$$

where f_b is the ratio of the baryon density to the total mass density i.e. $f_b = \Omega_b(0)/\Omega_m(0)$. The effective wavenumber is then,

$$q_{\rm eff} = \frac{k\Theta_{\rm cmb}^2}{\Gamma_{\rm eff}}, \tag{7.11}$$

where the effective shape parameter is given by,

$$\Gamma_{\rm eff} = \Omega_m(0)h^2\Big(\sqrt{\alpha_\nu} + \frac{1 - \sqrt{\alpha_\nu}}{1 + (0.43ks)^4}\Big). \tag{7.12}$$

This is all then incorporated into a function with a form found by Eisenstein and Hu (1998) to give:

$$T(k) = \frac{L}{L + Cq_{\rm eff}^2}, \tag{7.13}$$

where

$$C = 14.4 + \frac{325}{1 + 60.5q_{\rm eff}^{1.11}}, \tag{7.14}$$

$$L = \ln\left[\exp(1) + 1.84\beta_c\sqrt{\alpha_\nu}q_{\rm eff}\right], \tag{7.15}$$

and

$$\beta_c = \frac{1}{1 - 0.949f_b}. \tag{7.16}$$

7.1.3 Growing mode of linear perturbations

As described in Section 6.1.4, the growth of perturbations due to gravity is described by the growing mode of linear perturbations, $D_+(z)$, where

$$P(k,z) = P(k,0)D_+^2(z), \tag{7.17}$$

and $D_+(z)$ is given by,

$$D_+(z) = \frac{g(z)}{g(0)(1+z)}, \tag{7.18}$$

where $g(z)$ is the linear growth factor. Carroll et al. (1992) derived an approximation formula for $g(z)$ in a universe with a cosmological constant,

$$g(z) = \frac{5}{2}\Omega_m(z)\Big[\Omega_m^{4/7}(z) - \Omega_\Lambda(z) + \Big(1+\frac{\Omega_m(z)}{2}\Big)\Big(1+\frac{\Omega_\Lambda(z)}{70}\Big)\Big]^{-1}, \tag{7.19}$$

where $\Omega_\Lambda(z)$ is the dark energy density parameter. Note $\Omega_\Lambda(z)$ and $\Omega_m(z)$ are calculated as in Eisenstein and Hu (1999) as,

$$\Omega_\Lambda(z) = \frac{\Omega_\Lambda(0)}{\Omega_m(0)(1+z)^3 + (1-\Omega_m(0)-\Omega_\Lambda(0))(1+z)^2 + \Omega_\Lambda(0)}, \tag{7.20}$$

and

$$\Omega_m(z) = \frac{\Omega_m(0)(1+z)^3}{\Omega_m(0)(1+z)^3 + (1-\Omega_m(0)-\Omega_\Lambda(0))(1+z)^2 + \Omega_\Lambda(0)}. \tag{7.21}$$

7.1.4 Linear dark matter power spectrum

So to summarise, the above terminology can be used to approximate the linear dark matter power spectrum,

$$P(k,z) = P_0(k)T^2(k)D_+^2(z), \tag{7.22}$$

where $P_0(k)$ is the initial linear power spectrum (Section 7.1.1), $T^2(k)$ is the transfer function (Section 7.1.2) and D_+ is the growing mode of linear perturbations (Section 7.1.3).

This calculation requires a number of input parameters. In this work, results from the WMAP experiment are used to estimate the cosmological parameters

Parameter	Description	Estimate
n	Spectral index	0.96
σ_8	Variance of the field on the scale of 8 Mpc h^{-1}	0.812
h	Hubble parameter	0.705
$\Omega_m(z=0)$	Matter density^1	0.274
$\Omega_b(z=0)$	Baryon density^1	0.046
$\Omega_\Lambda(z=0)$	Dark energy density^1	0.726
N_ν	Number of different neutrino species	3.04
$\Omega_\nu(z=0)$	Neutrino density^1	0.014
T_{cmb}	Temperature of the CMB (K)	2.725
$\delta_c(z=0)$	Critical density for spherical collapse^2	1.686

Table 7.1: *Cosmological parameters required as inputs to the angular correlation models, along with current estimate from WMAP (Komatsu et al., 2009). These parameters will be used in all future calculations unless otherwise stated.* [1] *All densities are in units of the critical density.* [2] *In an Einstein-de Sitter cosmology, $\delta_c(0) = 1.686$; this value is also a good approximation for the concordance model (Reed et al., 2007).*

so that the analysis can concentrate on the effect of the halo parameters. These cosmological parameters are detailed in Table 7.1.

Note, the power spectrum is also often expressed as a dimensionless quantity,

$$\Delta(k) = \frac{k^3}{2\pi^2} P(k). \tag{7.23}$$

7.1.5 Linear dark matter angular correlation function

The *two-point correlation function* (Section 6.2.3) is a measure of the probability of finding two galaxies of a given distance r away from each other in 3D space, above what you would expect in a random distribution. It is defined as (Peacock, 1999),

$$\xi(r) = \int_0^\infty \Delta^2(k) \frac{\sin kr}{kr} \frac{dk}{k}. \tag{7.24}$$

When you take into account the distribution of the objects observed as a function of redshift, $\mathcal{N}(z)$, the observed correlation function becomes,

$$\xi_{\mathrm{obs}}(r) = \mathrm{N}^{-2} \int_Z \mathcal{N}^2(z)\xi(r,z)dz. \tag{7.25}$$

where N is a normalising factor for $\mathcal{N}(z)$. However, given that distance measurements are often uncertain, the observed angular correlation function, $\omega_{\mathrm{obs}}(\theta)$, is frequently used. This is defined as,

$$\omega_{\mathrm{obs}}(\theta) = \mathrm{N}^{-2} \int_Z G^2(z)\mathcal{N}^2(z) \int_{-\infty}^{\infty} \xi(r(u,\theta,z),z)dudz, \tag{7.26}$$

where

$$r(u,\theta,z) = a_0\sqrt{u^2 + x^2(z)\theta^2}, \tag{7.27}$$

$$G(z) = \left(\frac{dx}{dz}\right)^{-1}, \tag{7.28}$$

$$x(z) = \frac{c}{H_0 a_0} \int_0^z \frac{dz'}{[(1+z')^2(1+\Omega_m z') - z'(2+z')\Omega_\Lambda]^{-1/2}}, \tag{7.29}$$

and a_0 is chosen[2] to be 1 Mpc h^{-1} (with the assumption that $\Omega_t(0) = 1$). Note, this calculation of ω_{obs} from $\Delta^2(k)$ is valid for all linear/non-linear and dark matter/galaxy cases as long as the appropriate power spectrum required is used.

7.1.6 Linear galaxy angular correlation function

We have previously discussed how the distribution of galaxies is a good tracer of the distribution of the total mass. However, it is not a perfect tracer and the discrepancy between the galaxy power spectrum (Δ^2_{GAL}) and the dark matter power spectrum (Δ^2_{DM}) is known as *bias*.

---

[2] Since in this work the curvature k is chosen to be zero then a_0, which is the characteristic curvature length scale, has no specific value and can therefore be selected to be anything. In this instance we are looking at galaxy clustering and so a Mpc length scale is appropriate.

It was shown by Coles et al. (1999) that over large-scales the galaxy bias is well approximated by a linear factor ($b^2_{\rm eff}$), i.e.

$$\Delta^2_{\rm GAL}(k,z) = b^2_{\rm eff}\Delta^2_{\rm DM}(k,z). \tag{7.30}$$

We use an approximation for the bias by Mo and White (1996) who suggested the following relation between the bias and mass density fluctuations,

$$b(z,m) = 1 + \frac{1}{\delta_c}\Big[\frac{\delta_c^2}{\sigma^2_{\rm lin}(z,m)} - 1\Big], \tag{7.31}$$

where δ_c is the critical linear density required for collapse. $\sigma^2_{\rm lin}$ is linear variance in the smoothed density field (Equation 7.58).

The effective bias ($b_{\rm eff}$), i.e. the bias averaged over the distribution of masses, can then be derived as,

$$b_{\rm eff}(z) = \frac{1}{N_{ST}}\int\limits_{\ln(M_{\rm min})}^{\infty} b(z,m)n(z,m)d\ln m, \tag{7.32}$$

where n is the mass function given by Sheth and Tormen (1999) and discussed further in Section 7.3.3. N_{ST} is the normalisation factor for n.

$M_{\rm min}$ is the *minimum halo mass*, or the minimum mass that a halo can have to host a galaxy. This is because either baryonic material is prevented from entering, or reaching the centre, of a halo below a certain mass; or baryonic matter is in these dark matter halos but is prevented from cooling and collapsing. The exact physical process responsible for this is not yet known, although one suggestion that has been hypothesised is that it could be due photo-heating associated with re-ionisation of the Universe. This increases the temperature of the interstellar gas and prevents the cooling and collapse of low mass galaxies. Evidence for this cutoff has been provided by work such as that by Bouche et al. (2010) who used numerical simulations to show that without this cutoff relations between the mass, the star formation rate and the rotational velocity of galaxies can not be matched as well. Here we make the fairly simple assumption that $M_{\rm min}$ is a sharp cutoff so that no halos below this mass form galaxies, although we do discuss alternatives later in Chapter 9.

7.2 Non-linear galaxy angular correlation (Fitting function)

The angular correlation function in the non-linear regime can be calculated in the same way as described in Sections 7.1.5 - 7.1.6, but rather than using the linear power spectrum we need to use an estimate for the non-linear power spectrum.

A number of different approaches have been used to calculate this. Here we consider a *fitting function* which has been derived from numerical simulations of structure formation. Following on from work by Hamilton et al. (1991), Peacock and Dodds (1996) showed that the non-linear power spectrum and wavenumber can be approximated directly from the linear counter part as,

$$\Delta^2_{\rm nl}(k_{\rm nl}, z) = \mathcal{F}[\Delta^2_{\rm lin}(k_{\rm lin}, z)], \tag{7.33}$$

$$k_{\rm nl} = k_{\rm lin}[1 + \Delta^2(k_{\rm nl}, z)]^{1/3}, \tag{7.34}$$

where

$$\mathcal{F}(x) = x\Big[\frac{1 + B\beta x + [Ax]^{\alpha\beta}}{1 + ([Ax]^{\alpha} g^3(z)/[Vx^{1/2}])^{\beta}}\Big]^{1/\beta}, \tag{7.35}$$

and

$$A = \quad 0.482(1 + n/3)^{-0.947}, \tag{7.36}$$

$$B = \quad 0.226(1 + n/3)^{-1.778}, \tag{7.37}$$

$$\alpha = \quad 3.310(1 + n/3)^{-0.244}, \tag{7.38}$$

$$\beta = \quad 0.862(1 + n/3)^{-0.287}, \tag{7.39}$$

$$V = \quad 11.55(1 + n/3)^{-0.423}. \tag{7.40}$$

This approximation uses a scale dependent value for the spectral index n where,

$$n_{\rm eff}(k) = \frac{d\ln P}{d\ln k}(k = k/2). \tag{7.41}$$

This method calculates the non-linear dark matter power spectrum. The non-linear galaxy power spectrum can then be calculated using the same bias as described in Section 7.1.6.

7.3 Non-linear galaxy angular correlation (Halo model)

Alternatively, the Halo model can be used to calculate the non-linear galaxy power spectrum, which can then be used as described in Sections 7.1.5 - 7.1.6 to calculate the corresponding angular correlation function.

The Halo model considers two contributions to the galaxy power spectrum; the one-halo (1h) term which takes into account contributions from galaxies in the same halo, and the two-halo (2h) term which considers galaxies which are in separate halos, i.e.

$$\Delta^2(k,r) = \Delta^2_{\mathrm{1h}}(k,r) + \Delta^2_{\mathrm{2h}}(k,r), \tag{7.42}$$

where k is wavenumber and r is the scale-length. The one-halo term is then further subdivided; halos are assumed to have just one (or zero) central galaxies and then all other galaxies in the halo are known as satellite galaxies. The full expressions for each of these central-satellite pairs (cs) and satellite-satellite pairs (ss) terms are,

$$\Delta^2_{\mathrm{cs}}(k,r) = \frac{k^3}{2\pi^2 n_g^2} \int\limits_{M_{\mathrm{vir}}(r)}^{\infty} 2N_{\mathrm{cen}}(m)N_{\mathrm{sat}}(m)u(k,m)n(m)dm, \tag{7.43}$$

and,

$$\Delta^2_{\mathrm{ss}}(k) = \frac{k^3}{2\pi^2 n_g^2} \int\limits_{0}^{\infty} dm N_{\mathrm{cen}}(m)N^2_{\mathrm{sat}}(m)u^2(k,m)n(m). \tag{7.44}$$

The two-halo term is expressed as,

$$\begin{aligned}\Delta^2_{\mathrm{2h}}(k,r) = \frac{\Delta^2_{\mathrm{lin}}(k)}{[n'_g(r)]^2} &\int\limits_{0}^{M_{\mathrm{lim1}}(r)} N_{\mathrm{tot}}(m)b(m,r)u(k,m)n(m)dm \\ \times &\int\limits_{0}^{M_{\mathrm{lim2}}(r)} N_{\mathrm{tot}}(m)b(m,r)u(k,m)n(m)dm.\end{aligned} \tag{7.45}$$

Methods for approximating all the required terms in each of these expressions are detailed in the following sections. Section 7.3.1 details limits and normalisation factors such as M_{vir}, M_{lim1}, M_{lim2}, n_g and n'_g. $u(k,m)$ is the normalised

Fourier transform of the halo dark matter density profile and is detailed in Section 7.3.2. $n(m)$ is the mass function and is detailed in Section 7.3.3. $b(m, r)$ is the scale dependent bias and is detailed in Section 7.3.4. The number of galaxies, $N_{\rm cen}(m)$, $N_{\rm sat}(m)$ and $N_{\rm tot}(m)$, are estimated using the *halo occupation distribution* model (Section 7.3.5).

7.3.1 Limits and normalisation factors

n_g and n'_g are simply both normalisation factors and are defined as,

$$n_g = \int_0^\infty N_{\rm tot}(m) n(m) dm, \tag{7.46}$$

and

$$[n'_g(r)]^2 = \int_0^{M_{\rm lim1}(r)} N_{\rm tot}(m) n(m) dm \times \int_0^{M_{\rm lim2}(r)} N_{\rm tot}(m) n(m) dm. \tag{7.47}$$

The $M_{\rm lim}(r)$ values are the mass limits due to *halo exclusion* (i.e. halos are spatially exclusive) and are defined in Tinker et al. (2005) as the values that satisfy,

$$R_{\rm vir}(M_{\rm lim1}) = r - R_{\rm vir}(M_{\rm min}), \tag{7.48}$$

and

$$R_{\rm vir}(M_{\rm lim2}) = r - R_{\rm vir}(M_{\rm sat}). \tag{7.49}$$

Finally, the *virial mass* is defined here as the mass at which the mean density inside a given radius is 200 times[3] that of the background density, i.e.

$$\frac{3M_{\rm vir}}{4\pi r_{\rm vir}^3} = 200\rho_c\Omega_m(0). \tag{7.50}$$

---

[3] We use a factor of 200 here, as is commonly used by the community, following estimates from early numerical simulations such as that by Cole and Lacey (1996).

Substituting in the definition of the critical density ρ_c (see Equation 1.4.4) we get,

$$M_{\rm vir}(r) = \frac{100 H_0^2 \Omega_m(0) r_{\rm vir}^3}{G}. \tag{7.51}$$

7.3.2 Normalised Fourier transform of the halo density profile

The normalised Fourier transform of the dark matter distribution within a halo of mass m, is defined as,

$$u(\mathbf{k}|m) = \frac{\int \rho(\mathbf{x}|m) e^{i\mathbf{k}.\mathbf{x}} d^3\mathbf{x}}{\int \rho(\mathbf{x}|m) d^3\mathbf{x}},$$

where ρ is the density profile of the dark matter halo, and $\mathbf{x}$ is the distance from the centre of the halo. If we assume spherically symmetric profiles, truncated at the virial radius, this becomes,

$$u(\mathbf{k}|m) = \int_0^{2\pi} \int_0^{pi} \int_0^{r_{vir}} \frac{\rho(r|m) e^{i\mathbf{k}.\mathbf{r}}}{m} dr d\theta d\phi = \int_{-1}^{1} \int_0^{r_{vir}} \frac{2\pi \rho(r|m) e^{ikr\mu}}{m} dr d\mu$$

where $\mu = \cos\theta$, so

$$u(k|m) = \int_0^{r_{vir}} 4\pi r^2 \frac{\sin kr}{kr} \frac{\rho(r|m)}{m} dr. \tag{7.52}$$

Halo dark matter density profile

There are several different approximations for the density profile. Here we use the density profile proposed by Navarro et al. (1997), which we shall abbreviate as the NFW profile, who used numerical simulations of hierarchical clustering to fit a profile of the form,

$$\rho(r) = \frac{\rho_s}{(\frac{r}{r_s})(1 + \frac{r}{r_s})^2}, \tag{7.53}$$

where r_s is the scale radius and ρ_s is the density at r_s. The scale radius is defined as,

$$c = \frac{r_{\rm vir}}{r_s}, \tag{7.54}$$

where $r_{\rm vir}$ is the virial radius (Equation 7.51) and c is known as the concentration, which is discussed in Section 7.3.2.

ρ_s can also be found as a function of the concentration by substituting the expression for $\rho(r)$ into the definition of the mass to get;

$$m = \int \rho(\mathbf{x}) d^3\mathbf{x} = \int_0^{r_{\rm vir}} 4\pi r^2 \rho(r) dr = \int_0^{r_{\rm vir}} 4\pi r^2 \frac{\rho_s}{(\frac{r}{r_s})(1+\frac{r}{r_s})^2} dr.$$

So using the identity,

$$\int \frac{x}{(ax+b)^2} = \frac{b}{a^2(ax+b)} + \frac{1}{a^2} \ln|ax+b| + c,$$

with $x = r$, $a = \frac{1}{r_s}$, and $b = 1$ we get,

$$m = 4\pi\rho_s r_s^3 \Big[\ln(1+c) - \frac{c}{1+c} \Big]. \tag{7.55}$$

Halo concentration

Following work by Bullock et al. (2001), the concentration can be approximated using the following expression for the NFW dark matter density profile,

$$\bar{c}(m,z) \simeq \frac{9}{1+z} \Big(\frac{m}{m_\star(z)} \Big)^{-0.13}, \tag{7.56}$$

where $m_\star(z)$ is characteristic mass scale at which $\nu(m_\star, z) = 1$. Here,

$$\nu(m,z) = \frac{\delta_c^2}{\sigma^2(m,z)}. \tag{7.57}$$

where δ_c is the critical density required for spherical collapse, and $\sigma^2(m)$ is the variance in the initial density fluctuation field.

Variance of the initial density fluctuation field

The variance in the initial density fluctuation field is calculated as,

$$\sigma^2(m) = \int \Delta^2_{\rm lin}(k) |W(kR)|^2 \frac{dk}{k}, \tag{7.58}$$

where $W(kR)$ is the window function, and R is the scale over which the distribution is smoothed with a filter,

$$R = \Big(\frac{3m}{4\pi\rho_c\Omega_m} \Big)^{1/3}. \tag{7.59}$$

The window function has different values depending on the filter used. Here we use a Top Hat filter (i.e. a step function of radius) so,

$$W(kR) = \frac{3}{(kR)^3}[\sin(kR) - kR\cos(kR)]. \tag{7.60}$$

7.3.3 Mass function

The mass function (n) gives the differential number density of halos we would expect to find for a given cosmology as a function of mass. The approximation we use here for the mass function is that first proposed by Press and Schechter (1974),

$$\frac{mn(m)}{\bar{\rho}} = f(\nu)\frac{d\nu}{dm}, \tag{7.61}$$

where $\nu = \delta_c^2/\sigma^2(m)$.

Press and Schechter (1974) had their own approximation for $f(\nu)$ but here we use the approximation by Sheth and Tormen (1999),

$$\nu f(\nu) = A\left(1 + (q\nu)^p\right)\left(\frac{q\nu}{2\pi}\right)^{1/2}\exp(-q\nu/2), \tag{7.62}$$

where $p = -0.3$ and $q = 0.707$. Also note that the number density needs to be normalised, i.e.

$$\int n(m)dm = 1 \tag{7.63}$$

Figure 7.1 shows the resulting distribution of the number of halos for a selection of different redshifts. We see that the peak number of halos occurs at increasing mass as the redshift decreases, which is consistent with a *hierarchical* model of galaxy formation where many smaller objects gradually merge together to build larger objects.

7.3.4 Scale dependent bias

The bias used here was calculated by Tinker et al. (2005) from numerical simulations as,

$$b(m,r) = b(m)\sqrt{\frac{(1 + 1.17\xi(r))^{1.49}}{(1 + 0.69\xi(r))^{2.09}}}, \tag{7.64}$$

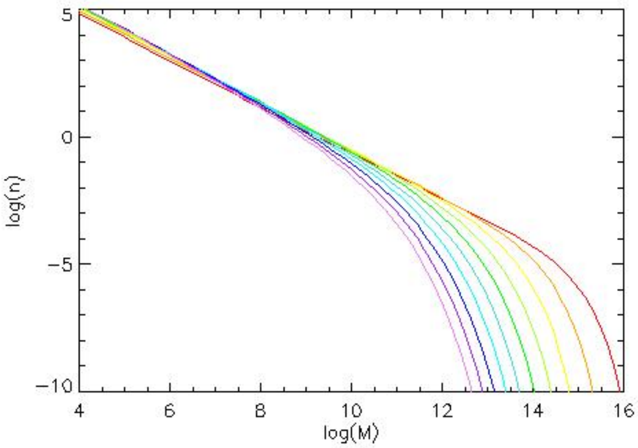

Figure 7.1: *The differential number of halos as a function of mass M ($M_\odot h^{-1}$) using the Sheth-Tormen approximation, for a different redshifts $z = 0 - 7$ (for red - violet respectively).*

where,

$$b(m) = 1 + \frac{1}{\delta_c\sqrt{a}}\Big[\sqrt{a}a\nu + \frac{\sqrt{a}}{2}(a\nu)^{1-c} - \frac{(a\nu)^c}{(a\nu)^c + b(1-c)(1-c/2)}\Big], \quad (7.65)$$

with $a = 0.707$, $b = 0.5$ and $c = 0.6$. ξ is the dark matter cross correlation.

Dark matter cross correlation

The dimensionless non-linear dark matter power spectrum, as approximated by the Halo model, can again be written as,

$$\Delta^2_{\rm DM}(k) = \Delta^2_{\rm 1h}(k) + \Delta^2_{\rm 2h}(k), \quad (7.66)$$

where $\Delta^2_{\rm 1h}(k)$ is the contribution from correlations within the same halo and $\Delta^2_{\rm 2h}(k)$ is the contribution from correlations between separate halos (Cooray and Sheth, 2002),

$$\Delta^2_{\rm 1h}(k) = \frac{k^3}{2\pi^2}\int\Big(\frac{m}{\bar{\rho}}\Big)^2 n(m)|u(k|m)|^2 dm, \quad (7.67)$$

$$\Delta^2_{\rm 2h}(k) = \Delta^2_{\rm lin}(k)\Big[\int b(m)\Big(\frac{m}{\bar{\rho}}\Big)n(m)u(k|m)dm\Big]^2. \quad (7.68)$$

$n(m, z)$ is the co-moving number density as defined in Equation 7.61, $u(k|m)$ is the normalised Fourier transform of the dark matter distribution as defined in Equation 7.52, $b(m)$ is the bias as defined in Equation 7.65, and $\bar{\rho}$ is the average dark matter density which is approximated here as the critical density ρ_c.

7.3.5 Halo Occupation Distribution model

The initial description of the Halo model, summarised by Cooray and Sheth (2002), made simple assumptions about the halo occupation distribution. The halo occupation distribution describes the number of galaxies expected to occupy the dark matter halos. Zheng et al. (2005) developed this further to show that the properties of galaxies residing at the centres of halos differ from those of satellite galaxies because of the differences in their formation histories. They used semi-analytic models, derive expressions for the separate contributions of central and satellite galaxies to the halo occupation distribution. They showed that the probability of a central galaxy can be estimated using a step function, and a power law approximation can be used for satellite galaxies. They approximated the total number of galaxies expected in a halo of mass m by,

$$N_{\text{tot}}(m) = N_{\text{cen}}(m) + N_{\text{sat}}(m), \tag{7.69}$$

where the number of central galaxies is given by,

$$N_{\text{cen}}(m) = 0.5\Big[1 + \text{erf}\Big(\frac{\log(m/M_{\text{min}})}{\sigma}\Big)\Big], \tag{7.70}$$

and the number of satellite galaxies is given by,

$$N_{\text{sat}}(m) = 0.5\Big[1 + \text{erf}\Big(\frac{\log(m/M_{\text{min}})}{\sigma}\Big)\Big]\Big(\frac{m}{M_{\text{sat}}}\Big)^{\alpha_s}. \tag{7.71}$$

7.3.6 Non-linear Galaxy angular correlation

All these components can be combined to calculate the power spectrum as described in Equations 7.42 to 7.45. This in turn can be used to estimate the galaxy correlation function as described in Equation 7.26.

To summarise, this chapter describes three different ways to approximate the galaxy angular correlation function, which can now be compared with observations as described in Chapter 8.

Chapter 8

Herschel observations compared to analytical models of galaxy clustering

In the previous chapter, different analytical models for the galaxy angular correlation function (ω_{obs}) were discussed. In this chapter, and in Short and Coles (2011), the different models are tested against sub-millimetre observations released from the science demonstration phase of the *Herschel* Space Observatory and are used to put estimates on the minimum halo mass associated with these sources.

8.1 *Herschel* Space Observatory

The *Herschel* Space Observatory was launched in 2009 and is the only space observatory to cover a spectral range from the far infrared ($\sim 1-100\mu$m) to sub-millimetre ($\sim 100-1000\mu$m). These wavelengths are obscured from view from the ground due to absorption by water vapor in the Earth's atmosphere, therefore *Herschel* provides a new and unique window through which to study high-redshift galaxy clustering.

Observing in the *Herschel* wavelength ranges allows the study of cooler objects with temperatures in the tens to hundreds of Kelvin. These objects include galaxies, perhaps in the early stages of their formation, obscured in clouds of gas and dust with star formation rate of around a few hundred solar mass per year (Hughes et al., 1998). A large proportion of the energy distribution of the radiation in the observable Universe is in the infrared to sub-millimetre part of the spectrum.

Herschel has three scientific instruments:

- the **Heterodyne Instrument for the Far Infrared (HIFI)** is a very high resolution spectrometer covering wavelengths from 157-625 micrometres in seven bands,
- the **Photodetector Array Camera and Spectrometer (PACS)** is an imaging photometer and medium resolution spectrometer covering wavelengths from 55 to 210 micrometres, and
- the **Spectral and Photometric Imaging Receiver (SPIRE)** is an imaging photometer and spectrometer. The imaging camera has three bands, each centered at 250, 350 and 500 micrometres respectively.

One of the problems of observing/imaging in the far-infrared/sub-millimetre wavelength range is that it has poor spacial resolution in comparison to say the optical wavelength band (for example see Figure 8.1). Poor spacial resolution is also one of the reasons it is difficult to get good spectroscopy in this spectral range, that and there is not a high abundance of spectral lines. Therefore, in the case of SPIRE particularly, redshifts are often derived photometrically rather than spectroscopically.

Observing time on *Herschel* has been awarded to a number of different scientific projects. Two of these surveys of particular interest to this work are HerMES and H-ATLAS which are discussed further below. This data is processed and cleaned to produce intensity maps, e.g. as shown in Figure 8.1. To calculate the

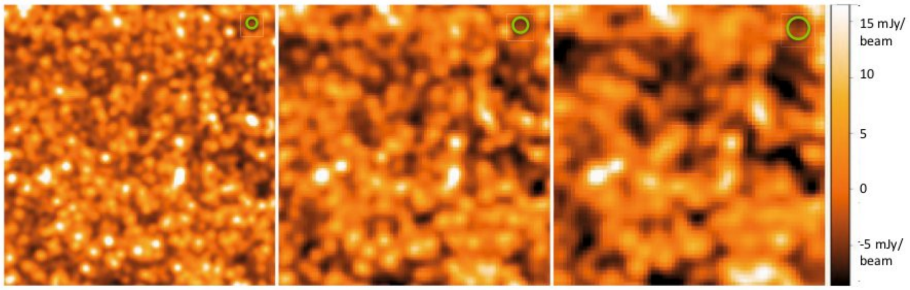

Figure 8.1: *The GOODS-North field shown as an example of a distribution of galaxies and background noise observed by SPIRE (Nguyen et al., 2010). The data is taken in the 250, 350 and 500 µm (left to right) wavelength bands. The green circles indicate the size of the beam relative to the* 16×16 *arcmin*^{2} *image. This shows how sources are less easily resolved at longer wavelengths.*

angular correlation function, the intensity maps need to be further processed to identify individual sources, which is normally done by imposing a flux cut-off. Then the correlation function is then calculated using the Landy & Szalay estimator described in Section 6.2.3.

8.1.1 *Herschel* multi-tiered extragalactic survey

The *Herschel* Multi-tiered Extragalactic Survey (Oliver et al., 2010, known as HerMES) will cover approximately 70 deg^2 of the sky, making observations in a hierarchical structure of several levels with increasing area but shallower coverage. The Lockman-SWIRE field, which is one of the shallower fields and 11 deg^2, was covered by the science demonstration phase; results of the clustering have recently been published by Cooray et al. (2010) along with a Halo model analysis.

8.1.2 *Herschel* astrophysical terahertz large area survey

The *Herschel* Astrophysical Terahertz Large Area Survey (Eales et al., 2010, known as H-ATLAS) will cover around 550 deg^2 of the sky. H-ATLAS is a legacy survey to record a large area of the sky; areas well studied in other wavelength bands at high Galactic latitudes were chosen. The current survey consists of five areas: the North Galactic Plane, the South Galactic Plane and three GAMA fields, designed to overlap with the Galaxy And Mass Assembly (GAMA) survey (Driver et al., 2009). This work uses data released from the science demonstration phase which consists of 16 deg^2 in one of the GAMA fields.

8.2 Model parameters

The analytical models have a number of input parameters. Firstly, cosmological parameters such as the matter and energy densities have a significant impact on the galaxy clustering. In this work they are fixed to match the current cosmological parameters as measured by the Wilkinson Microwave Anisotropy Probe (WMAP) as shown in Table 7.1.

The redshift distribution of the sources also plays a vital role in modeling the observed angular correlations. Figure 8.2 shows the normalised redshift distributions, $\mathcal{N}(z)$, which were used in my calculations reported here for both the HerMES and H-ATLAS data sets. The estimates for the redshift distribution for the HerMES data sets are the same as those used by Cooray et al. (2010) in their analysis. Several estimates of the redshift distribution of objects in the H-ATLAS survey have been made. First, Amblard et al. (2010) used colour-colour diagrams to estimate the redshift distribution; however this method includes only a subset of the sources used in the angular correlation analysis and is known to be slightly biased towards higher redshift objects. Second, Dye et al. (2010) used optical counterparts of sub-mm sources to estimate the redshift distribution; but again this method only uses a subset of all the objects and is biased towards lower redshift objects. So far, the best estimates of the redshift distribution, in that they use all the available sources, are those by Eales (2011) who used spectral energy distribution (SED) fitting to find a best fit redshift distribution.

In this analysis the estimates by Eales (2011) are used, but please note the

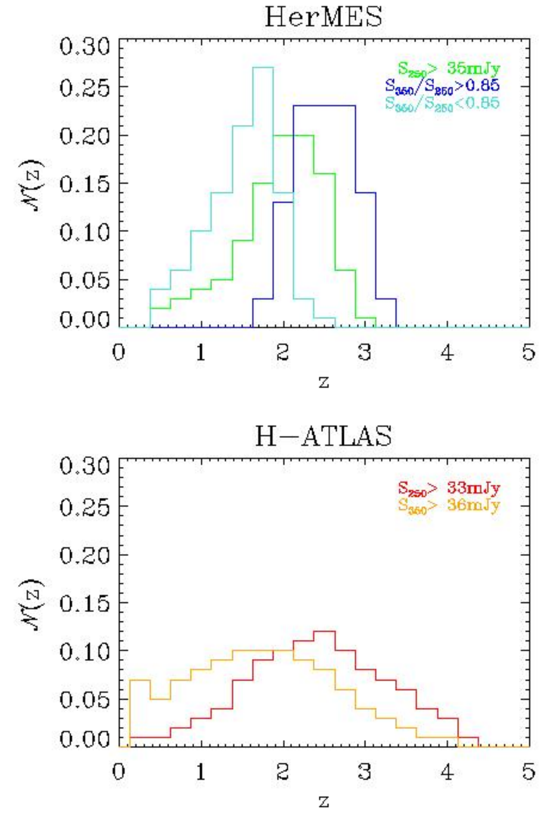

Figure 8.2: *The normalised distribution of galaxies as a function of redshift for data sets from HerMES and H-ATLAS. HerMES estimates for different wavelength and flux criteria are shown in the left plot and are sourced from Cooray et al. (2010): H-ATLAS estimates in the right plot are from Eales (2011). The redshift distributions are all roughly centred around $z \sim 2$ but are much wider in the case of H-ATLAS compared to HerMES.*

values for $\mathcal{N}(z)$ used here are *for illustrative purposes only* and should not be regarded as definitive given the preliminary state of the data. We have tried a variety of alternative models and find that, for reasonable choices, the results for $\omega_{\rm obs}$ are not especially sensitive to a particular mean redshift in $\mathcal{N}(z)$; this is probably because of the relatively slow evolution of the power spectrum at low redshift in the concordance cosmology. On the other hand the results are sensitive to the width of the distribution in z; the wider the distribution over z the lower the amplitude of the angular correlation function. This is expected given that the clustering signal is more concentrated in narrower redshift bands. The other model parameters are the minimum halo mass ($\mathrm{M}_{\rm min}$), and the average mass of a halo with one satellite galaxy ($\mathrm{M}_{\rm sat}$) and the slope of the first moment of the satellite galaxy HOD (α_s) in the case of the Halo model. These are used as free parameters to find a 'best fit' to the data. Because this work is focused on comparing the models, rather than finding precise results from the preliminary data, a 'best fit' is achieved 'by eye' as opposed to any particular numerical method, unless otherwise stated.

8.3 Analytical models compared to HerMES observations

Results for the galaxy angular correlation from the initial science run by HerMes were recently published by Cooray et al. (2010) along with a Halo model analysis. Figure 8.3 shows the results from Cooray et al. (2010) for the Lockman-SWIRE field of the HerMES survey for the different SPIRE wavelength bands. The results from the PACS instrument were too noisy to be included in the preliminary analysis. The criteria for sources to be included in the results are defined by flux; $\mathrm{S}_{250} > 35$mJy, $\mathrm{S}_{350}/\mathrm{S}_{250} > 0.85$ and $\mathrm{S}_{350}/\mathrm{S}_{250} < 0.85$. The flux criteria $\mathrm{S}_{350}/\mathrm{S}_{250}$ was chosen to divide low and high redshift sources as can be seen from Figure 8.2. Sub-mm sources at higher z peak in their spectral energy distributions at lower frequency: $\mathrm{S}_{350}/\mathrm{S}_{250} = 0.85$ was chosen as a cut-off

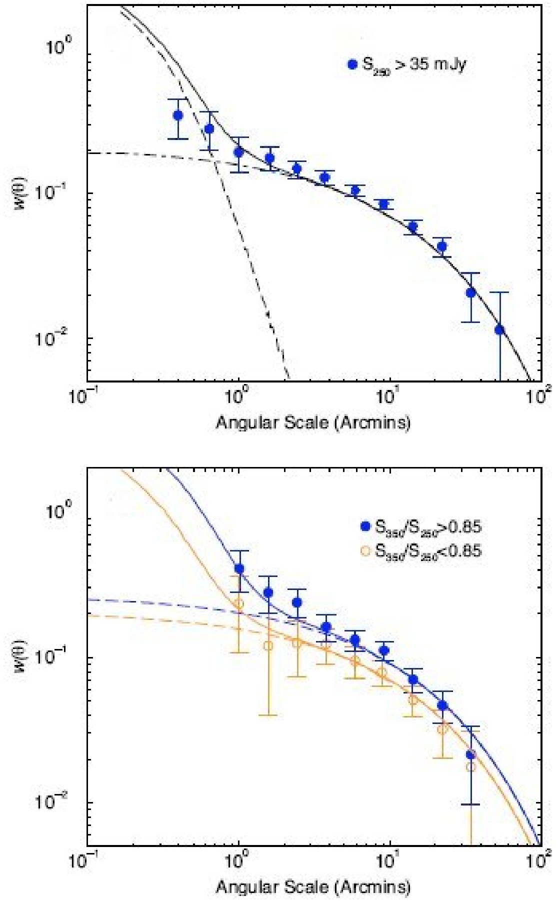

Figure 8.3: *These plots are from Cooray et al. (2010) and show the angular correlation function $\omega_{obs}(\theta)$ measured in the HerMES survey in the Lockman-SWIRE field. Data is compared to the Halo model which fits well in all cases. The dot-dashed lines show the two-halo term, the long-dashed lines shows the one-halo term and the solid line shows the total correlation function.*

between sources at $z < 1$ which have $\mathrm{S}_{350}/\mathrm{S}_{250} < 0.85$ and sources with $z > 1$ which have $\mathrm{S}_{350}/\mathrm{S}_{250} > 0.85$.

From the observations we see that the angular correlation has similar amplitudes in each of the three data sets; the middle panel has the highest amplitude and this most likely because the redshift distribution of the component sources is the most narrow of the three data sets making the clustering signal more concentrated and appear stronger (see redshift distribution is Figure 8.2).

8.3.1 Observations compared to a power law

For interest we first consider a simple power law approximation for galaxy angular correlation function. The power law is frequently used to approximate the ω in quick analyses,

$$\omega(\theta) = A\theta^{-\delta}. \tag{8.1}$$

But it doesn't always fit the data that well, and the parameters used to fit it don't really have a particularly meaningful physical interpretation. The examples in Figure 8.4 show the power-law generally fits this data well at small angular scales. However, the power law under estimates at mid scales and over estimates at higher scales thereby limiting its usefulness in further analysis. The power law fit was found using a least squares approach on the log of the data.

8.3.2 Observations compared to a linear model

The first model, the *linear* angular correlation function, is the simplest (ignoring the power law) of the three models we consider in this work. It provides a good fit to each of the data sets as can be seen in Figure 8.5. We see the linear model fit is much improved than the power law approximation plotted against the same data in Figure 8.4.

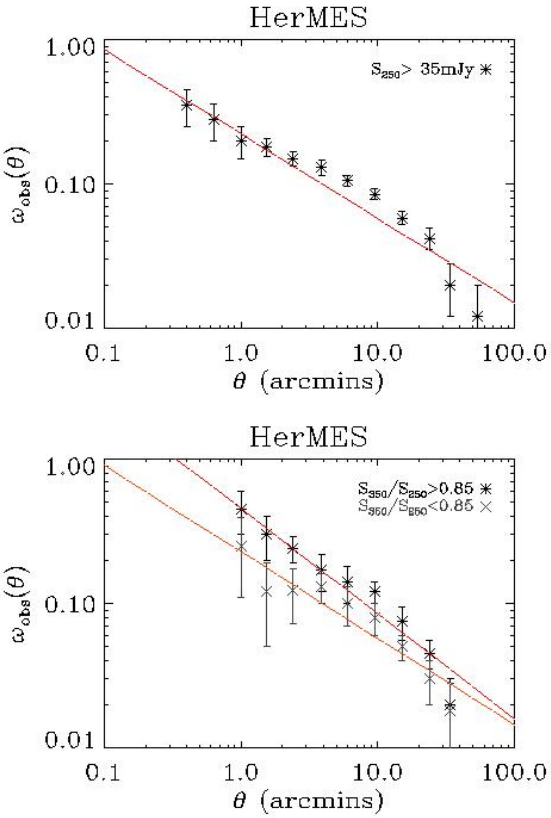

Figure 8.4: *These plots show the angular correlation function* $\omega_{obs}(\theta)$ *measured in the HerMES survey in the Lockman-SWIRE field. Sources were divided into three sets by flux density:* $S_{250} > 35$*mJy (left),* $S_{350}/S_{250} > 0.85$ *(right - black) and* $S_{350}/S_{250} < 0.85$ *(right - grey). Data is compared to a power law approximation; it over estimates at large scales and under estimates at mid angular scales.*

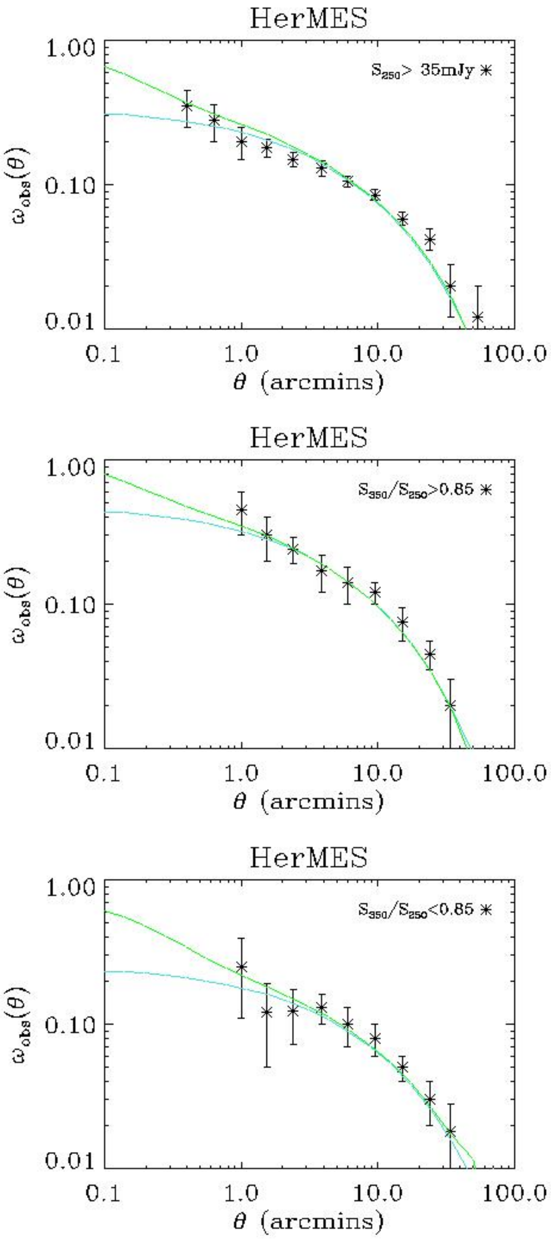

Figure 8.5: *These plots show the angular correlation function* $\omega_{obs}(\theta)$ *measured in the HerMES survey in the Lockman-SWIRE field. Data is compared to two different theoretical models: the non-linear approximation using the fitting function (green line) and the linear approximation (light blue line). Both approximation fit the data well.*

8.3.3 Observations compared to a non-linear fitting function

Figure 8.5 also shows the results for the (non-linear) fitting function approximation. By comparing the linear and non-linear models we see that the transition between linear and non-linear regimes does not become evident until quite small scales (~ 1 arcmin). However, limited resolution of the *Herschel* telescope in this range makes it difficult to probe the clustering regime on scales much smaller than this.

8.3.4 Observations compared to a non-linear Halo model

The results for the Halo model, as well as those from the fitting function method, are shown in Figure 8.6. The fit at intermediate scales is perhaps slightly better for the Halo Model over the fitting function, but then this approach has two more free parameters. However, overall both the Halo model and the fitting function method show good fits over the scales for which data is available. The differences in the behaviour at large angular scales is a consequence of slightly different best-fit values for the bias parameter.

The Halo model results are similar to the analysis by Cooray et al. (2010). The Halo model approximation is the sum of two components: the one halo and two halo terms. Cooray et al. (2010) and Amblard et al. (2011) argue that the good comparison between observations and the Halo model in Figure 8.3 is evidence for these two components. We considered the fitting function method, which doesn't include correlations between galaxies in the same halo (i.e. the one halo component), to give us a comparative model with which to test this hypothesis. Given that both models fit the data equally well, by eye, this shows that the two components are not (currently) necessary to model observed results. Rather we are not yet probing scales where the intra-halo correlations are becoming significant.

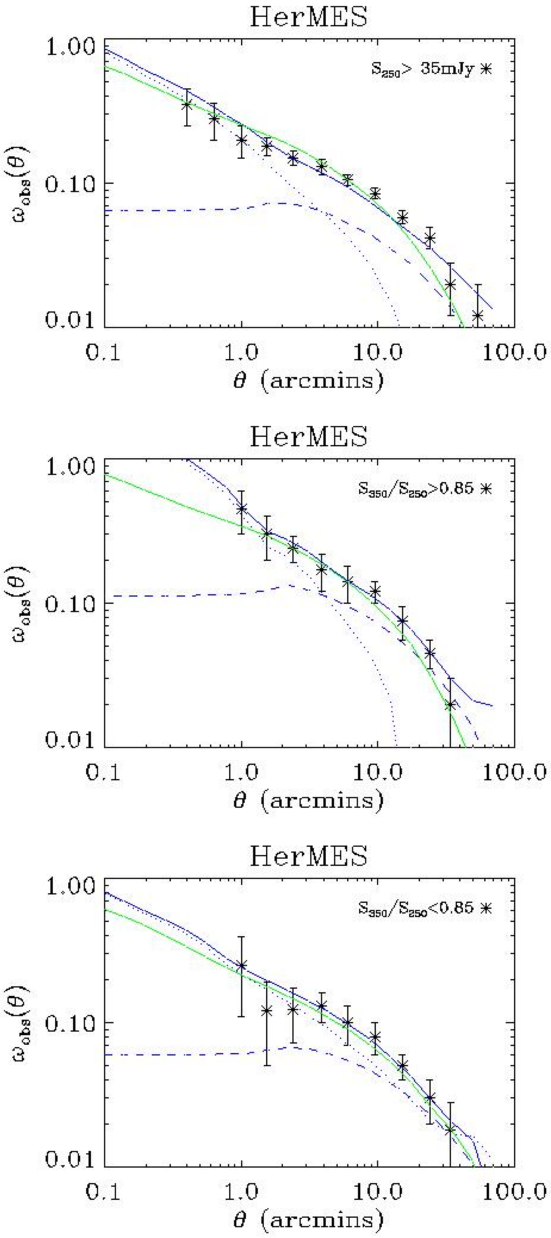

Figure 8.6: *These plots show the angular correlation function $\omega_{obs}(\theta)$ measured in the HerMES survey in the Lockman-SWIRE field. Data is compared to two different theoretical models: the non-linear approximation using the fitting function (green line), and the Halo Model (blue line). The Halo and fitting function methods both fit the data well.*

8.3.5 Estimates of the minimum halo mass

In both the fitting function and Halo models the amplitude of large-scale clustering depends on the bias and that, in turns, boils down to the minimum halo mass through Equations 7.32 and 7.69. We can therefore use the measured clustering amplitude to get a rough estimate of the mass of halos hosting these galaxies. The values of the minimum halo mass corresponding to the theoretical models in Figure 8.6 are shown in Table 8.1. We see that, for each of the data sets, the minimum halo masses found from both the fitting function and Halo models are in agreement (within the errors) and suggest a value of $\mathrm{M_{min}} \sim 10^{13.2\pm0.4} M_\odot h^{-1}$.

Data Set	$M_{\mathrm{min}}^{\mathrm{FF}}$	$M_{\mathrm{min}}^{\mathrm{Halo}}$	M_{sat}	α_s
$\mathrm{S_{250}} > 35\mathrm{mJy}$	$10^{13.4}$	$10^{13.0}$	$10^{13.0}$	< 1.2
$\mathrm{S_{350}/S_{250}} > 0.85$	$10^{13.2}$	$10^{13.2}$	$10^{13.4}$	< 1.2
$\mathrm{S_{350}/S_{250}} < 0.85$	$10^{13.4}$	$10^{13.0}$	$10^{13.2}$	> 1.4

Table 8.1: *Best fit parameters for the Halo and fitting function models for the HerMES results. M_{min} is the minimum halo mass for which galaxies can form, M_{sat} is the average mass of a halo with one satellite galaxy and α_s is the slope of the first moment of the satellite galaxy HOD. All masses are in units $M_\odot h^{-1}$. We see that the M_{min} found for each data set are consistent between models, given an accuracy of $10^{\pm0.2}$ $M_\odot h^{-1}$ on the masses.*

8.4 Analytical models compared to H-ATLAS observations

The other results analysed here are from the H-ATLAS survey (see Section 8.1.2). Maddox et al. (2010) recently released measurements of the angular correlation function of the galaxies observed. That analysis compares the results to a simple power law approximation for the angular correlation function. This was because, as can be seen in Figure 8.7, results for the angular correlation are quite noisy so none of the models fit particularly well - a more detailed analysis will be possible on the completion of the survey. This noise is thought to be due to Galactic cirrus in the H-ATLAS data as the number of detections are very similar to the HerMES data sets and both surveys are similarly confusion limited. Therefore here just the fitting function results are plotted against the data. They were used to find approximate values of the corresponding minimum halo masses, which are $10^{12.5}$ and $10^{13.0}$ $M_{\odot} h^{-1}$ for the examples plotted in Figure 8.7 left to right respectively.

The scale of the $\omega_{\rm obs}$ results for the $\mathrm{S}_{250} > 33$mJy flux cut is significantly smaller that those in the other examples. This is predominately due to the wide bimodal redshift distribution (see Figure 8.2) although it does still suggest a slightly lower minimum mass $\sim 10^{12.5}$ $M_{\odot} h^{-1}$. This minimum mass is consistent with that found in the cross correlation analysis by Guo et al. (2011).

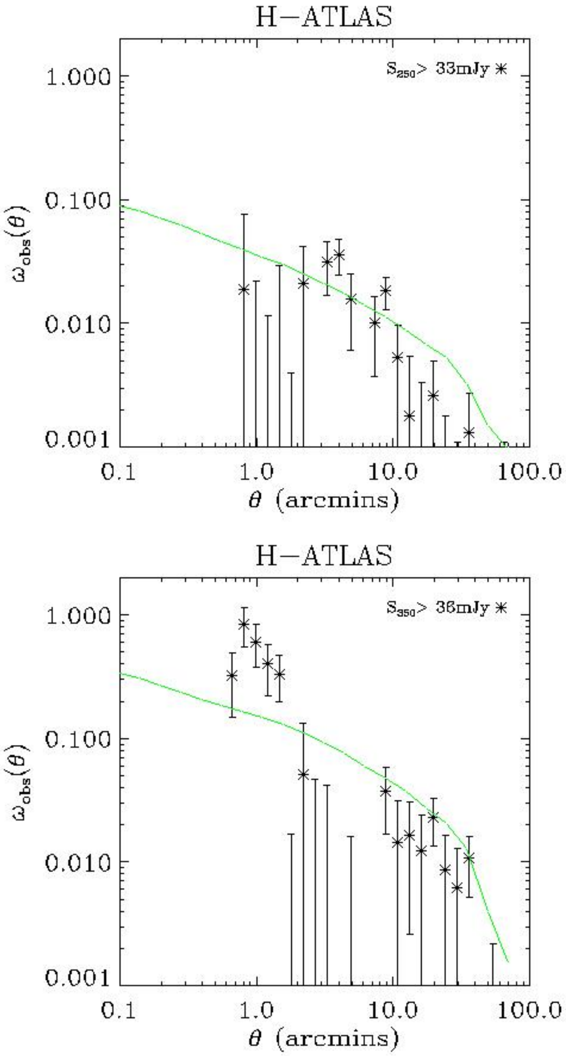

Figure 8.7: *These plots show the angular correlation function* $\omega_{obs}(\theta)$ *measured in the science demonstration phase of the H-ATLAS survey. The data are compared to the fitting function model (green line).*

Chapter 9

Investigating Galaxy Clustering Models

In this chapter we investigate how sensitive the non-linear fitting function model for the non-linear galaxy angular correlation function (described in Chapter 7) is to the different free parameters. The free parameter common to both the Halo model and the fitting function approach is the the minimum halo mass ($\mathrm{M}_{\mathrm{min}}$). However we also consider the redshift distribution of the observed galaxies ($\mathcal{N}$) which, whilst not really a free parameter as it is fixed by a given set of observations, is interesting to consider because it is known to be uncertain as shown in the results in Chapter 8.

9.1 Investigating effect of variations in $\mathcal{N}(z)$

In Chapter 8 we briefly touched on the importance of having accurate estimates of the sources redshift distribution. Since the redshift distribution is not well understood in all data sets it is worth discussing here to further understand the potential significance on the results.

To demonstrate the effects of different redshift distributions on the angular correlation function, Figures 9.1 and 9.2 show the resulting $\omega_{\rm obs}$ for a selection of source distributions over different ranges of redshifts. In Figure 9.1 the width of the distribution is fixed and the mean value varies. Figure 9.2 shows the effect of the changing width of the source distribution on $\omega_{\rm obs}$, whilst keeping the mean value fixed. Results for $\omega_{\rm obs}$ are derived using the fitting function approach for a minimum halo mass of $10^{13.5}$ $M_{\odot}h^{-1}$ (Note, these results are consistent with the Halo model).

Figure 9.1 shows how the variation in mean redshift causes the amplitude of $\omega_{\rm obs}$ to increase with increasing redshift. This is because at increasing redshift only the brighter objects are actually observed. These bright objects tend to be the largest of objects and therefore are the most strongly clustered[1], hence the effective bias is larger (see Figure 9.6), however the variation in amplitude is not dramatic. Figure 9.2 shows how the increasing width of the source redshift distribution causes the amplitude of $\omega_{\rm obs}$ to decrease. This is because clustering occurs due to gravitational interactions between objects which therefore have to be relatively close. Therefore as you increase the redshift range you look at what would have been a strongly clustered group in a small range gets averaged out to be weaker. In this example, the effect of varying the width of the source distribution appears much more pronounced than for the mean redshift example in Figure 9.1, although please note that the magnitudes of the variations in μ and σ are not directly comparable.

This demonstrates that the evolution in redshift of the angular correlation function is not too fast, so if there is uncertainty in the actual distance of a population of objects the resulting errors would not be excessively large. More important it seems, at least in this example, is the profile of a population over redshift. If a population is picked up that spans a large range of redshifts this has the effect of diluting out the clustering signal.

It is therefore important to understand the redshift distribution since it affects

---

[1] Larger halos (and their associated galaxies) are more strongly clustered because they are generally associated with large scale over-densities which amplify local peaks in the smaller scale fluctuations so that they collapse to form more massive halos than they otherwise would.

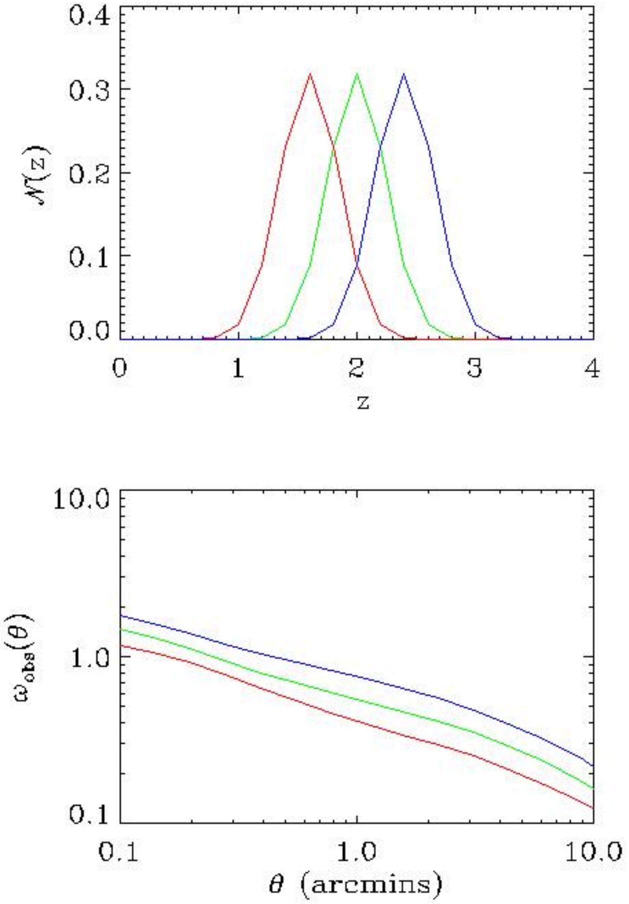

Figure 9.1: *Top: A selection of normalised Gaussian source distributions ($\mathcal{N}$) with fixed standard deviation (σ = 0.25), but with different mean redshifts at $\bar{z}$ =1.6, 2, and 2.4. Bottom: Corresponding angular correlation function ω_{obs} for $M_{min} = 10^{13.5}$ $M_{\odot}h^{-1}$ using the fitting function method.*

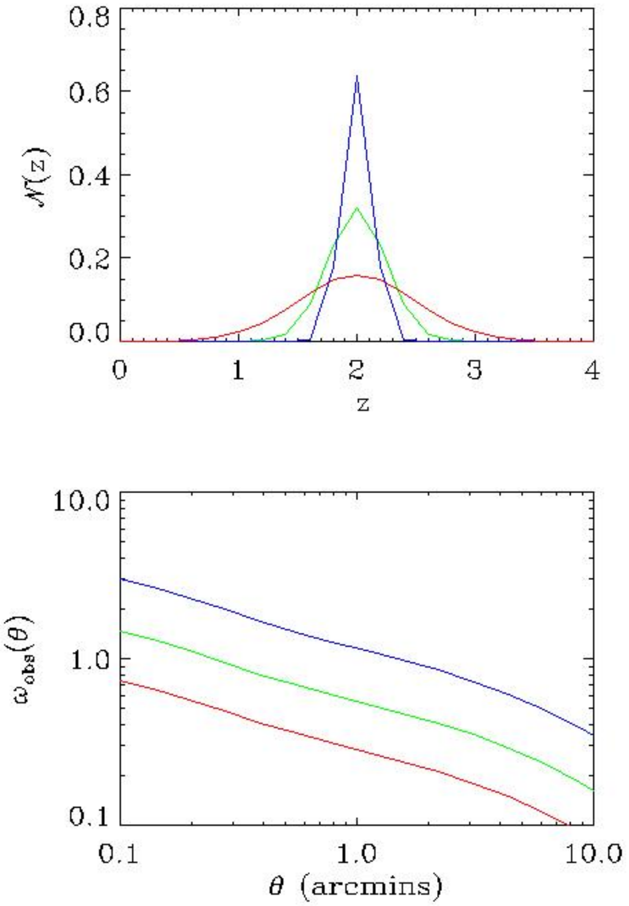

Figure 9.2: *Top: A selection of normalised Gaussian source distributions ($\mathcal{N}$) with fixed mean redshift ($\bar{z}$ = 2) but with different standard deviation of σ =0.125, 0.25 and 0.5. Bottom: Corresponding angular correlation function ω_{obs} for $M_{min} = 10^{13.5}$ $M_{\odot}h^{-1}$ using the fitting function method.*

the amplitude of $\omega_{\rm obs}$. As we shall discuss in Section 9.2, this variation in the amplitude of $\omega_{\rm obs}$ could likely be misinterpreted as being due to a change in the minimum halo mass. Having said this, there is not much that can be done to improve the model to account for this uncertainty, but awareness of the issue means that results can be considered with caution.

9.2 Investigating effect of variations in $M_{\rm min}$

The minimum halo mass ($\rm M_{min}$) is the minimum mass that a halo can have to host a galaxy. For example in case of the fitting function method it simply enters the calculation as the lower limit on the integral over the mass function (see Sections 7.2 and 7.1.6). Consequently it only has influence on the amplitude of $\omega_{\rm obs}$ not its shape as a function of scale.

Note, this is not quite the same for the Halo model. Here $M_{\rm min}$ also contributes to the halo occupation distribution which describes the small scale intra-halo clustering. This would therefore result in the amplitude of $\omega_{\rm obs}$ rising slightly faster at small scales for increasing mass.

Figure 9.3 shows the angular correlation function estimated using the fitting function method for a selection of $M_{\rm min} = 10^{11.0}, 10^{11.5}, ..., 10^{14.0}\ M_{\odot}h^{-1}$. The redshift distribution used is the same as for the 250μm HerMES data set considered in Chapter 8.

Figure 9.3 shows that the amplitude of $\omega_{\rm obs}$ increases with increasing $M_{\rm min}$. As a very approximate comparison, the amount of evolution in the magnitude due to an increase in $\log_{10} M_{\rm min} = 0.5$ is of the same sort of order of doubling the width of the source distribution discussed in Section 9.1.

9.2.1 Smooth cutoff at $M_{\rm min}$

We discussed in Section 9.1 that the source redshift distribution is not really a free parameter, it is fixed. So with observations of increased accuracy the uncertainty will be reduced. The uncertainty related to $\mathcal{N}$ can not therefore

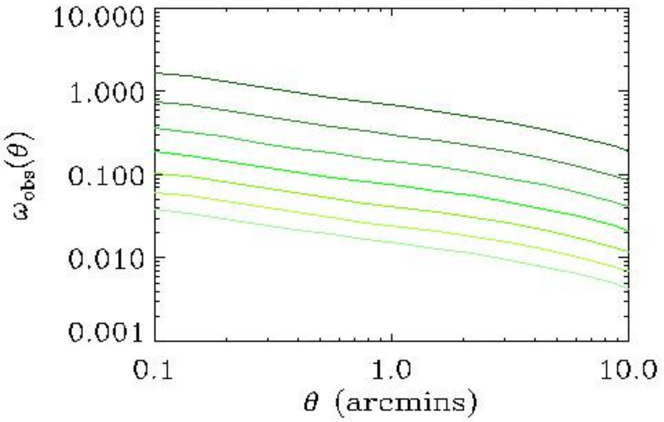

Figure 9.3: *The angular correlation function (ω_{obs}) using the fitting function method for a selection of fixed $M_{min} = 10^{11.0}, 10^{11.5}, ..., 10^{14.0}$ $M_{\odot}h^{-1}$ increasing from light to dark green (top to bottom). The redshift distribution is from the 250μm HerMES data set.*

really be improved via better modeling. However, in contrast to this perhaps the $M_{\rm min}$ free parameter can be better quantified.

In Section 9.1 and the first part of Section 9.2 we considered the minimum halo mass to be a sharp cutoff. A sharp cutoff means that the survey data set includes *all* objects associated with halos above $M_{\rm min}$, but *none* of the objects associated with halos below this cutoff. However this is a very simplistic approximation (Bouche et al., 2010); whilst the physical mechanism that quenches the formation of galaxies in low mass halos is not known, it seems more intuitive that this transition would be smooth.

Here we consider a more gradual cutoff or in other words, we assume that it is increasingly unlikely to observe objects associated with halos of mass below $M_{\rm min}$, but it is increasingly more likely to see objects associated with halos with masses above $M_{\rm min}$. If this is the case, as seems reasonable, then we consider how this would 'bias' the results if you used the simpler sharp cutoff analysis.

Of course it depends on the 'gradient' of this gradual cutoff; the steeper the gradient the more similar to the sharp cutoff the case becomes. So here we consider how a couple of different 'strengths' of 'gradient' would bias the resulting M_{min}.

As a simple approximation we use a gradient shaped like an *error function*, which is simply the integral of a Gaussian distribution. This is therefore equivalent to a Gaussianly distributed M_{min}. To calculate the ω_{obs} resulting from this Gaussianly distributed M_{min}, we simply integrate the original ω_{obs} calculation over a Gaussian distribution, i.e.

$$\omega'_{\mathrm{obs}} = \int_{0}^{\infty} P(\log M_{\mathrm{min}}) \omega_{\mathrm{obs}} d \log M_{\mathrm{min}}, \tag{9.1}$$

where

$$P(\log M_{\mathrm{min}}) = \frac{1}{\sqrt{(2\pi\sigma^2)}} \exp\Big(- \frac{(\log M_{\mathrm{min}} - \mu)^2}{2\sigma^2} \Big). \tag{9.2}$$

This is calculated here for two examples; Figure 9.4 shows the distribution of $P(\log M_{\mathrm{min}})$ and corresponding ω'_{obs} where $\mu = 13.4$ and $\sigma = 0.5$, and Figure 9.5 shows the same for $\mu = 13.4$ and $\sigma = 1.0$. As in Section 9.2, the redshift distribution used is from the 250μm HerMES data set.

Figure 9.4 shows $P(\log M_{\mathrm{min}})$ and the corresponding ω_{obs} for $\mu = 13.4$ and $\sigma = 0.5$ plotted in red. This is compared to the original method of just having a fixed cutoff at the mean value $\mu = 13.4$ or rather $M_{\mathrm{min}} = 10^{13.4}$ $M_{\odot}h^{-1}$ (plotted in blue). We see that the amplitude of ω_{obs} has increased with the Gaussianly distributed M_{min}. Thinking now about the problem reversed, we consider what value of a fixed M_{min} would be required to achieve this increase in amplitude, and find in this example it is equivalent to a fixed cutoff at $M_{\mathrm{min}} = 10^{13.5}$ $M_{\odot}h^{-1}$ (plotted in green).

Figure 9.5 shows that same as Figure 9.4 except that the σ for the Gaussianly distributed M_{min} is 1 as opposed to 0.5 (plotted in red). The fixed cutoff at the mean value $\mu = 13.4$ is again plotted in blue. As in Figure 9.4 the amplitude of ω_{obs} has increased between the fixed M_{min} and Gaussianly distributed M_{min},

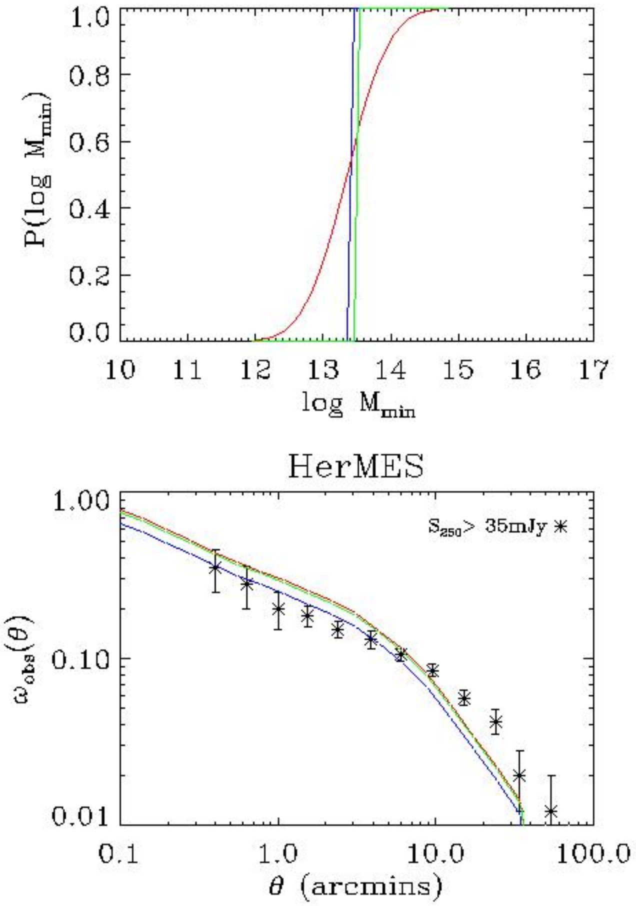

Figure 9.4: *Top:* $P(\log M_{min})$ *for each of three examples: a fixed cutoff of at* $M_{min} = 10^{13.4}\ M_{\odot}h^{-1}$ *(blue); a Gaussianly distributed* $\log M_{min}$ *with a mean* $\mu = 13.4$ *and standard deviation* $\sigma = 0.5$ *(red); and another fixed cutoff of at* $M_{min} = 10^{13.5}\ M_{\odot}h^{-1}$ *(green). Bottom: The corresponding* $\omega_{obs}(\theta)$ *approximated using the fitting function approach with data points from the HerMES survey for* $S_{250} >$ *35mJy, as is the source redshift distribution.*

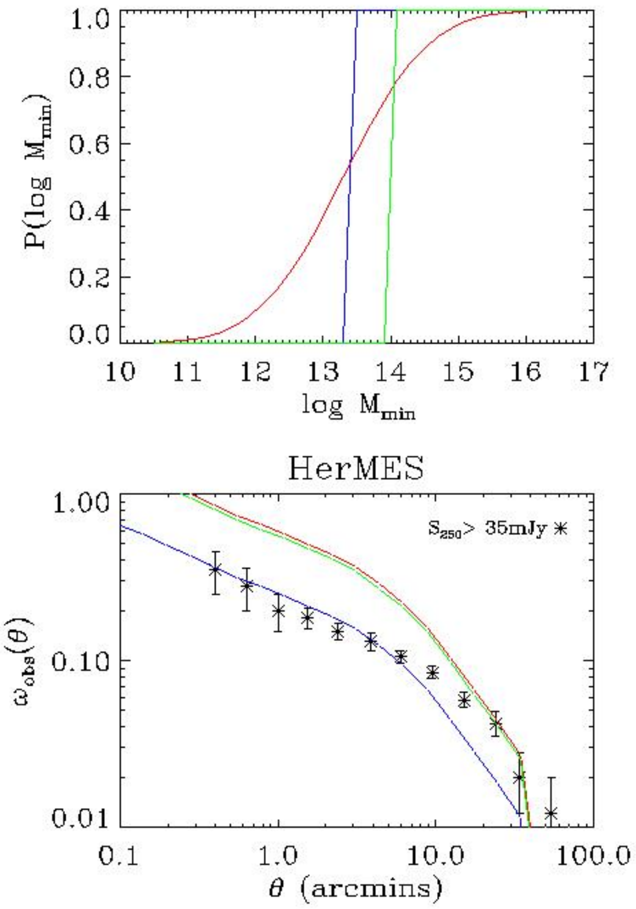

Figure 9.5: *Top:* $P(\log M_{min})$ *for each of three examples: a fixed cutoff of at* $M_{min} = 10^{13.4}$ $M_{\odot}h^{-1}$ *(blue); a Gaussianly distributed* $\log M_{min}$ *with a mean* $\mu = 13.4$ *and standard deviation* $\sigma = 1.0$ *(red); and another fixed cutoff of at* $M_{min} = 10^{13.9}$ $M_{\odot}h^{-1}$ *(green). Bottom: The corresponding* $\omega_{obs}(\theta)$ *approximated using the fitting function approach with data points from the HerMES survey for* $S_{250} >$ *35mJy, as is the source redshift distribution.*

but this time much more significantly. A fixed value of $M_{\rm min} = 10^{13.9}\ M_{\odot}h^{-1}$ (plotted in green) would be required to achieve this increase in amplitude
So we see that the amplitude of the resulting $\omega_{\rm obs}$ is increased for a Gaussianly distributed $M_{\rm min}$ compared to what would be expected for a fixed cutoff at the mean value. This effect becomes increasingly significant at an increasingly rapid rate for higher values of σ. To understand the reason for this behaviour we consider Figure 9.6 which shows the effective bias ($b_{\rm eff}$, Equation 7.32) for a range of (fixed) minimum halo masses: $10^{11.0}$, $10^{11.5}$, ..., and $10^{14.0}M_{\odot}h^{-1}$. We

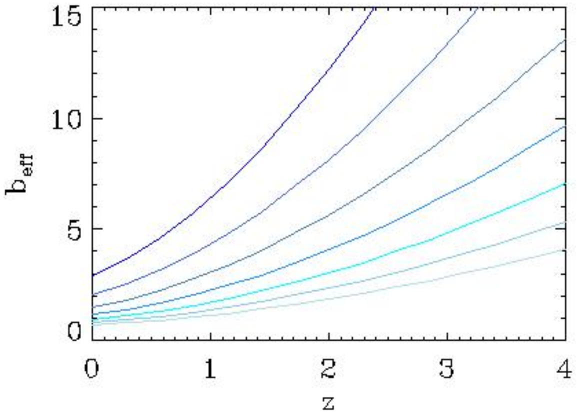

Figure 9.6: *The effective bias ($b_{\rm eff}$) for a range of (fixed) minimum halo masses: $10^{11.0}$, $10^{11.5}$, ..., and $10^{14.0}M_{\odot}h^{-1}$ (light to dark blue respectively).*

see that the bias becomes increasingly large with increasing mass, so when this is averaged over a Gaussianly distributed $M_{\rm min}$ as in the previous plots, the $b_{\rm eff}$ will not average out to be the same as for a fixed $\log(M_{\rm min}) = \mu$ but will have a higher overall value.

To summarise this chapter, we have considered the effect of uncertainty in two types of input to the galaxy angular correlation as approximated using the fitting function model as described in Chapter 7. Firstly we looked at the effect of varying the source redshift distribution and found that whilst sources at higher/lower redshift cause an increase/decrease in the amplitude of $\omega_{\rm obs}$, this evolution in the $\omega_{\rm obs}$ is not as significant as the 'washing-out' affect that occurs when the redshift range is increased.

Then we looked at the minimum halo mass to confirm that, like the source distribution, $M_{\rm min}$ affects the amplitude of $\omega_{\rm obs}$. Therefore an error in $\mathcal{N}$ could contaminate the extracted $M_{\rm min}$. We also discussed whether using a fixed cutoff in $M_{\rm min}$ was appropriate and considered an alternative gradient affect. If indeed the gradient affect is more realistic then the resulting $M_{\rm min}$ extracted using a fixed $M_{\rm min}$ technique would tend to appear higher than the true value.

Chapter 10

Conclusions for Galaxy Clustering

Over the past few years, much work has gone into trying to develop the standard *concordance* cosmological model to include a standard explanation for the formation and evolution of large scale structure. There are currently good methods for explaining the statistics of the overall matter distribution at early times and its evolution in linear regimes. However, in areas of large over density, with strong gravitational interaction, the non-linear motions of the matter are less easily quantified. Further there is no universally agreed explanation for the subsequent processes followed by baryonic matter to form galaxies.

With the recent launch of the *Herschel* Space Observatory (Pilbratt et al., 2010), a project with which many astronomers at Cardiff University are actively involved in, this is a good opportunity to test some of the current methodologies on this new data set. *Herschel* is the first space based observatory to look in the sub-millimeter wavebands, which opens up a whole new view of the Universe to us. Objects observed by *Herschel* are surrounded by lots of dust which often means they are early in their evolution i.e. just forming. Observations from *Herschel* will be used to investigate the properties of these galaxies, and compare the results to those in other wavebands.

That said, it should be stressed that the data sets considered in this work are from the science demonstration phase of *Herschel* are therefore only *preliminary*. We emphasise that this work is more focused on comparing the different analytical models for the angular correlation function, than drawing strong conclusions about *Herschel* data. The availability of *Herschel* data provided inspiration for the work as well as a new and exciting data set to provide realistic comparison for the analytical models. The key results are the comparison of the models themselves; the *Herschel* results are for illustrative purposes but do serve to situate the otherwise theoretical research in an empirical context.

This work considers three analytical models for the galaxy angular correlation function. This is a commonly used statistic to describe the overall statistical properties of a set of galaxy observations. We recall the angular correlation describes the probability of finding two objects at a given angular separation on the sky above what would be expected in a random distribution of objects. It considers the projected distribution of galaxies i.e. the observed distribution over a solid angle of the sky, as opposed to varying as a function of redshift.

The main model of interest considered in this work is the *fitting function* method. This follows from work by Hamilton et al. (1991) and Peacock and Dodds (1996) who showed that the non-linear matter correlation function can be approximated directly from its linear counter part, by way of a fitting function approach with numerical simulations. Then Coles et al. (1999) showed that the bias required to generate the galaxy angular correlation function from the matter one is well approximated by a linear factor. This method is then compared with the other two models.

The second of the three models is the *linear* approximation, which is included as a simplistic comparison to the fitting function method in that it assumes no non-linear clustering. Actually, the simplest approximation most frequently used in the literature is a *power law*. This is briefly considered in this work, but since it has no physical motivation the linear method is also considered. There are several different analytical linear approximations; this work follows that by

Eisenstein and Hu (1999).

The last of the three models is the *Halo model* (Cooray and Sheth, 2002), which is included as currently the most comprehensive analytical approximation for non-linear clustering.

The aim of this work is to 'revive' this fitting function approach. Whilst for full analytical studies the Halo model is unarguably effective, its complexity is its disadvantage. It makes it somewhat inaccessible for quick reports of observational results as it is a complicated and time consuming to perform. This work concentrates on whether a more simplistic fitting function approach could fill this void currently dominated by a power law which is known to be inaccurate. To be able to do this the fitting function needs to have comparable results to the Halo model without the 'fuss'. This would then be useful for quick analyses, and also for other studies which incorporate the angular correlation function into further calculation.

In this work, and in Short and Coles (2011), we highlighted that the fitting function method provides a much improved fit to the *Herschel* data than a power law, a slightly improved fit than the linear approximation and a similar fit to that of the Halo model. We also show that the minimum halo mass found using our fitting function approach is consistent with that found using the Halo model. We therefore confirm the much improved fit over a power law, comparable to that of the Halo model, of the fitting function method which is much simpler and easier to use than the Halo model.

Another point to highlight is the number of free parameters used by each of the different models. The linear approximation and the fitting function method both have just one - the minimum halo mass M_{min}. The power law has two free parameters - the amplitude and slope of the clustering, neither of which are actually physical parameters, in that they don't physically describe the objects being measured but are a parameterisation of the observable. Finally the Halo model has three three physically meaningful free parameters which include the minimum halo mass, along with a further two which describe the population of multiple galaxies in a halo. This again highlights the advantage of using the fitting function method.

As we already said, the results showed that, in fact, the linear angular correlation function also provides a reasonable fit to the data. Owing to the limited resolution of the *Herschel* telescope it is difficult to identify pairs of galaxies sufficiently close together on the sky to probe anything but the mildly non-linear regime. The currently available data provide some evidence of a transition between the linear and non-linear regimes, but as of yet they provide no unambiguous detection. Further investigation is required with the full data set which will be available shortly.

We also touched on the fact that the Halo model includes in its calculation a possibility of multiple galaxies per dark matter halo. However the fitting function method does not - it assumes either one or no galaxies per halo. This is known to be incorrect as there is evidence for multiple galaxies per halo. The additional intra-halo correlations generated by multiple galaxies will increase the amplitude of the clustering, but only on smaller scales. The inflection seen in the observed angular correlation function has in the past been used as evidence for these intra-halo correlations. However this work highlights that on the scales probed by current data the intra-halo correlations are not yet evident, and in fact the inflection observed is simply due to non-linear correlations due to galaxies in separate halos.

Having said this however, data is starting to reach the limit where intra-halo correlations will be significant as can be seen by the deviation of the two non-linear models just outside of the currently observable scales. Therefore the ability to model intra-halo correlations, whilst not immediately necessary, will become important in the future. It would therefore be interesting to investigate if a simple approximation for this can be used to further extend the fitting function approach.

The biggest stumbling block to a more complete analysis relates to the considerable uncertainties in the redshift distribution $\mathcal{N}(z)$ of the sources involved. The choices we adopted for this analysis are for illustration only so the results should not be regarded as definitive. Further data, especially ancillary data providing spectroscopic redshifts, will be needed before the precise nature of these galaxies can be determined.

To summarise, we have shown the advantage of using the fitting function method over the other commonly used alternatives in the literature. Although neither as sophisticated nor as flexible as the Halo model, it remains a useful tool that is perfectly adequate for modeling currently available data. This does not mean the Halo model is incorrect, of course, but what it does mean is that, at least for the time being, the paraphernalia involved in modeling intra-halo correlations is rather superfluous for these objects; simpler models can yield perfectly adequate results.

Chapter 11

Discussion

Observations of the Cosmic Microwave Background have been key to developing the standard model. The initial discovery of the CMB brought the Big Bang model to the forefront of models for the evolution of the Universe over the preceding popular alternative (Steady State). Subsequent observations of the fluctuations in the CMB have allowed cosmologists to extract per cent level accuracy on key cosmological parameters (Larson et al., 2011).

Studying the CMB temperature perturbations for anomalies from the standard model, generically known as 'non-Gaussianities', has for some time been the focus of much debate. This is the area on which the first half of this thesis has been focused. We considered three different methods; the first for identifying possible contamination/over-subtraction of Galactic foregrounds and the second two for assessing global anisotropy.

However, recent work by Francis and Peacock (2010) has shown how the integrated Sachs-Wolfe signal can account for some of the previously identified large-scale anisotropic features. Furthermore, it is expected that the imminent data release from the recently launched Planck will likely explain many of the 'non-Gaussian' anomalies. And if they are not explained away, then it is expected at least some definitive high significance results can be extracted.

In fact it seems that rather than being a CMB experiment, Planck may turn out to be most useful for looking at galaxy clusters via the Sunyaev-Zeldovich effect. This is related to the second half of this PhD thesis which looked at statistical measures of galaxy clustering. Whilst this is similar to the CMB part in that we consider the statistics of the large scale properties of the Universe, the galaxy clustering is much more complex. The distribution of galaxies will be structured around the primordial conditions set down in the CMB, but in addition their distribution is also governed by subsequent gravitational processes involved in the formation of galaxies.

Key to the formation of galaxies is the unknown material termed ‘dark matter’. Many studies are on-going at all scales to try to identify what this substance really is (Bergstrom, 2009). This work addresses questions surrounding structure formation, such as understanding the characteristic sizes of the dark matter halos associated with these galaxies, and also considers the structure of these halos.

The recent release of data from the *Herschel* Space Observatory has prompted further discussion of the ‘Halo model’, and observations of increased correlation at small scale have been used as ‘evidence’ for its multiple objects per halo scenario. This evidence is in no way definitive; this work demonstrates a similar small scale effect can be generated without including intra-halo correlations.

However if further data does constrain the error bars to convince us that the excess in small scale correlations is there, then there are still potentially other explanations apart from intra-halo correlations that need to be considered. For example, *gravitational lensing* occurs when two objects (a source and a lens) coincidently align causing the light from the source object to be magnified (and distorted) by the foreground lens. Lensing affects the distribution of observable galaxies in different ways, for example; the lens and the source might appear sufficiently compact to be detected as one much brighter object which may or may not have been detected otherwise. The generation of small scale correlations in the galaxy correlation function will occur when a lens and a source (or possible multiple images of the source) can be distinguished as separate items where one or both would not have been observed without the lensing effect.

We see therefore that the models discussed here are still very approximate, with much more study of the individual components and approximations required. This work, and the majority of other current analyses of galaxy clustering, are predominately based on source lists of positive (i.e. statistically significant) detections of individual sources. Whilst these observations provide an abundance of new information, the majority of the intensity map will actually be rejected because positive identifications of individual sources will not be possible for the fainter objects. Intensity maps contain much more information from fluctuations caused by these fainter objects than just those identified statistically significant lensed sources. One way forward with this work would to be to develop models to investigate what structure remains in the maps after removing detected sources. There are currently proposed methods, and the angular correlation function is a key component, so the non-linear fitting function model discussed here has a clear application in future analyses of intensity maps.

To summarise, whilst there have been huge developments in cosmology to build a robust standard model, today there are still many unanswered questions. Forthcoming new instruments such as Planck should help finally answer some of these. However key cosmological questions still remain unanswered such as the nature of dark matter and dark energy. Whilst dark energy remains very mysterious, much progress is being made on characterising dark matter which is believed to play a key part in the complex process of galaxy formation. Models are still scratching the surface of this complexity but by continuing to address the problem from a multitude of different angles we gradually continue to refine and advance our understanding of the Universe.

Bibliography

A. Amblard, A. Cooray, P. Serra, B. Altieri, V. Arumugam, et al., *Sub-millimetre galaxies reside in dark matter halos with masses greater than* 3×10^{11} *solar masses*, Nature, 470, 510–512, 2011, `http://adsabs.harvard.edu/abs/2011Natur.470..510A`, arXiv: 1101.1080.

A. Amblard, A. Cooray, P. Serra, P. Temi, E. Barton, et al., *Herschel-ATLAS: Dust temperature and redshift distribution of SPIRE and PACS detected sources using submillimetre colours*, A & A, 518, L9–L13, 2010, `http://adsabs.harvard.edu/abs/2010A%26A...518L...9A`, arXiv: 1005.2412.

J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, *The statistics of peaks of Gaussian random fields*, ApJ, 304, 15–61, 1986, `http://adsabs.harvard.edu/abs/1986ApJ...304...15B`.

R. B. Barreiro, *An overview of the current status of CMB observations*, Astrophysics and Space Science Proceeding, 3, 93–102, 2010, `http://adsabs.harvard.edu/abs/2010hsa5.conf...93B`, arXiv:0906.0956.

J. D. Barrow, R. Juszkiewicz, and D. H. Sonoda, *Universal rotation - How large can it be?*, MNRAS, 213, 917–943, 1985, `http://adsabs.harvard.edu//abs/1985MNRAS.213..917B`.

C. M. Baugh, *A primer on hierarchical galaxy formation: the semi-analytical approach*, Rept. Prog. Phys., 69, 3101–3156, 2006, `http://adsabs.harvard.edu//abs/2006RPPh...69.3101B`, arXiv:astro-ph/0610031.

C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, et al., *First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results*, ApJS, 148, 1–27, 2003, `http://adsabs.harvard.edu/abs/2003ApJS..148....1B`, arXiv:astro-ph/0302207.

A. J. Benson, *Galaxy Formation Theory*, Physics Reports, 495, 33–86, 2010, `http://adsabs.harvard.edu/abs/2010PhR...495...33B`, arXiv:1006.5394.

L. Bergstrom, *Dark Matter Candidates*, New Journal of Physics, 11, 105006, 2009, `http://adsabs.harvard.edu/abs/2009NJPh...11j5006B`, arXiv:0903.4849.

C. Blake and J. Wall, *Measurement of the angular correlation function of radio galaxies from the NRAO VLA Sky Survey*, MNRAS, 329, L37–L41, 2002, `http://adsabs.harvard.edu/abs/2002MNRAS.329L..37B`, arXiv:astro-ph/0111328.

J. R. Bond and G. Efstathiou, *Cosmic background radiation anisotropies in universes dominated by non-baryonic dark matter*, ApJ, 285, L45–L48, 1984, `http://adsabs.harvard.edu/abs/1984ApJ...285L..45B`.

J. R. Bond and G. Efstathiou, *The statistics of cosmic background radiation fluctuations*, MNRAS, 226, 655–687, 1987, `http://adsabs.harvard.edu/abs/1987MNRAS.226..655B`.

N. Bouche, A. Dekel, R. Genzel, S. Genel, G. Cresci, et al., *The Impact of cold gas accretion above a mass floor on galaxy scaling relations*, ApJ, 718, 1001–

1018, 2010, `http://adsabs.harvard.edu/abs/2010ApJ...718.1001B`, arXiv:0912.1858.

M. Bridges, J. D. McEwen, M. Cruz, M. P. Hobson, A. N. Lasenby, et al., *Bianchi VII$_h$ models and the cold spot texture*, MNRAS, 390, 1372–1376, 2008, `http://adsabs.harvard.edu/abs/2008MNRAS.390.1372B`, arXiv: 0712.1789.

J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, et al., *Profiles of dark haloes: evolution, scatter, and environment*, MNRAS, 321, 559–575, 2001, `http://adsabs.harvard.edu/abs/2001MNRAS.321..559B`, arXiv:astro-ph/9908159.

E. F. Bunn, P. Ferreira, and J. Silk, *How Anisotropic is our Universe?*, Phys. Rev. Lett., 77, 2883–2886, 1996, `http://adsabs.harvard.edu/abs/1996PhRvL..77.2883B`, arXiv:astro-ph/9605123.

S. M. Carroll, W. Press, and E. L. Turner, *The Cosmological Constant*, Annu. Rev. Astrophys., 30, 499–542, 1992, `http://adsabs.harvard.edu/abs/1992ARA%26A..30..499C`.

L.-Y. Chiang and P. Coles, *Phase Information and the Evolution of Cosmological Density Perturbations*, MNRAS, 311, 809–824, 2000, `http://adsabs.harvard.edu/abs/2000MNRAS.311..809C`, arXiv:astro-ph/9905250.

L.-Y. Chiang, P. Coles, P. D. Naselsky, and P. Olesen, *The one-dimensional Fourier Representation and Large Angular Scale Foreground Contamination in the 3-year Wilkinson Microwave Anisotropy Probe data*, JCAP, 1, 21, 2007a, `http://adsabs.harvard.edu/abs/2007JCAP...01..021C`, arXiv:astro-ph/0608421.

L.-Y. Chiang, P. Naselsky, and P. Coles, *Phase Mapping as a Powerful Diagnostic of Primordial Non-Gaussianity*, ApJ, 602, L1–L4, 2004, `http://adsabs.harvard.edu/abs/2004ApJ...602L...1C`, arXiv:astro-ph/0208235.

L.-Y. Chiang, P. Naselsky, and P. Coles, *Cosmic Covariance and the Low Quadrupole Anisotropy of the Wilkinson Microwave Anisotropy Probe (WMAP) Data*, ApJ, 694, 339–343, 2009, `http://adsabs.harvard.edu/abs/2009ApJ...694..339C`.

L.-Y. Chiang, P. D. Naselsky, and P. Coles, *Testing Gaussian random hypothesis with the cosmic microwave background temperature anisotropies in the three-year WMAP data*, ApJ, 664, 8–13, 2007b, `http://adsabs.harvard.edu/abs/2007ApJ...664....8C`, arXiv:astro-ph/0603662.

L.-Y. Chiang, P. D. Naselsky, O. V. Verkhodanov, and M. J. Way, *Non-Gaussianity of the derived maps from the first-year WMAP data*, ApJ, 590, L65–L68, 2003, `http://adsabs.harvard.edu/abs/2003ApJ...590L..65C`, arXiv:astro-ph/0303643.

S. Church, P. Ade, J. Bock, M. Bowden, J. Carlstrom, et al., *QUEST on DASI: a South Pole CMB polarization experiment*, New Astronomy Reviews, 47, 1083–1089, 2003, `http://adsabs.harvard.edu//abs/2003NewAR..47.1083C`.

S. Cole and C. Lacey, *The Structure of Dark Matter Haloes in Hierarchical Clustering Models*, MNRAS, 281, 716, 1996, `http://adsabs.harvard.edu/abs/1996MNRAS.281..716C`, arXiv:astro-ph/9510147.

P. Coles, *Galaxy formation with a local bias*, MNRAS, 262, 1065–1075, 1993, `http://adsabs.harvard.edu/abs/1993MNRAS.262.1065C`.

P. Coles and L.-Y. Chiang, *Characterizing the non-linear growth of large-scale structure in the Universe*, Nature, 406, 376–378, 2000, `http://adsabs.harvard.edu/abs/2000Natur.406..376C`, arXiv:astro-ph/0006017.

P. Coles, P. Dineen, J. Earl, and D. Wright, *Phase Correlations in Cosmic Microwave Background Temperature Maps*, MNRAS, 350, 989–1004, 2004, `http://adsabs.harvard.edu/abs/2004MNRAS.350..989C`, arXiv:astro-ph/0310252.

P. Coles and F. Lucchin, *Cosmology: The Origin and Evolution of Cosmic Structure*, Wiley, 2nd edition, 2002.

P. Coles, A. Melott, and D. Munshi, *Bias and Hierarchical Clustering*, ApJ, 521, L5–L8, 1999, `http://adsabs.harvard.edu/abs/1999ApJ...521L...5C`, arXiv:astro-ph/9904253.

M. Colless, G. B. Dalton, S. J. Maddox, W. J. Sutherland, P. Norberg, et al., *The 2dF Galaxy Redshift Survey: Spectra and redshifts*, MNRAS, 328, 1039–1063, 2001, `http://adsabs.harvard.edu/abs/2001MNRAS.328.1039C`, arXiv:astro-ph/0106498.

A. Connolly, R. Scranton, D. Johnston, S. Dodelson, D. Eisenstein, et al., *The Angular Correlation Function of Galaxies from Early SDSS Data*, ApJ, 579, 42–47, 2002, `http://adsabs.harvard.edu/abs/2002ApJ...579...42C`, arXiv:astro-ph/0107417.

A. Cooray, A. Amblard, L. Wang, B. Altieri, V. Arumugam, et al., *HerMES: Halo Occupation Number and Bias Properties of Dusty Galaxies from Angular Clustering Measurements*, A & A, 518, L22–L26, 2010, `http://adsabs.harvard.edu/abs/2010A%26A...518L..22C`, arXiv:1005.3303.

A. Cooray and R. Sheth, *Halo Models of Large Scale Structure*, Phys. Rept., 372, 1–129, 2002, `http://adsabs.harvard.edu/abs/2002PhR...372....1C`, arXiv:astro-ph/0206508.

C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman, *On the large-angle anomalies of the microwave sky*, MNRAS, 367, 79–102, 2006, `http://adsabs.harvard.edu/abs/2006MNRAS.367...79C`, arXiv:astro-ph/0508047.

C. J. Copi, D. Huterer, and G. D. Starkman, *Multipole Vectors–a new representation of the CMB sky and evidence for statistical anisotropy or non-Gaussianity at* $2 \leq \ell \leq 8$, Phys. Rev. D, 70, 043515, 2004, `http://adsabs.harvard.edu//abs/2004PhRvD..70d3515C`, arXiv:astro-ph/0310511.

C. M. Cress, D. J. Helfand, R. H. Becker, M. D. Gregg, and R. L. White, *The Angular Two-Point Correlation Function for the FIRST Radio Survey*, ApJ, 473, 7–14, 1996, `http://adsabs.harvard.edu/abs/1996ApJ...473....7C`, arXiv:astro-ph/9606176.

M. Cruz, E. Martinez-Gonzalez, P. Vielva, and L. Cayon, *Detection of a non-Gaussian Spot in WMAP*, MNRAS, 356, 29–40, 2005, `http://adsabs.harvard.edu/abs/2005MNRAS.356...29C`, arXiv:astro-ph/0405341.

M. Cruz, N. Turok, P. Vielva, E. Martinez-Gonzalez, and M. Hobson, *A Cosmic Microwave Background feature consistent with a cosmic texture*, Science, 318, 1612–1614, 2007, `http://adsabs.harvard.edu/abs/2007Sci...318.1612C`, arXiv:0710.5737.

S. Das and D. N. Spergel, *CMB Lensing and the WMAP Cold Spot*, Phys. Rev. D, 79, 043007, 2009, `http://adsabs.harvard.edu/abs/2009PhRvD..79d3007D`, arXiv:0809.4704.

P. de Bernardis, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, et al., *A Flat Universe from High-Resolution Maps of the Cosmic Microwave Background Radiation*, Nature, 404, 955–959, 2000, `http://adsabs.harvard.edu/abs/2000Natur.404..955D`, arXiv:astro-ph/0004404.

A. de Oliveira-Costa, M. Tegmark, M. Zaldarriaga, and A. Hamilton, *The significance of the largest scale CMB fluctuations in WMAP*, Phys. Rev. D, 69, 063516, 2004, `http://adsabs.harvard.edu/abs/2004PhRvD..69f3516D`, arXiv:astro-ph/0307282.

R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson, *Cosmic Black-Body Radiation.*, ApJ, 142, 414–419, 1965, `http://adsabs.harvard.edu/abs/1965ApJ...142..414D`.

J. M. Diego, M. Cruz, J. Gonzalez-Nuevo, M. Maris, Y. Ascasibar, et al., *WMAP anomalous signal in the ecliptic plane*, 2009, `http://arxiv.org/abs/0901.4344`, arXiv:0901.4344.

S. Dodelson, V. K. Narayanan, M. Tegmark, R. Scranton, T. Budavari, et al., *The 3D Power Spectrum from Angular Clustering of Galaxies in Early SDSS Data*, ApJ, 572, 140–156, 2002, `http://adsabs.harvard.edu/abs/2002ApJ...572..140D`, arXiv:astro-ph/0107421.

S. P. Driver, P. Norberg, I. K. Baldry, S. P. Bamford, A. M. Hopkins, et al., *GAMA: towards a physical understanding of galaxy formation*, Astron.Geophys., 50, 5.12–5.19, 2009, `http://adsabs.harvard.edu//abs/2009A%26G....50e..12D`, arXiv:0910.5123.

S. Dye, L. Dunne, S. Eales, D. J. B. Smith, A. Amblard, et al., *The Herschel-ATLAS: Evolution of the 250μm luminosity function out to z=0.5*, A & A, 518, L10–L14, 2010, `http://adsabs.harvard.edu//abs/2010A%26A...518L..10D`, arXiv:1005.2411.

S. Eales, 2011, private communication.

S. Eales, L. Dunne, D. Clements, A. Cooray, G. D. Zotti, et al., *The Herschel ATLAS*, PASP, 122, 499–515, 2010, `http://adsabs.harvard.edu/abs/2010PASP..122..499E`, arXiv:0910.4279.

D. J. Eisenstein and W. Hu, *Baryonic Features in the Matter Transfer Function*, ApJ, 496, 605–614, 1998, `http://adsabs.harvard.edu/abs/1998ApJ...496..605E`, arXiv:astro-ph/9709112.

D. J. Eisenstein and W. Hu, *Power Spectra for Cold Dark Matter and its Variants*, ApJ, 511, 5–15, 1999, `http://adsabs.harvard.edu/abs/1999ApJ...511....5E`, arXiv:astro-ph/9710252.

G. F. R. Ellis and M. A. H. MacCallum, *A class of homogeneous cosmological models*, Communications in Mathematical Physics, 12, 108–141, 1969, `http://adsabs.harvard.edu/abs/1969CMaPh..12..108E`.

H. K. Eriksen, A. J. Banday, K. M. Gorski, F. K. Hansen, and P. B. Lilje, *Hemispherical power asymmetry in the three-year Wilkinson Microwave Anisotropy Probe sky maps*, ApJ, 660, L81–L84, 2007, `http://adsabs.harvard.edu/abs/2007ApJ...660L..81E`, arXiv:astro-ph/0701089.

H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje, *On Foreground Removal from the Wilkinson Microwave Anisotropy Probe Data by an Internal Linear Combination Method: Limitations and Implications*, ApJ, 612, 633–646, 2004a, `http://adsabs.harvard.edu/abs/2004ApJ...612..633E`, arXiv:astro-ph/0403098.

H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje, *Astro-ph communication: Simulations of the WMAP Internal Linear Combination sky map*, 2005, `http://arxiv.org/pdf/astro-ph/0508196`, astro-ph/0508196.

H. K. Eriksen, F. K. Hansen, A. J. Banday, K. M. Gorski, and P. B. Lilje, *Asymmetries in the CMB anisotropy field*, ApJ, 605, 14–20, 2004b, `http://adsabs.harvard.edu/abs/2004ApJ...605...14E`, arXiv:astro-ph/0307507.

C. L. Francis and J. A. Peacock, *An estimate of the local ISW signal, and its impact on CMB anomalies*, MNRAS, 406, 14–21, 2010, `http://adsabs.harvard.edu//abs/2010MNRAS.406...14F`, arXiv:0909.2495.

W. L. Freedman, B. F. Madore, B. K. Gibson, L. Ferrarese, D. D. Kelson, et al., *Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant*, ApJ, 553, 47–72, 2001, `http://adsabs.harvard.edu/abs/2001ApJ...553...47F`, arXiv:astro-ph/0012376.

R. Gilli, G. Zamorani, T. Miyaji, J. Silverman, M. Brusa, et al., *The spatial clustering of X-ray selected AGN in the XMM-COSMOS field*, A & A, 494, 33–48, 2009, `http://adsabs.harvard.edu//abs/2009A%26A...494...33G`, arXiv:0810.4769.

K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, et al., *HEALPix – a Framework for High Resolution Discretization, and Fast Analysis of Data Distributed on the Sphere*, ApJ, 622, 759–771, 2005, `http://adsabs.harvard.edu/abs/2005ApJ...622..759G`, arXiv:astro-ph/0409513.

Q. Guo, S. Cole, C. G. Lacey, C. M. Baugh, C. S. Frenk, et al., *Which halos host Herschel-ATLAS galaxies in the local Universe?*, MNRAS, 412, 2277–2285, 2011, `http://adsabs.harvard.edu/abs/2011MNRAS.412.2277G`, arXiv:1011.3048.

A. H. Guth, *Inflationary universe: A possible solution to the horizon and flatness problems*, Phys. Rev. D, 23, 347–356, 1981, `http://adsabs.harvard.edu/abs/1981PhRvD..23..347G`.

A. H. Guth and S. Y. Pi, *Fluctuations in the new inflationary universe*, Phys. Rev. Lett., 49, 1110–1113, 1982, `http://adsabs.harvard.edu/abs/1982PhRvL..49.1110G`.

A. J. S. Hamilton, P. Kumar, E. Lu, and A. Matthews, *Reconstructing the primordial spectrum of fluctuations of the universe from the observed nonlinear clustering of galaxies*, ApJ, 374, L1–L4, 1991, `http://adsabs.harvard.edu//abs/1991ApJ...374L...1H`.

S. Hanany, P. Ade, A. Balbi, J. Bock, J. Borrill, et al., *MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on angular scales of 10 arcminutes to 5 degrees*, ApJ, 545, L5–L9, 2000, `http://adsabs.harvard.edu//abs/2000ApJ...545L...5H`, arXiv:astro-ph/0005123.

F. K. Hansen, A. J. Banday, K. M. Gorski, H. K. Eriksen, and P. B. Lilje, *Power Asymmetry in Cosmic Microwave Background Fluctuations from Full Sky to Sub-degree Scales: Is the Universe Isotropic?*, ApJ, 704, 1448–1458, 2009, `http://adsabs.harvard.edu/abs/2009ApJ...704.1448H`, arXiv:0812.3795.

M. Hansen, J. Short, P. Naselsky, P. Coles, J. Kim, et al., *Multipole cross-correlations and the Statistical Anisotropy of the Cosmic Microwave Background*, 2011, in preparation.

W. Hartley, O. Almaini, M. Cirasuolo, S. Foucaud, C. Simpson, et al., *The evolution of galaxy clustering since $z = 3$ using the UKIDSS Ultra Deep Survey: the divergence of passive and star-forming galaxies*, MNRAS, 407, 1212–1222, 2010, `http://adsabs.harvard.edu//abs/2010MNRAS.407.1212H`, arXiv:1005.1180.

G. Hasinger, N. Cappelluti, H. Brunner, M. Brusa, A. Comastri, et al., *The XMM-Newton wide-field survey in the COSMOS field: I. Survey description*, ApJS, 172, 29–37, 2007, `http://adsabs.harvard.edu/abs/2007ApJS..172...29H`, astro-ph/0612311.

G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, et al., *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results*, ApJS, 180, 225–245, 2009, `http://adsabs.harvard.edu/abs/2009ApJS..180..225H`, arXiv:0803.0732.

M. P. Hobson, A. W. Jones, and A. N. Lasenby, *Wavelet analysis and the detection of non-Gaussianity in the CMB*, MNRAS, 309, 125–140, 1999, `http://adsabs.harvard.edu/abs/1999MNRAS.309..125H`, arXiv:astro-ph/9810200.

J. Hoftuft, H. K. Eriksen, A. J. Banday, K. M. Gorski, F. K. Hansen, et al., *Increasing evidence for hemispherical power asymmetry in the five-year WMAP data*, ApJ, 699, 985–989, 2009, `http://adsabs.harvard.edu/abs/2009ApJ...699..985H`, arXiv:0903.1229.

W. Hu and N. Sugiyama, *Small Scale Cosmological Perturbations: An Analytic Approach*, ApJ, 471, 542–570, 1996, `http://adsabs.harvard.edu/abs/1996ApJ...471..542H`, arXiv:astro-ph/9510117.

E. Hubble, *A relation between distance and radial velocity among extra-galactic nebulae*, PNAS, 15, 168–173, 1929, `http://adsabs.harvard.edu/abs/1929PNAS...15..168H`.

D. Hughes, S. Serjeant, J. Dunlop, M. Rowan-Robinson, A. Blain, et al., *High-redshift star formation in the Hubble Deep Field revealed by a submillimetre-wavelength survey*, Nature, 394, 241–247, 1998, `http://adsabs.harvard.edu/abs/1998Natur.394..241H`, arXiv:astro-ph/9806297.

K. T. Inoue and J. Silk, *Local Voids as the Origin of Large-angle Cosmic Microwave Background Anomalies I*, ApJ, 648, 23–30, 2006, `http://adsabs.harvard.edu/abs/2006ApJ...648...23I`, arXiv:astro-ph/0602478.

T. R. Jaffe, A. J. Banday, H. K. Eriksen, K. M. Gorski, and F. K. Hansen, *Evidence of vorticity and shear at large angular scales in the WMAP data: a violation of cosmological isotropy?*, ApJ, 629, L1–L4, 2005, `http://adsabs.harvard.edu/abs/2005ApJ...629L...1J`, arXiv:astro-ph/0503213.

T. R. Jaffe, A. J. Banday, H. K. Eriksen, K. M. Gorski, and F. K. Hansen, *Bianchi Type VII_h Models and the WMAP 3-year Data*, A & A, 460, 393–396, 2006a, `http://adsabs.harvard.edu/abs/2006A26A...460..393J`, arXiv:astro-ph/0606046.

T. R. Jaffe, S. Hervik, A. J. Banday, and K. M. Gorski, *On the Viability of Bianchi Type VII_h Models with Dark Energy*, ApJ, 644, 701–708, 2006b, `http://adsabs.harvard.edu/abs/2006ApJ...644..701J`, arXiv:astro-ph/0512433.

N. Jarosik, C. Barnes, M. R. Greason, R. S. Hill, M. R. Nolta, et al., *Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Beam Profiles, Data Processing, Radiometer Characterization and Systematic Error Limits*, ApJS, 170, 263–287, 2007, `http://adsabs.harvard.edu//abs/2007ApJS..170..263J`, arXiv:astro-ph/0603452.

N. Jarosik, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, et al., *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results*, ApJS, 192, 14, 2011, `http://adsabs.harvard.edu//abs/2011ApJS..192...14J`, arXiv:1001.4744.

M. J. Jarvis, D. J. B. Smith, D. G. Bonfield, M. J. Hardcastle, J. T. Falder, et al., *Herschel-ATLAS: The far-infrared–radio correlation at z<0.5*, MNRAS, 409, 92–101, 2010, `http://adsabs.harvard.edu/abs/2010MNRAS.409...92J`, arXiv:1009.5390.

J. H. Jeans, *The Stability of a Spherical Nebula*, Philosophical Transactions of the Royal Society of London, 199, 1–53, 1902, `http://www.jstor.org/pss/90845`.

N. Kaiser, *On the spatial correlations of Abell clusters*, ApJ, 284, L9–L12, 1984, `http://www.adsabs.harvard.edu/abs/1984ApJ...284L...9K`.

G. Katz and J. Weeks, *Polynomial Interpretation of Multipole Vectors*, Phys. Rev. D, 70, 063527, 2004, `http://adsabs.harvard.edu//abs/2004PhRvD..70f3527K`, arXiv:astro-ph/0405631.

J. Kim, P. Naselsky, and P. R. Christensen, *CMB map derived from the WMAP data through Harmonic Internal Linear Combination*, Phys. Rev. D, 77, 103002, 2008, `http://adsabs.harvard.edu/abs/2008PhRvD..77j3002K`, arXiv:0803.1394.

A. Kogut, G. Hinshaw, and A. J. Banday, *Limits to Global Rotation and Shear From the COBE DMR 4-Year Sky Maps*, Phys. Rev. D, 55, 1901–1905, 1997, `http://adsabs.harvard.edu//abs/1997PhRvD..55.1901K`, arXiv:astro-ph/9701090.

E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, et al., *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, ApJS, 180, 330–376, 2009, `http://adsabs.harvard.edu/abs/2009ApJS..180..330K`, arXiv:0803.0547.

E. Komatsu, A. Kogut, M. Nolta, C. L. Bennett, M. Halpern, et al., *First Year Wilkinson Microwave Anisotropy Probe (WMAP)*

Observations: Tests of Gaussianity, ApJS, 148, 119–134, 2003, `http://adsabs.harvard.edu/abs/2003ApJS..148..119K`, arXiv:astro-ph/0302223.

M. Krumpe, T. Miyaji, and A. L. Coil, *The Spatial Clustering of ROSAT All-Sky Survey AGN: I. The cross-correlation function with SDSS Luminous Red Galaxies*, ApJ, 713, 558–572, 2010, `http://adsabs.harvard.edu//abs/2010ApJ...713..558K`, arXiv:1002.3598.

K. Land and J. Magueijo, *The Axis of Evil*, Phys. Rev. Lett., 95, 071301, 2005, `http://adsabs.harvard.edu/abs/2005PhRvL..95g1301L`, arxiv:astro-ph/0502237.

S. D. Landy and A. S. Szalay, *Bias and variance of angular correlation functions*, ApJ, 412, 64–71, 1993, `http://adsabs.harvard.edu/abs/1993ApJ...412...64L`.

D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, M. R. Nolta, et al., *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-Derived Parameters*, ApJS, 192, 16, 2011, `http://adsabs.harvard.edu/abs/2011ApJS..192...16L`, arXiv: 1001.4635.

D. L. Larson and B. D. Wandelt, *The Hot and Cold Spots in the WMAP Data are Not Hot and Cold Enough*, ApJ, 613, L85–L88, 2004, `http://adsabs.harvard.edu/abs/2004ApJ...613L..85L`, arXiv:astro-ph/0404037.

A. Lawrence, S. J. Warren, O. Almaini, A. C. Edge, N. C. Hambly, et al., *The UKIRT Infrared Deep Sky Survey (UKIDSS)*, MNRAS, 379, 1599–1617, 2007, `http://adsabs.harvard.edu/abs/2007MNRAS.379.1599L`, arXiv:astro-ph/0604426.

M. Limon, *Wilkinson Microwave Anisotropy Probe (WMAP): FiveYear Explanatory Supplement*, 2009, `http://lambda.gsfc.nasa.gov/product/map/dr3/pub_papers/fiveyear/supplement/WMAP_supplement` (Greenbelt, MD: NASA/GSFC).

A. Linde, *Inflationary Cosmology*, Lect. Notes Phys., 738, 1–54, 2008, `http://adsabs.harvard.edu/abs/2008LNP...738....1L`, arXiv: 0705.0164.

M. Lopez-Corredoira, *Some doubts on the validity of the foreground Galactic contribution subtraction from microwave anisotropies*, J. Astrophys. Astr., 28, 101–116, 2007, `http://adsabs.harvard.edu/abs/2007JApA...28..101L`, arXiv:0708.4133.

S. J. Maddox, L. Dunne, E. Rigby, S. Eales, A. Cooray, et al., *Herschel ATLAS: The angular correlation function of submillimetre galaxies at high and low redshift*, A & A, 518, L11–L15, 2010, `http://adsabs.harvard.edu//abs/2010A%26A...518L..11M`, arXiv:1005.2406.

F. Marin, *The Large-Scale 3-point correlation function of SDSS Luminous Red Galaxies*, 2010, `http://arxiv.org/abs/1011.4530`, arXiv:1011.4530.

E. Martinez-Gonzalez, *Gaussianity*, 2008, `http://arxiv.org/abs/0805.4157`, arXiv:0805.4157.

I. Masina and A. Notari, *The Cold Spot as a Large Void: Lensing Effect on CMB Two and Three Point Correlation Functions*, JCAP, 07, 35, 2009, `http://adsabs.harvard.edu/abs/2009JCAP...07..035M`, arXiv:0905.1073.

C. K. McBride, A. J. Connolly, J. P. Gardner, R. Scranton, J. A. Newman, et al., *Three-Point Correlation Functions of SDSS Galaxies: Luminosity and Color Dependence in Redshift and Projected Space*, ApJ, 726, 13, 2010, `http://adsabs.harvard.edu/abs/2011ApJ...726...13M`, arXiv: 1007.2414.

J. D. McEwen, M. P. Hobson, A. N. Lasenby, and D. J. Mortlock, *Non-Gaussianity detections in the Bianchi VII_h corrected WMAP 1-year data made with directional spherical wavelets*, MNRAS, 369, 1858–1868, 2006, `http://adsabs.harvard.edu/abs/2006MNRAS.369.1858M`, arXiv:astro-ph/0510349.

T. Miyaji, M. Krumpe, A. L. Coil, and H. Aceves, *The Spatial Clustering of ROSAT All-Sky Survey AGNs II. Halo Occupation Distribution Modeling of the Cross Correlation Function*, ApJ, 726, 83, 2011, `http://adsabs.harvard.edu//abs/2011ApJ...726...83M`, arXiv:1010.5498.

H. J. Mo and S. D. M. White, *An analytic model for the spatial clustering of dark matter haloes*, MNRAS, 282, 347–361, 1996, `http://adsabs.harvard.edu/abs/1996MNRAS.282..347M`, arXiv:astro-ph/9512127.

C. Monteserin, R. B. Barreiro, P. Vielva, E. Martinez-Gonzalez, M. P. Hobson, et al., *A low CMB variance in the WMAP data*, MNRAS, 387, 209–219, 2008, `http://adsabs.harvard.edu/abs/2008MNRAS.387..209M`, arXiv:0706.4289.

L. Moscardini, P. Coles, F. Lucchin, and S. Matarrese, *Modelling galaxy clustering at high redshift*, MNRAS, 299, 95–110, 1998, `http://adsabs.harvard.edu/abs/1998MNRAS.299...95M`, arXiv:astro-ph/9712184.

C. R. Mullis, J. P. Henry, I. M. Gioia, H. Boehringer, U. G. Briel, et al., *Spatial Correlation Function of X-ray Selected AGN*, ApJ, 617, 192–208, 2004, `http://adsabs.harvard.edu/abs/2004ApJ...617..192M`, arXiv:astro-ph/0408304.

P. D. Naselsky and O. V. Verkhodanov, *Peculiarities of phases of the WMAP quadrupole*, International Journal of Modern Physics D, 17, 179–194, 2008, `http://adsabs.harvard.edu/abs/2008IJMPD..17..179N`, arXiv:astro-ph/0609409.

P. D. Naselsky, O. V. Verkhodanov, and M. T. B. Nielsen, *Instability of reconstruction of the low CMB multipoles*, Astrophysical Bulletin, 63, 216–227, 2008, `http://adsabs.harvard.edu//abs/2008AstBu..63..216N`, arXiv:0707.1484.

J. F. Navarro, C. S. Frenk, and S. D. M. White, *A Universal Density Profile from Hierarchical Clustering*, ApJ, 490, 493–508, 1997, `http://adsabs.harvard.edu//abs/1997ApJ...490..493N`, arXiv:astro-ph/9611107.

H. T. Nguyen, B. Schulz, L. Levenson, A. Amblard, V. Arumugam, et al., *HerMES: The SPIRE confusion limit*, A & A, 518, L5, 2010, `http://adsabs.harvard.edu/abs/2010A%26A...518L...5N`, arXiv:1005.2207.

S. J. Oliver, L. Wang, A. J. Smith, B. Altieri, A. Amblard, et al., *HerMES: SPIRE galaxy number counts at 250, 350 and 500 microns*, A & A, 518, L21, 2010, `http://adsabs.harvard.edu/abs/2010A%26A...518L..21O`, arXiv:1005.2184.

C.-G. Park, *Non-Gaussian Signatures in the Temperature Fluctuation Observed by the Wilkinson Microwave Anisotropy Probe*, MNRAS, 349, 313–320, 2004, `http://adsabs.harvard.edu/abs/2004MNRAS.349..313P`, arXiv:astro-ph/0307469.

J. A. Peacock, *Cosmological Physics*, Cambridge University Press, 1999.

J. A. Peacock and S. J. Dodds, *Nonlinear evolution of cosmological power spectra*, MNRAS, 280, L19–L26, 1996, `http://adsabs.harvard.edu/abs/1996MNRAS.280L..19P`, arXiv:astro-ph/9603031.

P. J. E. Peebles, *The large-scale structure of the universe*, Princeton University Press, 1980.

H. V. Peiris, E. Komatsu, L. Verde, D. N. Spergel, C. L. Bennett, et al., *First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Inflation*, ApJS, 148, 213–231, 2003, `http://adsabs.harvard.edu/abs/2003ApJS..148..213P`, arXiv:astro-ph/0302225.

A. A. Penzias and R. W. Wilson, *A Measurement of Excess Antenna Temperature at 4080 Mc/s*, ApJ, 142, 419–421, 1965, `http://adsabs.harvard.edu//abs/1965ApJ...142..419P`.

W. J. Percival, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, et al., *The 2dF Galaxy Redshift Survey: The power spectrum and the matter content of the universe*, MNRAS, 327, 1297–1306, 2001, `http://adsabs.harvard.edu//abs/2001MNRAS.327.1297P`, arXiv:astro-ph/0105252.

W. J. Percival, R. C. Nichol, D. J. Eisenstein, D. H. Weinberg, M. Fukugita, et al., *Measuring the matter density using baryon oscillations in the SDSS*, ApJ, 657, 51–55, 2007, `http://adsabs.harvard.edu/abs/2007ApJ...657...51P`, arXiv:astro-ph/0608635.

S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, et al., *Measurements of Omega and Lambda from 42 High-Redshift Supernovae*, ApJ, 517, 565–586, 1999, `http://adsabs.harvard.edu/abs/1999ApJ...517..565P`, arXiv:astro-ph/9812133.

G. L. Pilbratt, J. R. Riedinger, T. Passvogel, G. Crone, D. Doyle, et al., *Herschel Space Observatory - An ESA facility for far-infrared and submillimetre astronomy*, A & A, 518, L1, 2010, `http://adsabs.harvard.edu/abs/2010A%26A...518L...1P`, arXiv:1005.5331.

A. Pontzen, *Rogues' gallery: the full freedom of the Bianchi CMB anomalies*, Phys. Rev. D, 79, 103518, 2009, `http://adsabs.harvard.edu/abs/2009PhRvD..79j3518P`, arXiv:0901.2122.

A. Pontzen and A. Challinor, *Bianchi Model CMB Polarization and its Implications for CMB Anomalies*, MNRAS, 380, 1387–1398, 2007, `http://adsabs.harvard.edu/abs/2007MNRAS.380.1387P`, arXiv:0706.2075.

W. H. Press and P. Schechter, *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*, ApJ, 187, 425–438, 1974, `http://adsabs.harvard.edu//abs/1974ApJ...187..425P`.

D. Reed, R. Bower, C. Frenk, A. Jenkins, and T. Theuns, *The halo mass function from the dark ages through the present day*, MNRAS, 374, 2–15, 2007, `http://adsabs.harvard.edu/abs/2007MNRAS.374....2R`, arXiv:astro-ph/0607150.

A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiattia, A. Diercks, et al., *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, AJ, 116, 1009–1038, 1998, `http://adsabs.harvard.edu/abs/1998AJ....116.1009R`, arXiv:astro-ph/9805201.

A. J. Ross and R. J. Brunner, *Halo-model Analysis of the Clustering of Photometrically Selected Galaxies from SDSS*, MNRAS, 399, 878–887, 2009, `http://adsabs.harvard.edu/abs/2009MNRAS.399..878R`, arXiv:0906.4977.

R. K. Sachs and A. M. Wolfe, *Perturbations of a Cosmological Model and Angular Variations of the Microwave Background*, ApJ, 147, 73–90, 1967, `http://adsabs.harvard.edu/abs/1967ApJ...147...73S`.

R. K. Sheth and G. Tormen, *Large scale bias and the peak background split*, MNRAS, 308, 119–126, 1999, `http://adsabs.harvard.edu/abs/1999MNRAS.308..119S`, arXiv:astro-ph/9901122.

J. Short and P. Coles, *Zonal Modes of Cosmic Microwave Background Temperature Maps*, MNRAS, 401, 2202–2206, 2010, `http://adsabs.harvard.edu/abs/2010MNRAS.401.2202S`, arXiv:0907.4011.

J. Short and P. Coles, *Models for the Clustering of Far-Infrared and Sub-millimetre selected Galaxies*, MNRAS, 412, L93–L97, 2011, `http://adsabs.harvard.edu/abs/2011MNRAS.412L..93S`, arXiv:1101.1862.

J. Silk, *Fluctuations in the Primordial Fireball*, Nature, 215, 1155–1156, 1967, `http://adsabs.harvard.edu/abs/1967Natur.215.1155S`.

G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, et al., *Structure in the COBE differential microwave radiometer first-year maps*, ApJ, 396, L1–L5, 1992, `http://adsabs.harvard.edu//abs/1992ApJ...396L...1S`.

G. F. Smoot and D. Scott, *The Cosmic Background Radiation*, 1996, `http://arxiv.org/abs/astro-ph/9603157`, arXiv:astro-ph/9603157.

D. N. Spergel, R. Bean, O. Dor, M. R. Nolta, C. L. Bennett, et al., *Wilkinson Microwave Anisotropy Probe (WMAP) Three Year Results: Implications for Cosmology*, ApJS, 170, 377–408, 2007, `http://adsabs.harvard.edu//abs/2007ApJS..170..377S`, arXiv:astro-ph/0603449.

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, et al., *Simulating the joint evolution of quasars, galaxies and their large-scale distribution*, Nature, 435, 629–636, 2005, `http://adsabs.harvard.edu/abs/2005Natur.435..629S`, arXiv:astro-ph/0504097.

A. Stannard and P. Coles, *Random-Walk Statistics and the Spherical Harmonic Representation of CMB Maps*, MNRAS, 364, 929–933, 2005, `http://adsabs.harvard.edu/abs/2005MNRAS.364..929S`, arXiv:astro-ph/0410633.

R. Sung and P. Coles, *Polarized Spots in Anisotropic Open Universes*, CQGra, 26, 172001, 2009, `http://adsabs.harvard.edu/abs/2009CQGra..26q2001S`, arXiv: 0905.2307.

R. Sung and P. Coles, *Temperature and Polarization Patterns in Anisotropic Cosmologies*, JCAP, 6, 36, 2011, `http://adsabs.harvard.edu/abs/2011JCAP...06..036S`, arXiv: 1004.0957.

R. Sung, J. Short, and P. Coles, *Statistical Characterization of Temperature Patterns in Anisotropic Cosmologies*, MNRAS, 412, 492502, 2010, `http://adsabs.harvard.edu/abs/2011MNRAS.412..492S`, arXiv: 1004.1925.

J. A. Tauber, N. Mandolesi, J.-L. Puget, T. Banos, M. Bersanelli, et al., *Planck pre-launch status: The Planck mission*, A & A, 520, A1–A22, 2010, `http://adsabs.harvard.edu/abs/2010A%26A...520A...1T`.

M. Tegmark, A. de Oliveira-Costa, and A. Hamilton, *A high resolution foreground cleaned CMB map from WMAP*, Phys. Rev. D, 68, 123523, 2003, `http://adsabs.harvard.edu/abs/2003PhRvD..68l3523T`, arXiv:astro-ph/0302496.

J. L. Tinker, D. H. Weinberg, Z. Zheng, and I. Zehavi, *On the Mass-to-Light Ratio of Large Scale Structure*, ApJ, 631, 41–58, 2005, `http://adsabs.harvard.edu/abs/2005ApJ...631...41T`, arXiv:astro-ph/0411777.

L. Verde, A. F. Heavens, and S. Matarrese, *Projected bispectrum in spherical harmonics and its application to angular galaxy catalogues*, MNRAS, 318, 584–598, 2000, `http://adsabs.harvard.edu/abs/2000MNRAS.318..584V`, arXiv:astro-ph/0002240.

P. Vielva, E. Martinez-Gonzalez, R. B. Barreiro, J. L. Sanz, and L. Cayon, *Detection of non-Gaussianity in the WMAP 1-year data using spherical wavelets*, ApJ, 609, 22–34, 2004, `http://adsabs.harvard.edu/abs/2004ApJ...609...22V`, arXiv:astro-ph/0310273.

M. P. Viero, P. A. R. Ade, J. J. Bock, E. L. Chapin, M. J. Devlin, et al., *BLAST: Correlations in the Cosmic Far-Infrared Background at 250, 350, and 500 microns Reveal Clustering of Star-Forming Galaxies*, ApJ, 707, 1766–1778, 2009, `http://adsabs.harvard.edu/abs/2009ApJ...707.1766V`, arXiv:0904.1200.

P. Watts and P. Coles, *Statistical Cosmology with Quadratic Density Fields*, MNRAS, 338, 806, 2003, `http://adsabs.harvard.edu/abs/2003MNRAS.338..806W`, arXiv:astro-ph/0208295.

A. P. S. Yadav and B. D. Wandelt, *Detection of primordial non-Gaussianity ($f_{\rm NL}$) in the WMAP 3-year data at above 99.5% confidence*, Phys. Rev. Lett., 100, 181301, 2008, `http://adsabs.harvard.edu/abs/2008PhRvL.100r1301Y`, arXiv:0712.1148.

D. G. York, *The Sloan Digital Sky Survey: Technical Summary*, AJ, 120, 1579–

1587, 2000, `http://adsabs.harvard.edu/abs/2000AJ....120.1579Y`, arXiv:astro-ph/0006396.

Z. Zheng, A. A. Berlind, D. H. Weinberg, A. J. Benson, C. M. Baugh, et al., *Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies*, ApJ, 633, 791–809, 2005, `http://adsabs.harvard.edu/abs/2005ApJ...633..791Z`, arXiv:astro-ph/0408564.